Ecole Doctorale d'Astronomie et
Astrophysique d'Île de France

THESE DE DOCTORAT

présentée pour obtenir le titre de :
Docteur de l'Observatoire de Paris
Spécialité Astronomie et Astrophysique

par

Florian PACAUD

# Exploitation cosmologique du relevé XMM-LSS

soutenue le 19/02/2008 devant le jury composé de :

| | |
|---|---|
| Pr. James BARTLETT | Président |
| Dr. William FORMAN | Rapporteur |
| Dr. Jean-Paul KNEIB | Rapporteur |
| Dr. Nabila AGHANIM | Examinatrice |
| Pr. Alfonso CAVALIERE | Examinateur |
| Dr. Marguerite PIERRE | Directrice de thèse |

Thèse préparée au sein du :
Service d'Astrophysique
DSM/DAPNIA/CEA-Saclay

# Remerciements

Une fois qu'un travail de thèse est terminé, que les dernières lignes viennent d'être écrites, le moment arrive enfin où l'on peut s'arrêter un moment, regarder en arrière et faire le bilan de quelques longues années.

En cet instant particulier, il devient alors possible d'apprécier à quel point le travail d'apprentissage et de recherche accompli dépend de dialogues, de rencontres et d'intéractions humaines. C'est, j'imagine, le moment où chacun entreprend d'écrire sa page de remerciements.

Comme beaucoup, je dois une profonde reconnaissance à ma directrice de thèse. Merci surtout, Marguerite, pour ton investissement quotidien dans mes travaux de thèse et pour ton attention.

Je suis également débiteur d'Ana Gomez pour m'avoir donné ma chance en astronomie, à une époque où je n'y croyais plus vraiment.

Un grand merci à tous les membres de mon jury pour avoir accepté d'en faire partie : être jugé par des gens que l'on estime constitue déjà une récompense. Merci à mes rapporteurs Bill Forman et Jean-Paul Kneib pour leur investissement plus poussé, ainsi qu'à Jim Barlett pour avoir présidé le jury. Merci à Alfonso Cavaliere pour s'être déplacé depuis l'Italie et tout particulièrement à Nabila Aghanim pour son intérêt pour ma thèse malgré son congé maternité.

Tous mes remerciements à l'ensemble des employés du SAp où j'ai effectué mes travaux. Merci à Pierre-Olivier Lagage et à David Elbaz pour m'avoir accueilli dans le laboratoire 'cosmologie et évolution des galaxies'. Merci à Alexandre Réfregier pour son investissement lors de mes débuts. Merci aussi à tous ceux qui ont tenté de me faire partager leur savoir, entre autres : Jean Ballet, J.-L. Sauvageot, Monique Arnaud, J.-L. Stark, et Pierre-Alain Duc.

Merci à Ivan Valtchanov et Sergio Dos Santos pour leurs enseignements utiles.

Merci à Jean-Paul Lefèvre pour son initiation au fonctionnement des bases de données.

Le travail que je présente dans cette thèse est le fruit d'une large collaboration internationale. Je remercie donc tous les membres du consortium XMM-LSS avec qui mes interactions furent fructueuses et en particulier Trevor Ponman, Stefano Andreon, Jon willis, Danielle Alloin et Jean Surdej. Un clin d'oeil spécial également à tous les jeunes de la collaboration : Abdul, Chris, Cyril, Jean-Baptiste, Joël, Olivier, Pierre-Guillaume

Je voudrais aussi remercier mes compagnons de bureau durant cette longue épreuve : Samuel, Renaud, Joël (encore, mais il le mérite) et Ivan Debono. Enfin

merci pour l'ambiance à tous les thésards du labo, avec une pensée particulière pour Yann, Matthias (avec deux t), Sandrine et Delphine.

Merci à Frank Bertoldi de m'avoir offert un contrat à l'université de Bonn avant ma soutenance.

D'une façon générale ce volume est dédicacé à tous les astrophysiciens et astrophysiciennes que j'ai croisés et avec qui mes rapports furent aussi divers qu'enrichissants

Une reconnaissance éternelle à ma famille pour avoir fait de moi ce que je suis et à ma belle-famille pour y avoir contribué.

Merci enfin et surtout à Elise, l'étoile de mes jours, pour son soutien et son abnégation, ainsi qu'à mes deux merveilleux petits garçons qui rendent la vie tellement plus intéressante.

# Table des matières

# Première partie:

## Exploitation cosmologique du relevé XMM Large Scale Structure

# Chapitre 1
# Introduction

## Sommaire



## 1.1 Le modèle de concordance

La cosmologie, en tant que discipline scientifique, a une histoire relativement jeune comparée à d'autres domaines de la physique. Moins d'un siècle s'est écoulé depuis les premiers sursauts de l'astronomie extragalactique : la découverte des premiers systèmes stellaires en dehors de la Voie Lactée (Hubble, 1925) et l'expansion de la toile cosmique Hubble (1929). A l'exception notable de la découverte de la matière noire (Zwicky 1933, puis plus tard Roberts 1976 et Rubin et al. 1978) et de la détection du fond diffus cosmologique (Penzias & Wilson, 1965; Dicke et al., 1965), les 50 années suivantes furent essentiellement marquées par des avancées sur un plan purement théorique.

A l'opposé, les inovations technologiques récentes (CCD, grands instruments, cryogénie, missions spatiales et calculs numériques) ont permis d'obtenir sur les 20 dernières années un ensemble de contraintes observationnelles fortes, et constitue une période très stimulante. La quasi totalité des observations pointent aujourd'hui vers un modèle auto-cohérent qui permet d'expliquer la plupart des données : fluctuations du fond diffus cosmologique (Tegmark, 1996; Benoît et al., 2003), densité et distribution des galaxies (Eke et al., 2006; Percival et al., 2007) et des amas de galaxies (Borgani et al., 2001; Schuecker et al., 2003), cisaillement gravitationnel cosmique (Benjamin et al., 2007), supernovae (Perlmutter et al., 1999; Astier et al., 2006), ...

Cette série de résultats a vu son apogée avec les mesures du fond diffus cosmologique par le satellite américain WMAP (Spergel et al., 2007). A elle seule, cette expérience permet de mesurer la plupart des paramêtres cosmologiques avec une erreur meilleure que 10%.

Ce modèle de concordance, comme il est d'usage de l'appeler, recèle pourtant

nombre de surprises. Il postule que la distribution initiale de masse dans l'univers primordial soit homogène à mieux que 0.001% près sur des échelles dépassant la zone de causalité. Il prédit un Univers exactement plat dont seul 5% de la matière existante se compose de baryons. Pire encore, il faut admettre, afin de reproduire toutes les observations simultanément, que les 95% restant comprennent deux types de matériaux exotiques distincts : une forme d'énergie associée à une pression négative qui génére une gravité répulsive (l'énergie sombre, comparable à la constante cosmologique d'Einstein), et une composante massive non-colisionnelle n'intéragissant pas avec le champs électromagnétique (matière noire).

A mesure que les contraintes s'affinent, certaines incompatibilités apparaissent également entre les observables, notamment dans l'estimation de l'amplitude des fluctuations du champ de densité au temps présent. Résoudre ces dernieres incohérences et comprendre la nature de l'énergie sombre et celle de la matière noire consitue un nouveau défi pour notre entendement. Il faudra pour cela atteindre une finesse encore supérieure dans les mesures de l'évolution de l'Univers et de son contenu, ainsi qu'un contrôle minutieux des erreurs systématiques propres à chaque méthode.

La discipline entre donc dans une nouvelle ère : celle de la 'cosmologie de précision'.

## 1.2 Amas de galaxies et cosmologie

Les amas de galaxies furent initialement identifiés comme des agglomérations comprenant de quelques dizaines à près de 1000 galaxies. Il s'agit en fait de structures virialisées dont la masse totale est comprise entre $10^{14}$ et quelques $10^{15}$ masses solaires. Contrairement, à ce que leur nom laisse entendre, les galaxies ne représentent que quelques pourcents de leur masse total. Une image plus réaliste consisterait à les définir comme des surdensités de matière noire, associées à une composante de gaz diffus ($\sim 15\%$ de la masse) dans lesquelles les galaxies tendent à se regrouper. La Figure 1.1 illustre ce propos.

Dans le cadre du modèle hiérarchique de formation des structures, les fluctuations initiales de densité croissent par accrétion de la matière environante jusqu'à former des objets découplés de l'expansion cosmologique. Ceux-ci fusionnent entre eux pour constituer des structures de plus en plus massives tandis que l'Univers veillit. A l'époque actuelle, les amas de galaxies sont les objets les plus massifs que l'on connaisse à avoir atteint leur êtat d'équilibre propre. Pour les études cosmologiques, ils ont donc le double avantage d'être des sytèmes relativement simples (dans la mesure où les processus physiques non-linéaires ne dominent pas encore leur modèlisation), tout en fournissant une information directe sur le taux de croissance des structures dans l'Univers actuel. La fonction de masse des amas dans l'univers local est ainsi paticulièrement sensible à la densité moyenne de masse $\Omega_m$. Connaître précisemment son évolution permettrait de différencier les modèles possibles d'énergie sombre.

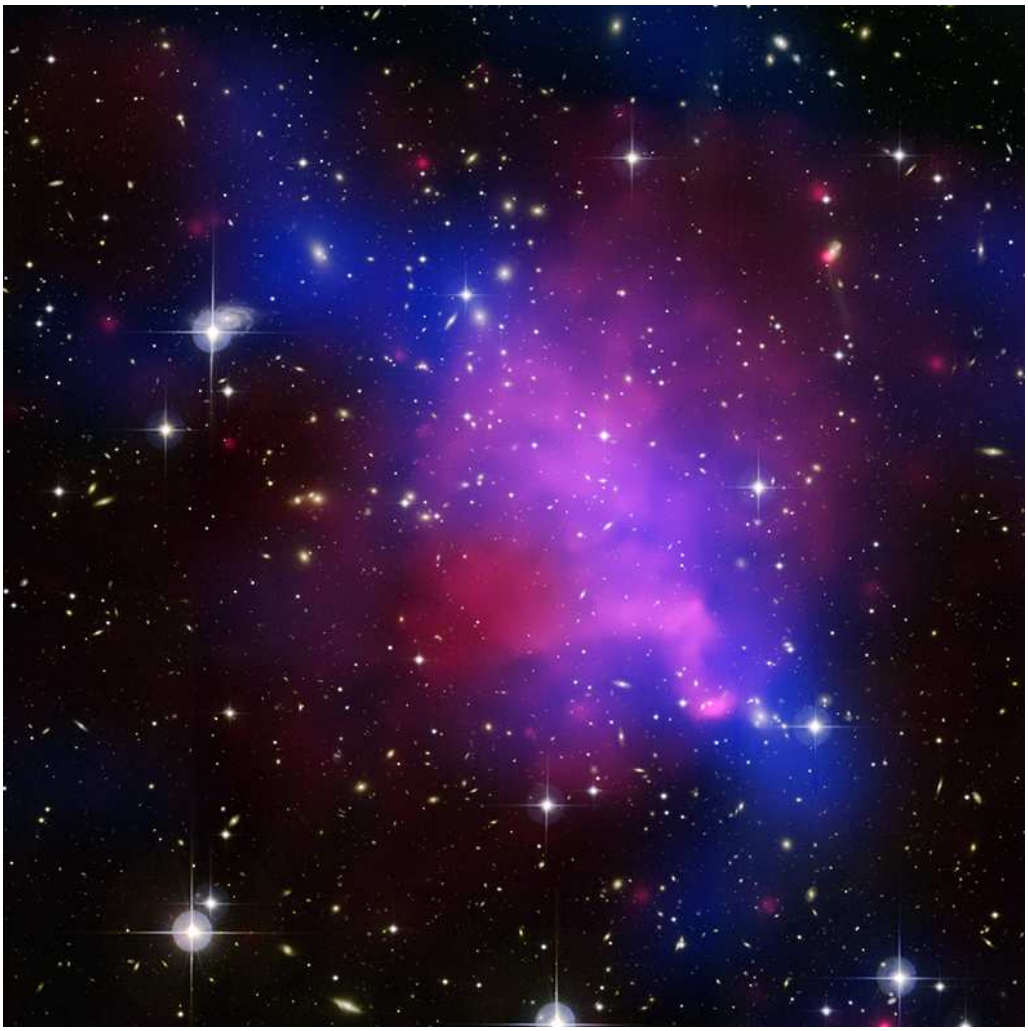

**Fig. 1.1**: *L'amas de galaxies Abell 520. Les galaxies montrées en jaune proviennent d'une observation en optique du Canada-France-Hawaii Telescope (CFHT). La distribution de gaz, tracée par l'émission X, est indiquée en rouge et provient d'observations Chandra. La composante bleue montre la distribution de masse mesurée par l'effet de cisaillement gravitationnel faible. Cet amas vient de se former par fusion de deux sous-structures. La survie (temporaire) de deux surdensités de matière noire illustre la nature non-colisionnelle de cette composante.*

## 1.3 Etude réalisée

Ce travail de thèse se situe dans le cadre du projet XMM-LSS, un relevé permettant la détection des amas dans la bande X. Celui-ci profite d'un satellite de nouvelle génération, XMM-Newton, pour assembler un large échantillon d'amas qui permettra de sonder leurs propriétés à une profondeur inédite.

L'étude présentée vise à établir des méthodes 'propres', basées sur les premières données du relevé XMM-LSS, pour la détection, la caractérisation et l'analyse cosmologique des sources X extragalactiques. Outre les résultats préliminaires du relevé qui seront présentés, il s'agit de manière plus générale de tester le potentiel des relevés d'amas de galaxies en rayons X pour atteindre le régime de la cosmologie de précision.

Pour celà, je commencerai par une description plus détaillée des propriétés physiques des amas (Chapitre 2), puis du relevé XMM-LSS (Chapitre 3). Ensuite, j'introduirai les méthodes développées pour cette application (Chapitre 4). Enfin, je présenterai une sélection des résultats obtenus (Chapitre 5) ainsi que quelques perspectives (Chapitre 6).

Une deuxième partie de ce manuscrit présente les publications scientifiques associées aux résultats mentionés, ainsi qu'une introduction contextuelle pour chacun d'eux.

# Chapitre 2
# Contexte astrophysique

## Sommaire



## 2.1 Observations des amas de galaxies en rayon X

### 2.1.1 Développement de l'astronomie X

L'atmosphère terrestre est opaque aux rayons X. L'astronomie X a donc débuté assez tardivement (dans les années 1950) par des observations solaires à l'aide de compteurs Geiger attelés à des roquettes. La détection de la première source extra-solaire (l'étoile binaire Sco X-1) date de 1962. Il faut attendre 1970 pour voir le premier relevé complet du ciel avec le satellite americain Uhuru. Celui-ci utilisait un simple compteur proportionel à gaz et ne disposait pas de système focalisant. Parmi ces 339 détections, les premières sources X extra-galactiques

furent identifiées : une dizaines de galaxies de Seyfert et 50 amas de galaxies environ (Forman et al., 1978). La confirmation de l'origine thermique du rayonnement des amas de galaxies vint rapidement avec la détection des raies d'émission du $Fe_{XXV}$ et $Fe_{XXVI}$ à $\sim$7 keV par ARIEL V dans l'amas Perseus (Mitchell et al., 1976) puis dans nombres d'autres (e.g. Serlemitsos et al., 1977; Mitchell & Culhane, 1977).

Le premier satellite comportant un télescope X, et donc une réelle capacité d'imagerie, fut Einstein (ou HEAO-2, lancé en 1978). Il permit de détecter quelques milliers de sources (e.g. Gioia et al., 1990), dont près de 500 amas (Jones & Forman, 1999).

En 1990, ROSAT effectue un relevé complet du ciel sur la bande [0.1,2.5] keV avec une résolution spatiale de 15" et une sensibilité 1000 fois supérieur à Uhuru, puis opérera des pointés profonds pendant près de 10 ans. Le catalogue résultant, contenant près de 150000 sources fait jusqu'à présent office de référence. Il comprend plusieurs milliers d'amas de galaxies. Les plus brillants d'entre eux ont vu leur température mesurée par les satellites ASCA et BeppoSAX (1993), disposant d'une bonne résolution spectroscopique sur une large bande.

Les satellites de nouvelle génération, XMM et Chandra, utilisent désormais des capteurs CCD. Ils furent tous deux lancés à quelques mois d'intervalle en 1999. Le premier, financé par l'Agence Spatiale Européenne, est 15 fois plus sensible que ROSAT avec une résolution sensiblement améliorée (6"), tandis que le second, financé par la NASA, affiche une résolution exceptionnelle de 1" pour une surface effective 'seulement' 3 fois supérieure à celle de ROSAT.

Bien que cette perspective historique se concentre sur l'étude des amas de galaxies, l'astronomie X occupe maintenant une place importante dans l'études de nombreux objets :

- Les étoiles proches,
- Les objets accrétants galactiques tels que les étoiles binaires, variables cataclysimiques, étoiles à neutron et trous noirs stellaire,
- Les milieux dilués chaud de la Galaxie (bulle locale, restes de supernova, chocs),
- Les galaxies proches, noyaux actifs de galaxies et amas de galaxies.

### 2.1.2 L'émission X des amas

Durant le processus de formation des amas, la matière baryonique s'effondre avec la matière noire. La compression du gaz ainsi que les chocs dans la zone d'accrétion et durant les épisodes de fusion chauffent le gaz jusqu'à des températures de $4 \times 10^7$ à $2 \times 10^8$ K (ou 3 à 15 keV).

Le plasma intra-amas emet alors dans les X par brehmsstrahlung thermique. Ce rayonnement correspond à l'énergie évacuée par les électrons lorsqu'ils sont freinés par les diffusions successives avec les ions du plasma. Compte tenu de la faible densité de baryons dans le milieu intergalactique (moins de un atome par litre), on peut considérer ce plasma comme optiquement mince. De plus, s'agissant d'une intéraction à deux corps, le nombre de photons émis par un élément de volume

infinitésimal $dV$ varie comme le carré de la densité électronique locale :

$$dN(E)/dE = n_e^2 \varepsilon(E,T,Z)dV \tag{2.1}$$

où l'émissivité du plasma $\varepsilon(E,T,Z)$ à l'énergie $E$ dépend de la température $T$ et de la métalicité du plasma $Z$. La densité numérique des atomes lourds étant très faible, l'émission du continuum dépend peu de la métalicité du gaz ; on peut l'écrire $\varepsilon_{cont} \propto g(E,T)T^{-1/2}exp(-E/kT)$, où le facteur de Gaunt $g(E,T)$ est une fonction douce de la température. Le spectre correspondant se traduit donc par une coupure exponentielle à haute énergie caractérisant la température du gaz observé (c.f. 2.1).

La recombinaison des ions du plasma engendre également du rayonnement dans la bande X, pour les couches internes des métaux lourds. En particulier, les raies du Fe K autour de 6.7 keV sont proéminentes quelle que soit la température considérée. D'autres raies du Fe et du Si émettent à plus basse énergie. Bien qu'on ne les détecte généralement pas pour des températures élevées, elles engendrent un refroidissement radiatif très important dans les groupes de galaxies ($M = 5 \times 10^{12}$ à $3 \times 10^{14}$ $M_\odot$) pour lesquels $T \sim 0.1$ à 3 keV.

Dans les études spectroscopiques d'amas en bande X, la forme globale du spectre, l'intensité des raies (du Fe en particulier), et la position du complexe à 6.7 keV permettent donc de containdre respectivement la température, la métalicité et le décalage spectrale du plasma observé.

Puisque l'émission du gaz varie comme $n_e^2$, la densité élevée dans le coeur des amas les plus massifs implique un refroidissement important du gaz. La température chute alors rapidement, ce qui augmente encore le refroidissement de par l'émission des raies à basse énergie. Ce processus devrait en principe donner lieu à des courants de refroidissement au centre des amas les plus relaxés et engendrer des coeurs très denses et très froids. XMM et Chandra ont montré que ces coeurs froids, lorsqu'ils existent, sont beaucoup moins développés que ne le prédisent les modèles les plus simples d'effondrement isobarique (Fabian et al., 2002).

### 2.1.3 Modélisation du milieu intra-amas

#### L'équilibre hydrostatique

Le processus de formation hiérarchique des structures implique que les amas accrètent constamment de la matière et fusionnent régulièrement avec d'autres surdensités. En dehors des cas extrêmes où des structures de masses similaires entrent en collision, il semble néanmoins que ces perturbations affectent peu le gaz dans les parties internes de l'amas. En première approximation, la gaz apparaît en équilibre hydrostatique avec le puit de potentiel de la matière noire. En d'autres termes, la pression du gaz contrebalance les contraintes exercées par la gravitation :

$$\frac{dP(r)}{dr} = -\frac{GM(r)\rho(r)}{r^2} \tag{2.2}$$

En supposant que le plasma se comporte comme un gaz parfait, soit $P(r) = n_g(r)kT(r)$, il devient possible de reconstruire la masse totale à partir des profils

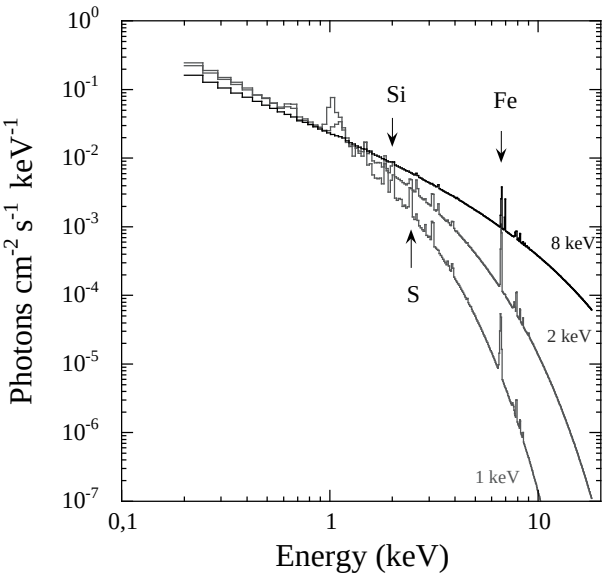


**FIG. 2.1**: *Spectre de brehmsstrahlung thermique d'un plasma optiquement mince pour différentes températures (extrait de Arnaud 2005).*

de densité et de température du gaz :

$$M(r) = -\frac{kT(r)\ r}{\mathrm{G}\mu m_p}\left[\frac{d\ln n}{d\ln r} + \frac{d\ln T}{d\ln r}\right] \tag{2.3}$$

où $m_p$ est la masse du proton et $\mu$ le poids moléculaire moyen d'une particule de gaz, typiquement $\mu \sim 0.6 - 0.7$.

### Le modèle $\beta$

Il est d'usage de modéliser les galaxies d'amas comme une population isotherme, de dispersion de vitesse $\sigma_v$ fixe. On leur associe alors comme distribution spatiale un profil de King (1962) :

$$\rho_{gal} = \rho_0 \left(1 + \frac{r}{r_c}\right)^{3/2} \tag{2.4}$$

qui s'apparente à celle d'une sphère isotherme.

En leur appliquant l'équation d'équilibre hydrostatique, on peut en déduire un profil de masse et l'identifier à celui obtenu de par le gaz (equation 2.3). Dans le cas où la masse gazeuse serait elle aussi isotherme, la correspondance se réduit à $n_g \propto \rho_{gal}^{\beta}$, soit :

$$n_g = n_0 \left(1 + \frac{r}{r_c}\right)^{3\beta/2} \tag{2.5}$$

avec $\beta = \mu m_p \sigma_v / kT$. Après projection le long de la ligne de visée, la brillance de surface attendue en bande X s'exprime comme :

$$S_b = \sqrt{\pi} n_0^2 r_c \Lambda(T,Z) \frac{\Gamma(3\beta - 0.5)}{\Gamma(3\beta)} \left(1 + \frac{r}{r_c}\right)^{0.5-3\beta} \tag{2.6}$$

où $\Lambda(T,Z) = \int_{E_{min}}^{E_{max}} \varepsilon(E,T,Z) dE$ représente l'émission intégrée dans la bande d'énergie considérée $[E_{min}, E_{max}]$.

Ce modèle $\beta$, comme on l'appelle habituellement, se révèle une bonne approximation de la distribution spatiale observée pour les amas X massif lorsque $\beta = 2/3$ (Jones & Forman, 1984). La valeur de $\beta$ semble cependant diminuer pour les systèmes plus légers (Helsdon & Ponman, 2000). Il existe aussi une généralisation du modèle, présentée par Ettori (2000), dans le cas d'un gaz polytrope, i.e. dont la température s'exprime comme $T(r) = T(0)(n_g/n_0)^{\gamma-1}$. La distribution isotherme correspond alors au cas particulier $\gamma = 1$ et un gaz adiabatique à $\gamma = 5/3$. Les modèles physiques acceptables impliquent une valeur de $\gamma$ entre ces deux extrêmes.

La masse hydrostatique pour un modèle $\beta$ polytrope prend la forme :

$$M(r) = \frac{7.33 \times 10^{13}}{\mu} \beta\gamma T_0 r_c \frac{x^3}{(1+x^2)^B} \; M_\odot \tag{2.7}$$

où $x = r/r_c$, $T_0$ s'exprime en keV et $r_c$ en Mpc

## 2.2 Cadre cosmologique : le modèle de halos

L'Univers tel que nous l'étudions en astrophysique peut généralement être décomposé en un ensemble de structures et de sous-systèmes. Le modèle de concordance suppose cependant que l'Univers jeune, tel que le fond diffus cosmologique nous le présente à $z > 1000$, soit extrêment homogène. Les fluctuations de température mesurées ne dépassent guère en effet $\Delta T / < T > \sim 10^{-5}$. Le mécanisme de formation des structures envisagé (par accrétion et fusion autour des inhomogénéités primordiales) ne permettant pas de développer des structures à toutes les échelles simultanément, il existe nécessairement une taille caractéristique au-delà de laquelle aucune surdensité n'a pu s'effondrer. L'Univers doit donc être encore homogène et isotrope à très grande échelle, ce que l'on observe effectivement dans les relevés de galaxies et simulations cosmologiques.

Le formalisme habituel pour modéliser la croissance des structures consiste donc à considérer un univers homogène de densité moyenne $< \rho_m >$, et à traiter les inhomogénéités par traitement perturbatif[1]. L'evolution globale de l'Univers, en particulier de son taux d'expansion, ne dépend alors que de la densité moyenne de matière et d'énergie sombre ($\Omega_m$, $\Omega_\Lambda$) ainsi que de leur équation d'état.

[1] Il s'avère que la validité d'une telle méthode n'a rien d'évident en relativité générale. Certains voient même dans cette supposition fondamentale une solution possible au problème de l'énergie sombre, e.g. Kolb et al. (2006).

### 2.2.1 Traitement perturbatif de l'évolution linéaire

Les perturbations autour de la densité moyenne sont généralement décrites à l'aide de la surdensité locale :

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - < \rho_m >}{< \rho_m >} \tag{2.8}$$

Afin de simplifier encore le problème, on suppose généralement que le champ de fluctuation primordial $\delta(\mathbf{x})$ est gaussien. Son spectre de puissance P(k), définit comme la variance des différents modes de Fourrier qui le composent, constitue alors une description complète de $\delta(x)$ :

$$\delta_k(\mathbf{k}) = \int \delta(\mathbf{x}) e^{i\mathbf{k}.\mathbf{x}} d^3\mathbf{x} \quad ; \quad P(k) = < |\delta_k|^2 > . \tag{2.9}$$

Aucune observation à ce jour n'a pu fermement invalider cette hypothèse.

Dans le régime linéaire, et en ne considérant que la gravité, on peut alors montrer que les différentes longueurs d'onde du champ s'amplifient indépendamment, suivant un facteur de croissance unique :

$$\delta(\mathbf{x}; z) = D(z)\delta(\mathbf{x}; 0) \tag{2.10}$$

où par convention $g(z) = 1$ à $z = 0$. Dans des configurations simples $D(z)$ peut s'obtenir analytiquement, par exemple dans le modèle $\Lambda CDM$ où l'énergie sombre est modélisée comme une constante cosmologique (e.g. Carroll et al., 1992). Dans les autres cas il faut recourir à l'intégration numérique d'une équation différentielle ordinaire.

D'autres processus que la gravité ont par ailleurs un impact significatif sur l'évolution des surdensités, notamment la pression du fluide cosmologique, la diffusion de ses particules et son intéraction avec le rayonnement. Ces altérations additionnelles sont rassemblées dans la fonction de transfert, définie comme :

$$T(k) = \frac{\delta_k(k; 0)}{D(z_i)\delta_k(k; z_i)} \tag{2.11}$$

où $z_i$ est un redshift suffisamment grand pour que le spectre de puissance initial n'ait pas encore été modifié. $T(k)$ ne peut s'obtenir que par des simulations numérique, mais certains articles fournissent des ajustements commodes aux résultats (e.g. (Bardeen et al., 1986) ou (Eisenstein & Hu, 1998)).

En pratique, on utilise un spectre initial invariant d'échelle $P(k) \propto k^{n_s}$ avec $n_s \sim 1$, qui trouve une justification dans le cadre de la théorie de l'inflation primordiale (Guth, 1981). Le spectre de puissance dans l'Univers local s'écrit alors $P(k, z) \propto k^{n_s} T(k)^2 D(z)^2$, forme qui reproduit précisémment la plupart des données actuellement disponibles.

Pour finir, la normalisation de $P(k)$ est le plus souvent spécifiée par la variance du spectre de puissance linéaire lissé sur une échelle de 8 Mpc, notée $\sigma_8$. Ce paramêtre a l'avantage de pouvoir se calculer facilement pour $n$, $D(z)$ et $T(k)$

donnés et d'avoir une valeur proche de l'unité. Les estimations récentes de $\sigma_8$ donnent toutes des résultats entre 0.7 et 1.0. Cependant ont observe des différences systématiques en fonction de la méthode utilisée et les incertitudes de mesure, de l'ordre de 10%, ne permettent pas d'expliquer cette dispersion.

### 2.2.2 Effondrement non-linéaire

L'existence de systèmes fortement non-linéaires tels que les galaxies, groupes et amas de galaxies nous enseigne que l'approximation linéaire n'est pas une description suffisante du monde qui nous entoure. En général, le calcul complet de l'évolution non-linéaire d'une surdensité ne peut s'effectuer analytiquement. On l'étudie donc à l'aide de simulations numériques.

Il reste possible d'appréhender quelques mécanismes communs en considérant des configurations particulières. L'évolution d'une surdensité sphérique dans un univers Einstein-de Sitter (Gunn & Gott, 1972) peut par example se résoudre analytiquement. Le système en effondrement parvient à l'équilibre du viriel pour une densité extrapolée linéairement de $\delta \simeq 1.686$ alors que la surdensité réelle atteint $\Delta = 178$. Ce résultat se généralise, sur la base de simulations, dans un univers $\Lambda CDM$ plat où les structures virialisées présentent un contraste de densité universel Bryan & Norman (1998) :

$$\Delta_c = 18\pi^2 + 82[\Omega_m(z) - 1] - 39[\Omega_m(z) - 1]^2, \tag{2.12}$$

définit par rapport à la densité critique de l'univers.

Cette caractéristique sert généralement de définition aux halos de matière découplés de l'expansion. D'autres contrastes de densité sont parfois considérés, en particulier $\Delta_c = 200$ par analogie avec l'Univers Einstein-de Sitter, et $\rho = 180\rho_m(z)$. Ce dernier semble mieux retranscrire la sélection des halos par les algorithmes de détection utilisés sur les simulations numériques.

### 2.2.3 Fonction de masse des halos

Le formalisme introduit par Press & Schechter (1974) et finalisé par Bond et al. (1991) fournit une approche hybride pour calculer la fonction de masse des halos directement à partir du spectre de puissance linéaire, sans se plonger dans les calculs insolubles du formalisme non-linéaire.

On commence par associer à chaque masse $M$ et chaque reshift $z$ le rayon $r_M(z) = (3M/4\pi\rho_m(z))^{1/3}$ qui englobe cette masse de part la densité moyenne de l'univers à cette époque. Pour un spectre de puissance linéaire donné, on calcule ensuite la variance du champ de densité linéaire lissé à cette échelle $\sigma(M, z)$. Par analogie avec l'effondrement des surdensités sphériques, on définit une densité seuil $\delta_c$ au delà de laquelle un objet s'est formé. Pour un champ de perturbation gaussien, la probabilité qu'une région de taille $r_M$ se soit effondrée en un objet de masse $M$ à un redshift $z$ s'écrit alors $\mathrm{erfc}[\delta_c/\sqrt{2}\sigma(M, z)]$, où un facteur 2 a été introduit artificiellement pour tenir compte des sous-densités. En dérivant par

rapport à la masse et en considérant le nombre de région de taille $r_M$ par unité de volume comobile, on obtient la densité numérique de halo :

$$\frac{dn}{d\,ln\,M}(M,z) = \frac{<\rho_m(0)>}{M} \times \frac{d}{d\,ln\,M}\left(\mathrm{erfc}\left[\frac{\delta_c}{\sqrt{2}\sigma(M,z)}\right]\right) \tag{2.13}$$

En général, on utilise $\delta_c = 1.686$ comme dans un espace Einstein-de Sitter, car cette valeur dépend peut de $\Omega_m$ et $\Omega_\Lambda$ dans un espace plat.

Ce modèle a depuis lors connu plusieurs améliorations, d'abord avec l'incorporation d'un modèle d'effondrement ellipsoidal pour calculer la valeur seuil (Sheth & Tormen, 1999; Sheth et al., 2001), puis par des ajustements directs de $dn(M,z)/d\ln\sigma^{-1}$ sur les simulations (Jenkins et al., 2001; Warren et al., 2006). Ces fonctions de masses analytiques reproduisent précisément la densité de halos observée dans les simulations de matière noire, à condition que l'on utilise une densité $\rho = 180\rho_m(z)$ pour définir la masse.

### 2.2.4 Fonction de corrélation

La distribution spatiale des halos contient énormément d'information sur le champ de densité et donc les paramètres cosmologiques. La méthode la plus courante pour l'extraire est d'utiliser la fonction de corrélation à deux points définie comme :

$$\xi(\mathbf{r}) = < \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) > \tag{2.14}$$

On peut montrer qu'il s'agit directement de la transformée de Fourier du spectre de puissance.

Si l'on ne connait pas le redshift des sources considérées, et donc leur distance, on se contente de la fonction de corrélation angulaire sur la sphère céleste $w(\theta)$, dont la relation avec $P(k)$ est moins triviale. Sous couvert d'un ensemble de conditions généralement satisfaites, l'équation de Limber (Limber, 1953; Baugh & Efstathiou, 1993) fournit ce lien :

$$w(\theta) = \int_0^\infty P(k; z=0) g(k\theta) k \; dk \tag{2.15}$$

où le masque $g(k\theta)$ s'exprime :

$$g(k\theta) = \frac{1}{2\pi}\int_0^\infty J_0\left(k\theta\chi(z)\right)\left[\frac{d\mathcal{N}}{dz}\right]^2 \frac{dz}{d\chi} F(z)\; dz \tag{2.16}$$

Ici, $J_0(x)$ est la fonction de Bessel de première espèce et $\chi(z)$ la distance comobile jusqu'au redshift z. $\frac{d\mathcal{N}}{dz}$ représente la densité de probabilité pour le redshift des sources étudiées. Enfin, $F(z)$ incorpore l'évolution du spectre de puissance et de la population de sources.

En pratique, l'équation de Limber dépend de plusieurs supositions sur la nature des objets utilisés et son inversion, pour obtenir $P(k)$, ne peut d'ordinaire s'opérer

analytiquement. De plus, on ne recouvre pas directement la fonction de corrélation du champ de densité, mais celle d'un ensemble de halos identifiés dans le ciel. D'après Mo & White (1996), la différence peut s'exprimer par un facteur de biais dépendant uniquement de la masse des halos. On peut le dériver directement de l'expression de la fonction de masse via :

$$b(M) = 1 - \frac{d}{d\delta_c}\left[\ln\left(\frac{dn}{d\ln\sigma^{-1}}\right)\right] \tag{2.17}$$

### 2.2.5 Auto-similarité

Des arguments théoriques (Gunn, 1977; Bertschinger, 1985) prévoient une forme auto-similaire du profil de densité des halos, une fois renormalisé en fonction de leur contraste de densité $\Delta_c$ et du rayon l'englobant $r_{\Delta_c}$. Des simulations numériques ont permis à Navarro, Frenk & White (1997) de déterminer la forme commune attendu dans le modèle $\Lambda CDM$, dite profil NFW :

$$\frac{\rho(r)}{\rho_{crit}} = \frac{\delta_{\rm NFW}}{(c\,r/r_{200})\,(1 + c\,r/r_{200})^2} \tag{2.18}$$

où $c$ est un paramètre de concentration et $\delta_{\rm NFW}$ une simple normalisation. La définition de $r_{200}$ implique :

$$\delta_{\rm NFW} = \frac{200}{3}\frac{c^3}{[\ln(1+c) - c/(1+c)]} \tag{2.19}$$

Pour les amas de galaxies la concentration est approximativement constante autour de $c = 4 - 5$, mais elle augmente lorsque l'on considère des structures moins massives. Les observations semblent corroborer cette tendance pour les groupes et amas (Mandelbaum et al., 2006; Pointecouteau et al., 2005), comme le montre la Fig. 2.2.

Plusieurs modèles phénoménologiques, basés sur des calculs analytiques, tentent d'expliquer la relation $c(M, z)$ observée dans les simulations (Navarro et al., 1997; Bullock et al., 2001; Eke et al., 2001; Neto et al., 2007) sans qu'aucun ne préfigure comme supérieur aux autres. La dispersion importante autour de cette relation semble également difficile à reproduire.

### 2.2.6 Relations d'échelle

#### Description auto-similaire du gaz

Dans un papier fondateur, Kaiser (1986) montre que sous certaines conditions, le gaz accrété dans les amas doit lui aussi suivre une distribution auto-similaire en densité et en température. Il en résulte une série de relations d'échelle simples entre différentes grandeurs caractérisitques ($M_{tot}$, $M_{gaz}$,T, $L_X$, ...) pour lesquelles on peut estimer l'évolution avec le redshift. Cette propriété a été depuis confirmée par les observations (voir par exemple la Fig. 2.3 pour le cas de la température).

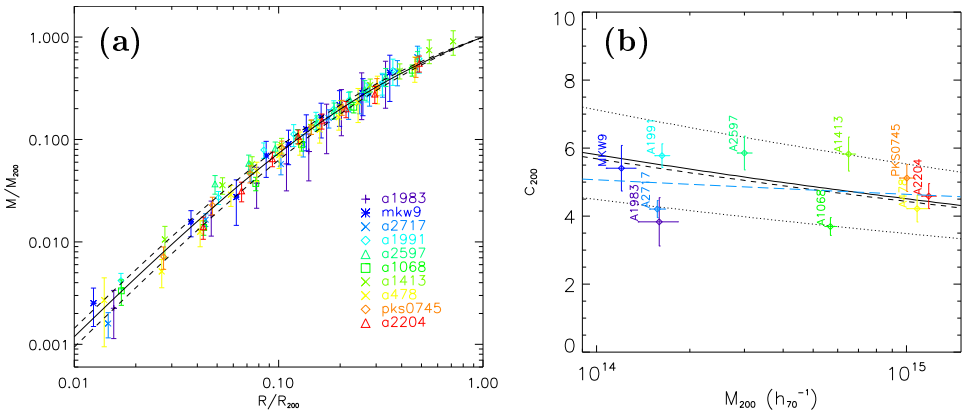


**FIG. 2.2**: *Vérification du profil de masse universel sur un échantillon d'amas X par Pointecouteau et al. (2005). (a) : Profils de masses renormalisés. (b) : Relation $M_{200}$-c.*

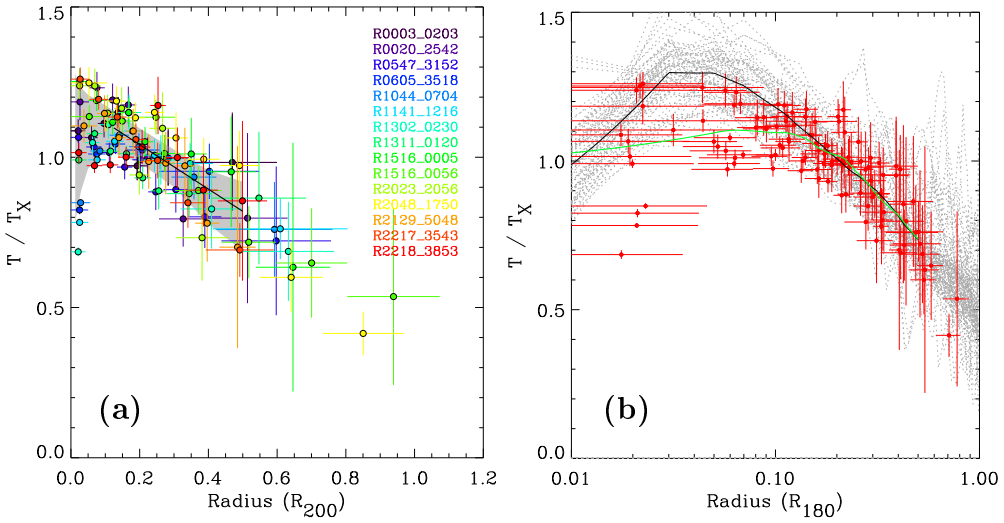


**FIG. 2.3**: *Profils de température renormalisé, mesurés pour l'échantillon de Pratt et al. (2007). (a) : échelle de rayon linéaire, le profil moyen et la dispersion des points de mesure sont indiqué par le trait plein et la zone grisée. (b) : Même graphe avec une échelle logarithmique. Les profils de températures mesurés pour des amas simulés par Borgani et al. (2004) sont superposé à l'aide de lignes en pointillé.*

En combinant, cette auto-similarité avec l'équilibre hydrostatique, et une fraction de gaz $M_{gaz}/M_{tot}$ constante dans le rayon utilisé pour renormaliser, on obtient en particulier les relations d'échelle suivantes :

$$\begin{array}{llll} E(z)^{-1} & L_{bol} & \propto & \Delta(z)^{1/2} \quad T^2 \\ E(z) & M_{tot} & \propto & \Delta(z)^{-1/2} \quad T^{3/2} \\ E(z) & M_{gaz} & \propto & \Delta(z)^{-1/2} \quad T^{3/2} \end{array} \tag{2.20}$$

où $\Delta(z)$ est le contraste de densité, par rapport à la densité critique, définissant le rayon de renormalisation et E(z) représente l'évolution du paramètre de Hubble, $E(z) = H(z)/H_0$. Les simulations numériques comme les observations tendent à favoriser un contraste de densité fixe. $\Delta = 2500$, 500 et 200 sont les valeurs les plus fréquemment utilisées par la communauté X.

Il convient de remarquer à ce stade que l'hypothèse d'une fraction de baryons constante n'a rien d'évidente. En principe, les amas résultent de l'effondrement d'une large région de l'Univers primordial pour laquelle la proportion de gaz correspond à la valeur moyenne universelle. Dans le cas où le gaz suivrait parfaitement la matière noire au cours de l'effondrement cette proportion resterait inchangée. Dans la pratique, la nature colisionnelle du gaz implique qu'il ne peut pas suivre exactement la matière noire. Cela explique que la distribution du gaz ne corresponde pas à un profil NFW, et en particulier comporte un coeur plutôt qu'un pique central prononcé. En revanche, le profil résultant reste auto-similaire et la fraction totale de baryons proche de la valeur cosmique tant que la gravitation seule entre en jeu.

Plusieurs études utilisent cette propriété des amas pour obtenir des contraintes sur les paramètres cosmologiques (e.g. Allen et al., 2007).

### 2.2.7 Lois d'échelles observationnelles

L'étude observationnelle des lois d'échelles permet de mieux comprendre la physique des amas et de définir les limites de validité du modèle auto-similaire. Une bonne connaissance de ces corrélations permet entre autre de lier la masse totale des amas à des observables plus directement accessibles à partir de données moins profondes. Cela s'avère particulièrement important pour les relevés cosmologiques car la masse est le seul paramètre pertinent pour décrire les halos dans le formalisme théorique.

Les meilleures mesures de la relation $M_{500}$-$T$ actuellement disponibles proviennent des travaux de Arnaud et al. (2005) et Vikhlinin et al. (2006). Ceux-ci obtiennent un bon accord pour les amas massif (T>4 keV) et mesurent tous deux une pente compatible avec le modèle auto-similaire ainsi qu'une normalisation comparable à celle observée dans les simulations de Borgani et al. (2004). Dans le régime des groupes (T<3.5keV), Arnaud et al. (2005) trouvent une pente logarithmique plus élevée de $1.72 \pm 0.10$, tandis que Vikhlinin et al. (2006) ne voient aucune déviation. Cette écart possible au modèle auto-similaire reste pour le moment à confirmer.

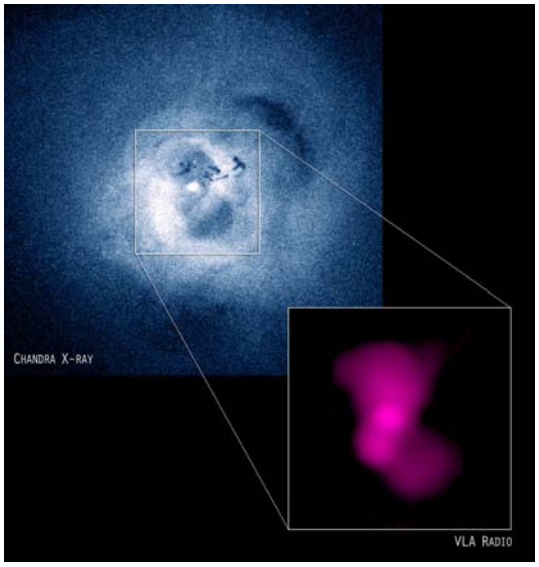


**FIG. 2.4**: *Image profonde Chandra de l'amas Perseus révélant des cavités. L'emission radio (observée par le VLA) provenant de ces bulles de faible densité pleines de particules relativiste est montrée dans l'encadré en bas à droite.*

Il est en revanche certain que la relation $L_X$-$T$ observée ne reproduit pas la pente logarithmique prédite par le modèle auto-similaire : Markevitch (1998) mesure $\alpha = 2.64 \pm 0.27$ et Arnaud & Evrard (1999) $\alpha = 2.88 \pm 0.15$. Il semblerait même que cette valeur augmente encore lorsque l'on s'intéresse aux groupes de galaxies (Helsdon & Ponman, 2000).

Pour expliquer ces déviations du modèle auto-similaire, il est nécessaire d'introduire des processus non-gravitationnels qui modifient l'entropie du gaz. Plusieurs processus ont été proposés dont : le préchauffage du milieu intra-amas avant son effondrement, la rétro-action des noyaux actifs de galaxie (AGNs) et galaxies d'amas par l'intermédiaire des vents galactiques et explosion de supernovae, et enfin le refroidissement radiatif du gaz.

Tous ces mécanismes jouent probablement un rôle. Le recyclage du milieu intragalactique dans le milieu intra-amas constitue la seule explication à son abondance élevé en métaux, tandis que l'emission X elle-même est un processus de refroidissement du gaz. Récemment, les observations à haute résolution de Chandra ont révélé des structures complexes dans le gaz de plusieurs amas (voir par example Sanders et al., 2005). Il s'agit généralement de bulles, vidées en grande partie de leur gaz et remplies de particules relativistes émettant dans la bande radio par émission synchrotron. La figure 2.4 en montre un exemple. Ces bulles sont perçues comme les reliques d'éjection de matériel depuis le noyaux actif de la galaxie centrale d'amas. Il s'agirait d'un des processus majeur d'injection d'entropie dans le milieu intra-amas, et ces cycles d'accrétion-éjection permettraientt d'expliquer la quasi-absence de courant de refroidissement au coeur des amas.

Une autre source de questionnement concerne la dispersion autour des relations d'échelles observées. Globalement, elle retranscrit les fluctuations dans l'histoire de formation particulière à chaque halos. Une part importante de cette dispersion provient des amas en cours de fusion, ainsi que de l'importance variable du refroidissement dans la région centrale.

La question de l'évolution des lois d'échelles reste encore ouverte. Les quelques études effectuées en ce sens ne montrent pas un accord suffisant pour conclure à un modèle précis. Mieux connaître l'évolution des relations d'échelle permetttrait de départager les différentes modèles proposés pour reproduire les relation locales (e.g. Voit, 2005; Cavaliere & Lapi, 2008).

## 2.3 Autres propriétés observationnelles des amas

### 2.3.1 Les galaxies

Les amas n'ont longtemps été étudiés qu'en optique où l'emission des galaxies domine[2]. Ces observations (qu'elles soient spectrocopiques ou seulement photométriques) restent incontournables pour valider l'éxistence d'un système et en particulier déterminer sont redshift. Les autres méthodes comme la tomographie par lentille gravitationnelle ou la spectroscopie X demeurent marginales.

Comme mentionné dans la section 2.1.3, la population des amas peut être représentée par une distribution isotherme, suivant un profil de King. On modélise d'ordinaire leur répartition en luminosité par une foncion de Schechter (1976), soit :

$$\phi(L)dL = \phi^* \left(\frac{L}{L_*}\right)^{\alpha} \exp\left(-\frac{L}{L_*}\right) \frac{dL}{L_*} \tag{2.21}$$

où $\alpha$ détermine la pente à basse luminosité, $L_*$ est une luminosité caractéristique de l'amas au-delà de laquelle une coupure exponentielle s'opère. Enfin, la normalisation $\phi_*$ indique la densité de galaxies par unité de volume et de luminosité évalué à $L_*$.

Plusieurs observables liées aux galaxies d'amas permettent d'estimer leur masse. Une première, de nature dynamique, consiste à mesurer la dispersion de vitesse des membres de l'amas par spectroscopie, et d'appliquer le théorème du viriel. Une seconde fréquemment employée consiste à définir un paramêtre de richesse, qui correspond à un nombre caractérristique de galaxie dans un volume et un intervalle de luminosité donné (Becker et al., 2007). On peut également additionner la luminosité des galaxies membres pour obtenir une luminosité optique de l'amas (Popesso et al., 2005). Dans ces deux derniers cas, il faut alors étalonner une relation masse-observable.

La compilation de grands échantillons d'amas détectés comme surdensités de galaxies subsiste de nos jours, mais les méthodes se sont considérablement améliorées. Il s'agit désormais de détecter une surdensité vérifiant certaines propriétés

[2] certains travaux s'orientent cependant vers l'analyse de la lumière diffuse intra-amas, associée aux étoiles libres intergalactiques. Citons par exemple (Adami et al., 2005; Mihos et al., 2005) et (Krick & Bernstein, 2007)

**FIG. 2.5**: *Coeur de l'amas Abell 1689 observé par le Hubble Space Telescope. De nombreux arcs sont visibles, provenant de la distortion de galaxies d'arrière plan.*

caractéristiques de amas, diminuant ainsi la contamination importante due aux effets de projection et aux structures filamentaires. Une des méthodes les plus en vogue est la détection de la séquence rouge (Gladders & Yee, 2000; Koester et al., 2007). Elle prend en compte la similitude en couleur des galaxies d'un même amas pour identifier des surdensitées dans l'espace position-couleur. La technique des filtres adaptés permet également de considérer la distribution spatiale et/ou en luminosité des galaxies d'amas (Olsen et al., 2007)

Evidemment, nombre de travaux se concentrent aussi sur la physique des galaxies, et en particulier l'analyse comparée des galaxies d'amas et de champs.

### 2.3.2 Lentilles gravitationnelles

Le traitement géométrique de la gravitation en relativité générale implique que la lumière, comme toute autre forme de matière, est déviée par la présence d'objets massifs. Les amas de galaxies en particulier modifient la trajectoire des photons sur la ligne de visée.

Dans l'approximation dite de lentille mince, où l'on considère que toute la masse de l'amas se trouve dans un plan, et pour de petits angles de déviations, la position des sources dans le plan image $\vec{\theta}_I$ et leur position réelle $\vec{\theta}_S$ vérifient :

$$\vec{\theta}_I = \vec{\theta}_S + \frac{2}{c^2}\frac{D_{LS}}{D_{OS}}\vec{\nabla}\phi \tag{2.22}$$

où $\phi$ est le potentiel newtonien projeté, tandis que $D_{LS}$ et $D_{OS}$ représentent respectivement les distances angulaire de la source affectée à la lentille et à l'observateur.

Cette transformation peut s'exprimer à l'aide d'une matrice d'amplification :

$$\frac{d\vec{\theta}_S}{d\vec{\theta}_I} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \tag{2.23}$$

où le paramètre $\kappa$, appelé convergence, correspond à la densité surfacique dans le plan de la lentille en unité d'une densité critique $\Sigma_{crit} = \frac{c^2}{4G}\frac{D_{OS}D_{OL}}{D_{LS}}$ (nous introduisons la distance angulaire entre l'observateur et la lentille, $D_{OL}$). Le deuxième paramêtre, $\gamma = \gamma_1 + i\gamma_2$, est un terme de cisaillement, lié à la distribution de masse d'une manière moins directe.

Dans les régions des amas où $\Sigma \geq \Sigma_{crit}$, le signal de fond peut être très fortement amplifié et déformé, générant de arcs, souvent multiples aux centres des amas massifs (voir la Fig. 2.5 pour un exemple). Ces effets de lentilles gravitationnelles fortes contiennent beaucoup d'informations sur la distribution de masse. Le profil dans les parties internes de l'amas se voit généralement contraint par l'ajustement d'un modèle paramétrique. Dans les régions plus externes, l'effet devient beaucoup plus faible. Etant donné que l'amplitude de la distortion sur une galaxie distante est très petite devant la largeur de la distribution en ellipticite des galaxies, il faut employer une approche statistique, en moyennant les propriétés d'un ensemble de sources sur la même ligne de visée. La mesure du cisaillement et l'inversion vers une carte de masse projeté comportent alors de nombreuses difficultées liées à la faiblesse du signal considéré.

D'autres utilisations des effets de lentilles permettent de contraindre les paramêtres cosmologiques. On citera par exemple les statistiques du cisaillement cosmique, résultant de l'effet intégré de toute la distribution de masse sur la ligne de visée (e.g. Hoekstra et al., 2006; Hetterscheidt et al., 2007). Les lentilles fortes présentant de nombreux arcs contiennent énormément d'informations. Il devient possible d'en extraire des contraintes lorsque les redshifts des galaxies sources sont connus (e.g. Koopmans & Fassnacht, 1999). Enfin, les effets de lentilles introduisent de faible distortions non gaussiennes sur le fond diffus cosmologique (e.g. Bernardeau, 1997).

L'avantage des méthodes de lentilles gravitationnelles tient à leur indépendance vis-à-vis de tout autre paramêtre que la masse. Il s'agit donc de la seule mesure directe disponible.

### 2.3.3 L'effet Sunyaev-Zel'dovich

Tout comme les observations X, cette technique sonde les propriétés du gaz intra-amas. Il s'agit de mesurer l'impact du milieu intra-amas sur le fond diffus cosmologique.

La diffusion par les électrons du gaz de ce rayonnement de corps noir à 2.73 K produit en effet des distorsions dans le spectre (Sunyaev & Zeldovich, 1980). Dans la limite non-relativiste et pour des diffusions simples (justifiées par l'épaisseur optique du plasma intra-amas), la variation de température de brillance correspondant à la distribution thermique des électrons a une dépendance spectrale

caractéristique. Elle s'écrit :

$$\frac{\Delta T}{T_{cmb}} = y_c \left[ x \frac{e^x + 1}{e^x - 1} - 4 \right] \tag{2.24}$$

où $x = \frac{h\nu}{kT_{cmb}}$ est une mesure sans dimension de la fréquence, $T_{cmb}$ représente la température locale du fond diffus, et toutes les caractérisiques du gaz sont incorrporées dans le paramêtre $y_c$, dénommé comptonisation. Ce dernier ne dépend que de l'intégrale de la pression du gaz, le long de la ligne de visée :

$$y_c = \frac{k_B \sigma_T}{m_e c^2} \int n_e T dl \tag{2.25}$$

où $k_B$, $\sigma_T$, $m_e$ sont la constante de Boltzmann, la section efficace de la diffusion compton et la masse de l'électron. On remarquera que la forme et l'amplitude de la distortion sont indépendantes du redshift, ce qui implique l'intéret croissant pour ce type de mesure. L'effet se manifeste comme un décrément de brillance de surface pour des longueurs d'onde inférieures à 220 GHz, et comme un incrément au-delà (les photons diffusés sont redistribués vers de plus hautes énergies).

Comparé à l'émission X, la comptonisation dépend moins fortement de la densité du gaz, ce qui rend le signal moins sensible aux sous-strucutres et permet une détection jusqu'à un rayon plus large. Néanmoins, ces deux caractéristiques tendent à augmenter les effets de projection comparativement aux relevés X. De plus, l'extraction du signale nécessite de le distinguer parmi plusieurs contaminants (en fonction de la longueur d'onde) : fluctuations primaires du fond diffus, emission des poussières, galaxies infra-rouge. Ces deux dernières composantes ont un impact plus important dans la zone d'incrément du signal. Pour cette raison, la plupart des expériences fonctionnent à grande longueur d'onde (e.g. OVRO/BIMA, CBI, SZA).

La combinaison de la comptonisation avec l'émission X devrait permettre une déprojection directe des profils de gaz et de température, sans avoir recours à la spectroscopie X. Les premières tentatives sur ce point utilisent des profils paramétrés (LaRoque et al., 2006) mais des méthodes non-paramétriques semblent envisageablent dans l'avenir (Ameglio et al., 2007). La comparaison entre la comptonisation et la brillance de surface X fait en outre intervenir la distance diamètre angulaire, ce qui permet un test géométrique pour la contrainte des paramètres cosmologiques. Les résultats sur ce point se bornent à mesurer la constante de Hubble (Bonamente et al., 2006), mais rien n'empêche en principe de construire des diagrammes redshift-distance comparables, à condition de constituer un échantillon d'amas à grand redshift ($1 < z < 1.5$). On obtiendrait ainsi des données similaires à celles déduites des supernovae, mais dont les effets de sélections et les systématiques différeraient.

Au jour présent, la sensibilité des instruments n'a pas encore atteind un niveau équivalent à celles des études en bandes X. Les années à venir verront cependant une augmentation considérables en nombre de source et profondeur avec plusieurs grandes expériences, notamment le satellite PLANCK de l'ESA, et les relevés dédiés du South Pole Telescope (SPT), de l'Atacama Pathfinder Experiment (APEX-SZ) et de l'Atacama Cosmology Telescope (ACT).

### 2.3.4 Les amas en radio

L'émission radio la plus couramment observée en direction des amas provient des AGNs qu'ils contiennent ou qui tombent sur la même ligne de visée.

On observe aussi des zones étendues émettrices en radio, dont l'origine est attribuée au rayonnement synchrotron de particules relativistes distribuées dans le milieu intra-amas. Une telle émission nécessite l'existence d'un champ magnétique de l'ordre de quelques $\mu G$. Ces valeurs sont compatibles avec les mesures déduites de la rotation Faraday (Vogt & Enßlin, 2003), mais nécessite l'existence dans l'univers jeune de fluctuations difficiles à reproduire d'un point de vu théorique (e.g. Langer et al., 2003).

Les zones d'émission radio peuvent se classifier en trois catégories :

- Les halos radio ont une faible brillance de surface et s'étendent sur une échelle comparable avec la taille totale de l'amas. De telles halos semblent exister dans $\sim 5\%$ des amas massifs (Giovannini et al., 1999). Leur origine est imputée aux particules énergétiques générées lors des collisions d'amas, ce qui expliquerai leur distribution comparable au gaz des sources. On observe aussi des reliques similaires aux halos radio, mais de formes irrégulières et distribuées en périférie. Elles résultent probablement de l'onde de choc successive aux fusion d'amas, plutôt qu'à la collision elle-même
- Dans certains amas à courant de refroidissement, on observe des mini-halos, s'étendant sur $\sim 500$ kpc, qui se situent toujours dans les parties centrales. Gitti et al. (2002, 2004) ont développé un model dans lequel des électrons énergétiques seraient réaccélérés par le courant de refroidissement. Leur model semble conforté par l'existence d'une corrélation entre la puissance du courant de refroidissement et la luminosité radio des mini-halos.
- Enfin, les bulles radio sont irrégulières et de taille encore plus modestes. On ne les trouvent que dans les parties centrales, où elles coincident avec des cavités dans le milieu intra-amas. Elles semblent donc résulter d'une intéraction complexe entre l'AGN de la galaxie centrale et le milieu environnant, comme illustré par la Fig. 2.4. Cette intéraction reste l'explication la plus probable de la faiblesse des courants de refroidissement.

## 2.4 Quelques mots sur les AGNs

Les noyaux actifs de galaxie constituent l'essentiel des sources X extragalactiques. Le modèle unifié les décrivant postule que tout AGN se compose d'un trou noir supermassif situé au centre de la galaxie, autour duquel la matière s'effondre, formant un tore de gaz et de poussière. Dans le plan perpendiculaire, des jets de particules énergétiques expulsent la matière qui n'a pas été absobée. Selon ce modèle, les caractéristiques variées des noyaux actifs observés ne proviendraient que des différents angles sous lesquels l'objet central apparait. L'emission X proviendrait des régions internes proches du noyaux, où la matière est effectivement absorbée. Une composantes plus molle pourrait provenir du tore lui-même.

L'étude des AGNs dans la bande X s'axe sur différentes thématiques. L'une

d'elle est de mieux comprendre leurs mécanismes d'émission complexes dans le cadre du model standard (e.g. Barcons et al., 2003), souvent par la combinaison de plusieurs longueur d'ondes (comme Polletta et al., 2007). D'autres études se penchent sur l'évolution de ces propriétés (Barger et al., 2005; Hasinger et al., 2005).

La distribution spatiale des AGNs fait l'objet de nombreuses investigations, en particulier leur fonction de corrélation (Basilakos et al., 2005; Gilli et al., 2005; Yang et al., 2006). Comme nous l'avons mentionné dans la section 2.2.4, une mesure précise de cette fonction permet en effet de contraindre les paramêtres cosmollogiques (Basilakos & Plionis, 2005, 2006). Dans le cas des AGNs dont la formation et l'évolution sont mal connus, cela permet aussi de comparer le facteur de biais entre la population globale de galaxie et celles contenant un noyau actif.

Enfin, la majeure partie du fond X de haute énergie est aujourd'hui attribuée à la population d'AGNs (Setti & Woltjer, 1989; Barcons et al., 1991). A plus basse énergie, l'émission de la bulle locale et du plasma galactique compte pour une fraction significative. Sur la base d'observations profondes, des études tentent de tester cette hypothèse en résolvant en sources ponctuelles une part significative du fond X extragalactique (e.g. Carrera et al., 2007).

# Chapitre 3
# Le relevé XMM-LSS

## Sommaire



## 3.1 L'observatoire spatial XMM-Newton

### 3.1.1 Description

Le satellite XMM (abbréviation de "high throughput X-ray spectroscopy Multi-Mirror mission", Jansen et al. 2001) est la seconde des 4 pierres angulaires du projet à long terme de l'ESA *HORIZON 2000*.

Il dispose de trois télescopes X co-alignés (XRT1 à XRT3) focalisant le rayonnement sur deux spectromètres à réseaux de haute résolution (RGS pour Reflection Grating Spectrometer), et trois caméras de technologie CCD denommées EPIC (pour European Photon Imaging Camera). Un quatrième téléscope, de type Ritchey-Chretien et baptisé Optical Monitor (OM), permet d'observer simultanément le champ en UV et en optique pour des longueurs d'onde comprises entre 170 et 550nm. Celui-ci a un diamètre de 30cm. Le montage optique du satellite est illustré par la Fig. 3.1 tandis que la Table 3.1 décrit les attributs de chaque instrument.

Le lancement d'XMM par une fusée Ariane 5 date du 10 décembre 1999. Il parcourt depuis lors une orbite très elliptique (7 000 km de périgée pour 114 000 km d'apogée) en environ 48 heures par cycle. A l'occasion de sa première lumière le 9 Février 2000, il fut renommé XMM-Newton en l'honneur des travaux fondateurs de ce dernier dans le domaine de la spectroscopie optique.

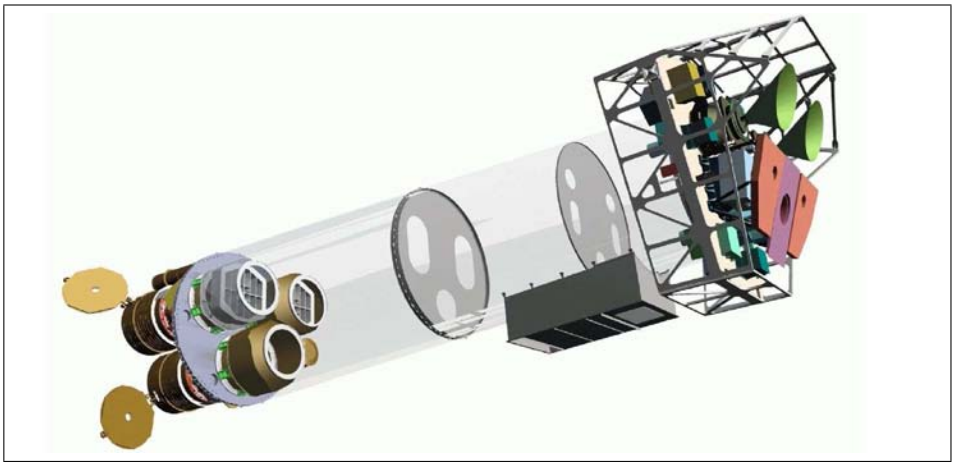

**FIG. 3.1**: *Structure schématique de l'observatoire spatial XMM-Newton. En bas à gauche, on peut observer les trois modules de mirroirs et leurs obturateurs. Les RGAs sont situés à la sortie des deux blocs mirroirs supérieurs (XRT1 et XRT2) et dévient une partie du faisceau vers les RFCs (identifiables dans le plan focal par leurs radiateurs roses). Les détecteurs MOS (dont les radiateurs sont les deux cônes verts) se trouvent dans l'axe optique de ces deux télescopes. La caméra PN (radiateur violet) récolte tous le flux du troisième télescope (XRT3). Divers roues à filtres sont disposées entre les mirroirs et le plan focal. Enfin, le Moniteur Optique est situé près des mirroirs X, caché ici derrière le télescope XRT3.*

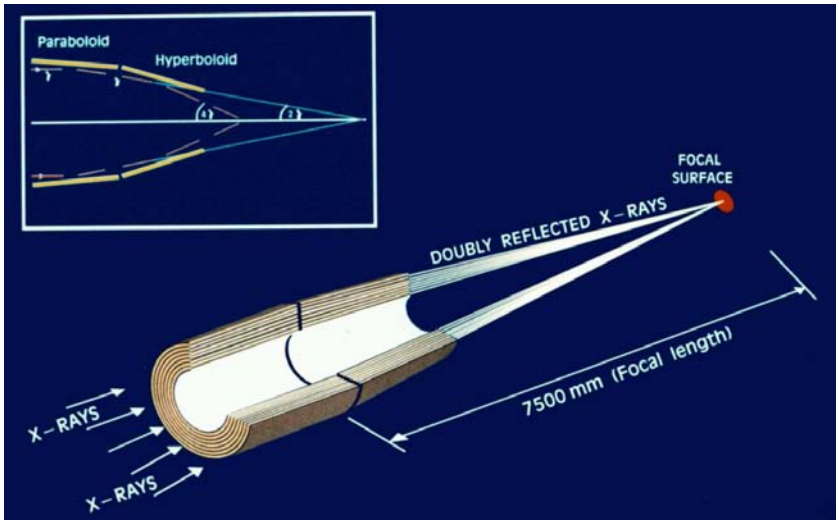


**FIG. 3.2**: *Focalisation par les télescopes de types Wolter I. Les photons sont d'abord réfléchis en incidence rasante sur une section de mirroirs paraboliques, puis sur une portion hyperbolique. La distance focale associée à ce type de mirroirs est nécessairement très grandes (7,50m pour XMM).*

| Instrument | EPIC MOS | EPIC pn | RGS | OM |
|---|---|---|---|---|
| Bande passante | 0.15-12 keV | 0.15-15 keV | 0.35-2.5 keV | 180-600 nm |
| Champ de vision | $30'$ | $30'$ | $5'$ | $17'$ |
| Résolution spatiale | $6''$ | $6''$ | s.o. | $1.4''$ |
| Résolution specrale | 70 eV | 80 eV | 2 eV | s.o. |

**Tab. 3.1**: *Principales caractéristiques des instruments équipant XMM. La résolution spatiale indiquée correspond à la largeur à mi-hauteur de la fonction d'étalement du point dans l'axe optique des télescopes.*

Les possibilités variées d'XMM (haute résolution spectrale, sensibilité inégalée et résolution spatiale excellente) en font un instrument de choix pour presque tous les domaines de l'astronomie X : depuis les étoiles et amas d'étoiles (Güdel et al., 2001; Favata et al., 2003) jusqu'aux amas de galaxies (Arnaud et al., 2001; Finoguenov et al., 2005; Pratt et al., 2006) en passant par les restes de supernovae (Decourchelle et al., 2001; Willingale et al., 2001) objets compacts (Rodriguez et al., 2003; Burwitz et al., 2003; Bélanger et al., 2005), galaxies et AGNs (Yuan & Narayan, 2004; Mainieri et al., 2002; Perola et al., 2004).

### 3.1.2 Miroirs et focalisation

Dans le domaine des rayons X, les photons sont extrêmement pénétrant et il devient impossible d'utiliser, comme en optique, un unique miroir parabolique pour défléchir efficacement le rayonnement. Pour cela, les télescopes d'XMM s'appuient sur une architecture Wolter I (cf. Fig. 3.2) : les photons arrivent en incidence rasante sur une structure très évasée de manière à augmenter l'indice de réflexion ; comme la section efficace d'un tel système optique est très faible, ce type de télescope se compose généralement d'un ensemble de miroirs co-alignés.

Dans le cas d'XMM, chaque module est constitué de 58 coques (en nickel) concentriques recouvertes d'une couche réfléchissante en or. La surface effective total approche ainsi 1500 $cm^2$ par télescope à 1.5 keV. Ce chiffre décroît avec l'énergie considérée, avec en particulier une chute prononcée au delà de 2.2 keV (seuil d'absorption de l'Au M). La sensibilité diminue également d'un facteur 2 à 3 en bordure de champ (effet de vignettage similaire à celui que l'on observe dans le visible). Dans le cas des télescopes XRT1 et XRT2, des réseaux par réflexion (Reflexion Grating Arrays ou RGAs en anglais) sont disposés à la sortie des télescopes, afin de disperser une partie du rayonement incident vers les Caméra Focales du RGS (RFCs). Cela a pour conséquence des variations azimuthales du vignettage pour la fraction du faisceau non-dispersée. La Figure 3.3 illustre ces variations de sensibilité à travers le champs de vision.

La fonction d'étalement du point (PSF) dans le plan focal, où sont installées les caméras EPIC, dépénd principalement de la précision sur l'alignement des miroirs. Pour XMM, l'assemblage des modules de 58 coques constituait donc un véritable défit technologique. Au final, la largeur à mis hauteur (FWHM) de la PSF dans l'axe correspond à un angle apparent d'environ $6''$et ne varie que faiblement en

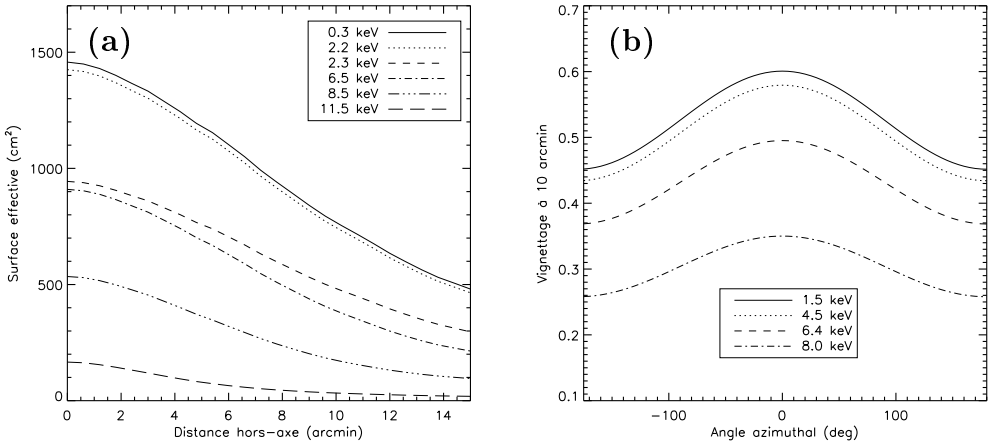


**FIG. 3.3**: *Surface effective des télescopes X d'XMM. (a) : Surface effective moyenne en fonction de la distance hors-axe et de l'énergie des photons incidents. (b) : Variation azimuthal du vignettage des télescopes XRT1 et XRT2 due à la présence des RGA.*

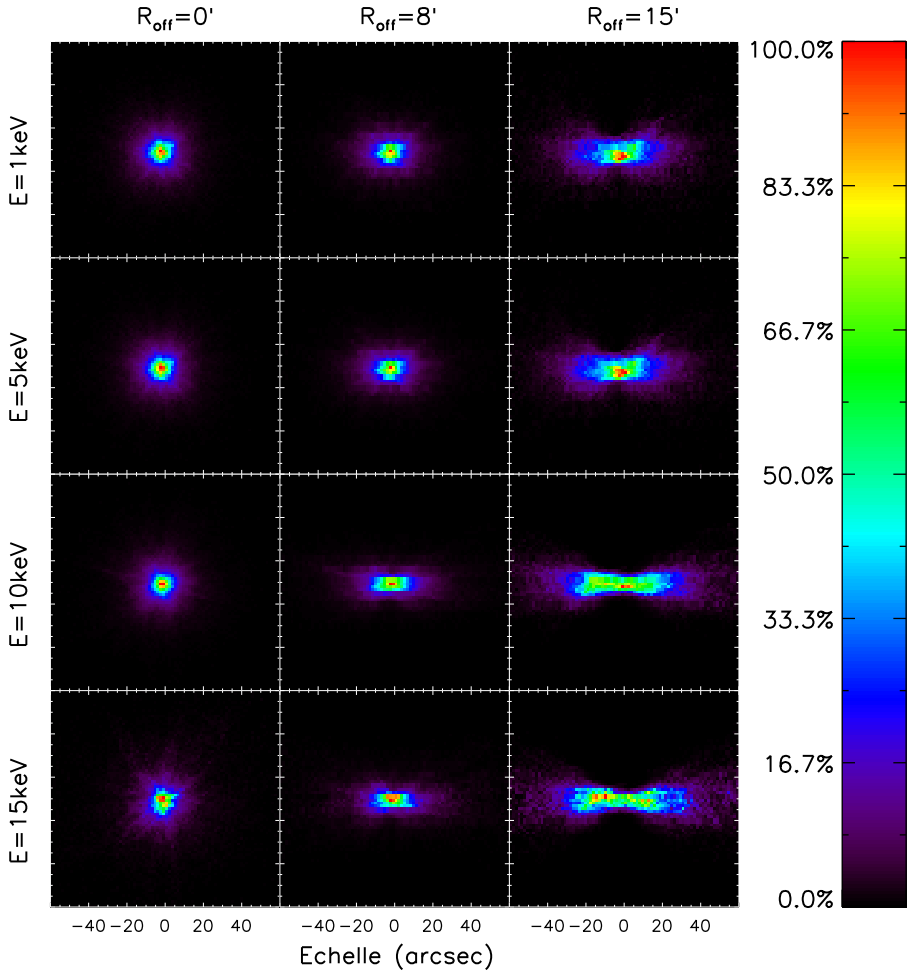


**FIG. 3.4**: *PSF d'XMM en fonction de l'énergie et de la distance hors-axe. L'échelle de couleur indique la brillance de surface rapportée à sa valeur maximale.*

fonction de l'énergie des photons. Lorsque l'on s'éloigne du centre du champ, la PSF s'élargie et adopte une forme asymétrique 'en cacahuète'. Cette distortion est d'autant plus marquée que l'énergie considérée augmente (voir Fig. 3.4).

### 3.1.3 Les caméras EPIC

#### Caractéristiques des détecteurs

Le flux des télescopes XRT1 et XRT2 non dispersé par les RGAs attérit sur deux détecteurs MOS (Metal Oxide Semi-conductor, Turner et al. 2001) fabriqués à Leicester (UK). Chacun d'eux se composent de 7 plaques CCD non co-planaires dont les pixels physiques ($40\mu m$) correspondent à $1.1''$dans le ciel. Elles sont tournées de 90° l'une par rapport à l'autre de manière à ce que les zones non observables de coincident pas.

Un détecteur à jonction p-n (imaginativement dénommé PN, Strüder et al. 2001), récolte l'intégralité du flux de XRT3. De fabrication allemande (Garching), il comprend 12 plaques CCD dont les pixels de $150\mu m$ confèrent une moins bonne résolution spatiale comparé aux MOS ($4.1''$). Le centre physique du détecteur coincide avec un interstice entre les CCD. La caméra fut donc décallée par rapport à l'axe optique de son télescope (d'environ $2.5'$) de façon à collecter au moins 90% du flux total d'une source ponctuelle dans l'axe.

En terme de sensibilité et de résolution spectrale, les trois détecteurs affichent des performance comparables. Bien que l'efficacité quantique du PN ne soit que légèrement supérieure à celle des MOS, ces dernières sont en pratiques 2 à 3 fois moins sensibles (en fonction de l'énergie) à cause de la faible transmission des RGA.

De part l'utilisation de la technologie CCD, les détecteurs EPIC sont de fait sensibles à toutes les longueurs d'ondes depuis le proche infra-rouge. En vue des applications scientifiques variées de l'instrument, le satellite est donc pourvu de divers filtres servant entre autre à bloquer l'émission à basse énergie, de l'infra-rouge aux UV : 'thin' (mince), 'medium' (moyen), 'thick' (épais), 'closed' (opercule de métal afin d'étudier le bruit instrumental). Pour les applications cosmologiques comme le XMM-LSS, où les sources considérées restent relativement faibles, le filtre 'thin' est presque toujours utilisé afin de récolter un maximum de flux.

La Fig. 3.5 présente la surface effective finale issue de la combinaison des télescopes X, des détecteurs EPIC et des filtres 'thin'.

#### Fond instrumental et de particule

La modélisation du fond est un problème central à l'étude des sources étendues à faible brillance de surface comme les amas de galaxies ou les restes de supernova.

Une fraction importante du fond observationnel provient de sources astrophysiques, en particulier l'émission intégrée des AGNs non-résolus ainsi que de l'emission thermique du milieu environant (bulle locale, milieu galactique, héliosphère). Comme il s'agit de véritables photons X, cette composante subit le vignettage des télescopes. L'émission des AGNs peut se modéliser spectralement par une loi de

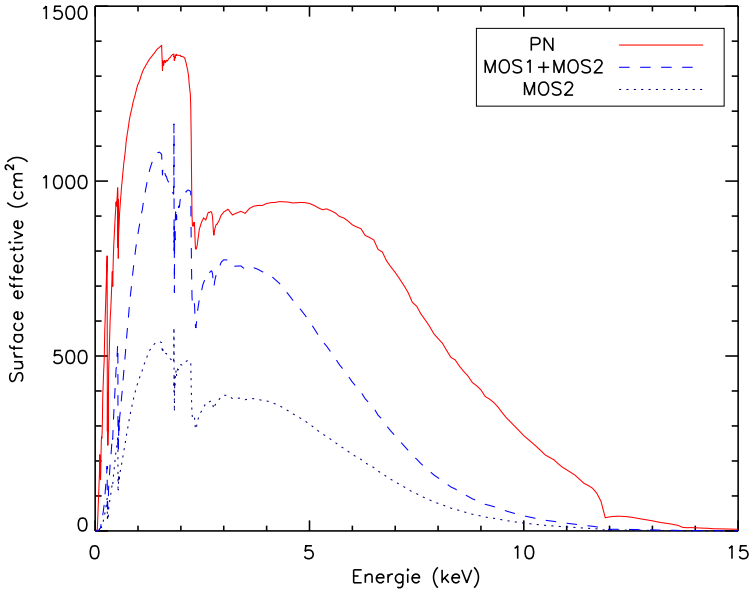


**FIG. 3.5**: *Surface effective sur l'axe d'XMM pour les caméras EPIC pourvues de filtres minces. Celle-ci prend en compte la surface collectrice des miroirs, la transmission des RGA et des filtres, ainsi que l'efficacité quantique des détecteurs. Le saut à 2.2 keV dû à l'absorption de l'or composant les miroirs y est clairement visible.*

puissance, tandis que celles des plasmas galactiques correspondent à des brehmsstrahlung de basses températures.

Au delà de 4 keV, le bruit instrumental et l'intéraction avec des particules comptent cependant pour plus de la motié du signal observé. Il s'agit pour une part de protons mous présents dans la magnétosphère céleste et focalisés par les télescopes. Le vignettage pour ces particules est moins important que pour les photons. Leur distribution, spatiale comme spectrale, varie beaucoup. Le reste du signal résulte du bombardement de rayons cosmiques. Ceux-ci créent des charges directement dans les capteurs CCD et engendrent des raies de fluorescence lors de leur intéraction avec les détecteurs. La distribution spectrale et spatiale associée aux charges induites s'avère relativement uniforme. Les raies par contre se situent à des fréquences bien spécifiques et montrent des fluctuations spatiales caractéristiques. En particulier, les raies du cuivre et du nickel dans la caméra PN proviennent du système électronique situé sous le détecteur. Leur distribution reflète la configuration des circuits, comme le montre la Fig. 3.6.

Le bruit de lecture des CCD est inéxistant au delà de 500 eV et quasiment nul à partir de 300 eV.

Le spectre total résultant est illustré par la Fig. 3.7 On trouvera une description plus compète du fond dans Carter & Read (2007).

La variabilité spectrale et temporelle du fond rend difficile l'analyse spectrale

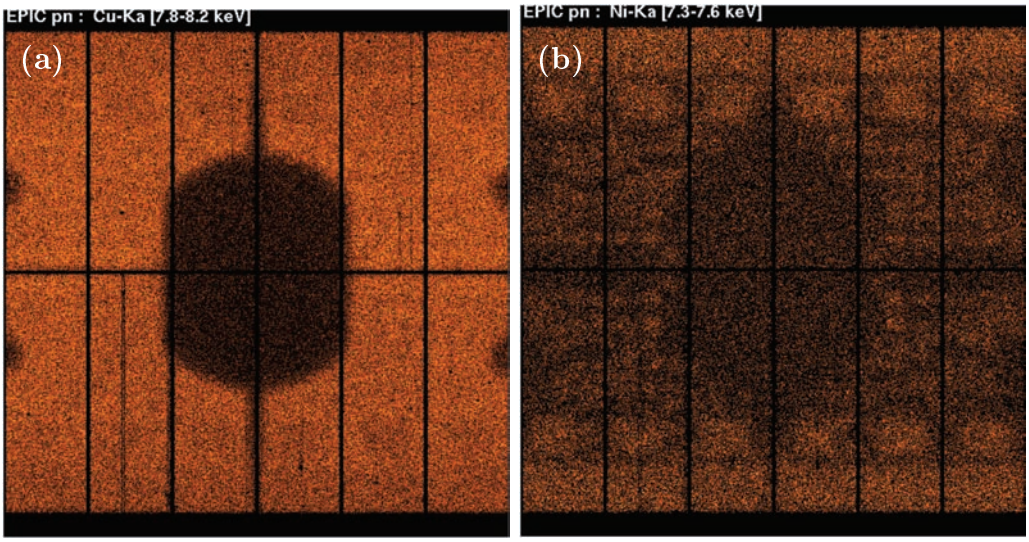


**FIG. 3.6**: *Distribution spatiale inhomogène des raies de fluorescence à hautes énergies de la caméra PN. (a) : Raie du cuivre à 8 keV. (b) : Raie du Nickel vers 7.5 keV.*

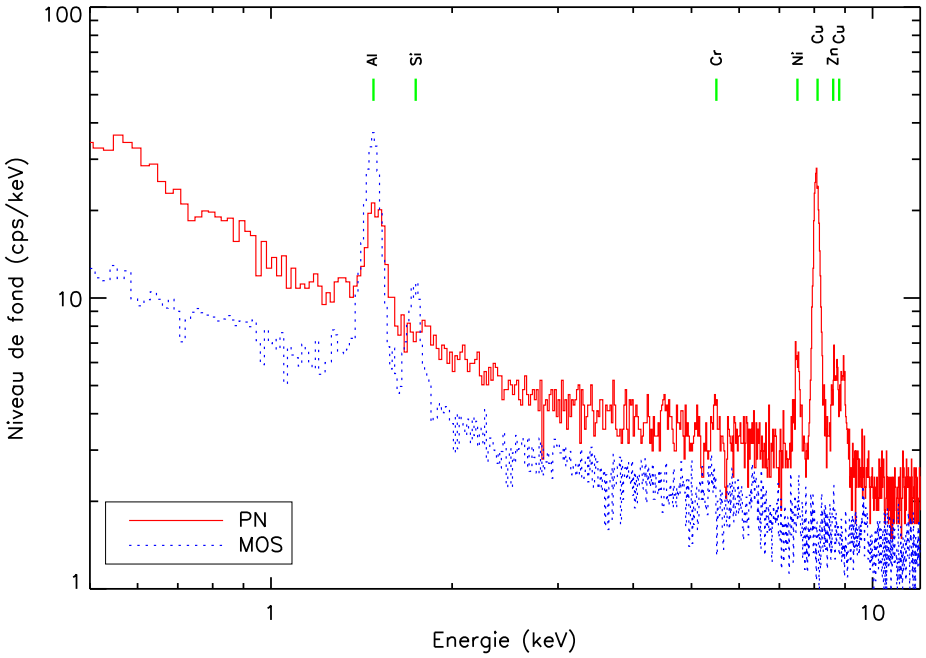


**FIG. 3.7**: *Spectres de fond illustratifs pour les caméras EPIC, tirés d'observations du XMM-LSS. Les principales raies de fluorescence provenant des matériaux du satellite sont indiquées dans la partie supérieure.*

des sources étendues recouvrant une fraction significative du champ de vision, d'autant plus lorsque la brillance de surface est faible. Pour palier à ce problème, on a habituellement recours à des cartes de fond profond, issue de l'empilement de nombreuses observations. Celles-ci sont renormalisées au niveau mesuré dans les observations, sur la base des évènements extérieurs au champ de vision, ou à très haute énergie. Les composantes jugées variables peuvent être ajoutées au model spectral ajusté.

## 3.2 Le relevé XMM-LSS

### 3.2.1 Motivation

Le satellite ROSAT, avec son relevé de tout le ciel, a permis d'accroître considérablement, en taille et en profondeur les échantillons d'amas utilisables pour contraindre le champ de densité. La Fig.3.8 donne un aperçu comparatif des principaux relevés X existants. Les échantillons d'amas les plus utilisés pour les mesures cosmologiques furent REFLEX (Böhringer et al., 2001, ROSAT All Sky Surcey), NEP (Henry et al., 2001) et RDCS pour les archives de pointés ROSAT, ainsi que le relevé Einstein EMSS (Gioia et al., 1990; Stocke et al., 1991). Les résultats obtenus concernent les fonctions locales ($zleq0.3$) de luminosité (Böhringer et al., 2002), de température (Henry, 1997) et de corrélation à 2 points (Collins et al., 2000). Il semblerait aussi que la fonction de luminosité ait peu évolué depuis z$\sim$0.8 (Rosati et al., 1998).

Un des rapports de la 'Dark Energy Task Force'[1] montre que l'étude de la distribution des amas, et en particulier de son évolution jusqu'à z=1-1.5, constitue l'une des méthodes les plus prometteuses en vue d'appréhender les propriétés de l'énergie sombre (Haiman et al., 2005).

Avec sa sensibilité exceptionnelle, sa bonne résolution (FWHM de $6''$ comparée à $25''$ pour le PSPC ROSAT) et son grand champ de vision, XMM a le potentiel d'accroître significative l'apport des amas à la cosmologie. Il restera l'observatoire X le plus sensible jamais construit pour au moins les 10 ans à venir. Les observations XMM de profondeur moyenne (bien au-dessus de la limite de confusion) à haute latitude galactique comprennent en majorité des AGs (ponctuels) et des amas de galaxies (étendus).

Dans ce contexte unique, XMM-LSS s'est fixé pour but d'étendre les études actuelles d'amas à un redshift de z$\sim$1.

Les études préliminaires ont montré que quelques centaines d'amas étaient nécessaires pour obtenir une précision de 20% sur la fonction de corrélation. Les contraintes sur cette dernière dans l'univers local (Collins et al., 2000) indiquent une distance comobile de corrélation de l'ordre de 20 $h^{-1}Mpc$. Cette échelle correspond à des ouvertures angulaires de respectivement 1° et 3° à des redshift de 1 et 0.15. Observer une corrélation significative sur cette gamme de redshift né-

[1] Cette commission a été rassemblée par la National Science Foundation américaine. Elle a pour but d'établir une analyse comparative des différentes approches permettant d'étendre notre compréhension de l'énergie sombre.

cessite donc de couvrir une échelle transverse de 5-10 deg. Ce type d'échelle est également requis pour réduire la variance cosmique à un niveau inférieur au bruit de Poisson.

Ces considérations ont conduit à la définition d'un relevé de plusieurs dizaines de deg$^2$ en assignant un flux limite de l'ordre de $1 \times 10^{-14}$erg.s$^{-1}$.cm$^{-2}$pour les amas de galaxies. Des observations de 10 ks suffisent pour atteindre une tel sensibilité avec XMM. Cette configuration permet d'accomplir les objectifs scientifiques énoncés tout en requierant un temps total d'observation qui reste dans les limites du temps allouable. La profondeur est 4, 10 et 1000 fois plus élevée que le RDCS, EMSS et REFLEX respectivement.

Un tel relevé permet d'obtenir des comptages cosmologiques d'amas jusqu'à z$\sim$1 et d'étudier systématiquement la population des groupes de galaxies jusqu'à z$\sim$0.5 (T=1-2keV, une population jamais détectée jusqu'alors). L'échantillon ainsi obtenu permettra enfin de tracer les structures à grande échelle de par la ditribution d'AGNs X jusqu'à z$\sim$4.

### 3.2.2 Location et couverture XMM

En vue du suivi multi-longueurs d'onde des amas détectés, un champ proche de l'équateur a l'avantage d'être accessible depuis les instruments des deux hemisphères. Pour éviter toute contamination, les hautes latitudes galactiques sont en outre préférables, de même que l'éviction des champs contenant des sources X brillantes. Le seul champ vérifiant toutes ces conditions et visible par XMM plus de 15% du temps fut choisit pour le relevé. Il a pour coordonnées centrales : $\alpha$ = 2h18m00s =et $\delta$ = -7° 00′00″. Une des régions du CFHTLS (W1) a été spécifiquement choisie pour recouvrir le XMM-LSS.

Cette position coincide avec deux autres relevés profonds XMM : le XMM Medium Deep Survey (XMDS) couvre $\sim$2 deg$^2$ avec des pointés de 20 ks, et le Subaru/XMM Deep Survey (SXDS) se compose de 6 pointés de 50 ks disposés autour d'un $7^{me}$ de 100 ks.

En y ajoutant les pointés propres au XMM-LSS, la surface totale couverte totalise 10.9 deg$^2$ (dont la Fig.3.9 montre la distribution sur le ciel). Les résultats présentés dans cette thèse n'en utilisent que 6 deg$^2$. Les données de l'AO5 d'XMM (un grand programme de 46 pointés) ont été obtenues plus tardivement et celles du SXDS, plus profondes, requièrent un traitement spécifique de la confusion des sources ponctuelles.

### 3.2.3 Le suivi multi-longueurs d'onde

Outre son propre suivi spectroscopique des sources détectées, le relevé XMM-LSS est lié à nombre de projets couvrant l'ensemble du spectre observable, depuis la radio basse fréquence (74 MHz) jusqu'aux rayons $\gamma$ (10 MeV).

Nous donnons ici une liste non exhaustive des instruments ou relevés qui couvrent ou prévoient de couvrir une partie du champ XMM-LSS :

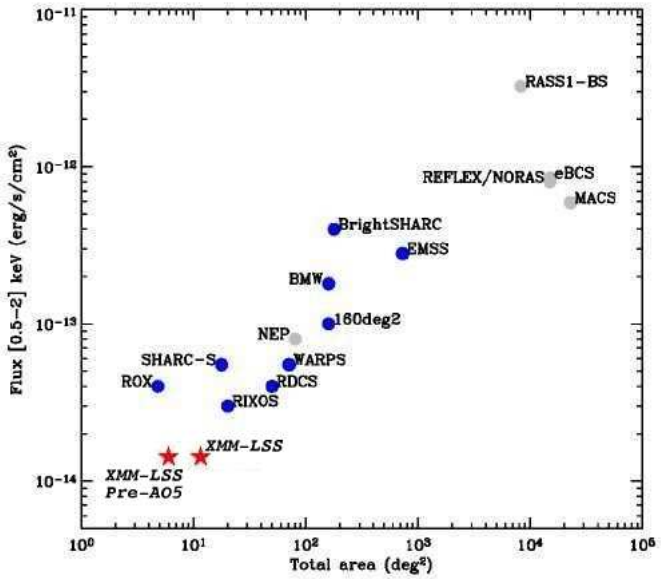


**FIG. 3.8**: *Profondeur et surface des grands relevés ROSAT et Einstein disponibles à ce jour. Les ronds gris représentent les relevés recouvrant une surface continue. Les autres utilisent des pointés d'archive. Pour comparaison, les étoiles rouges indiquent les caractéristiques du XMM-LSS.*

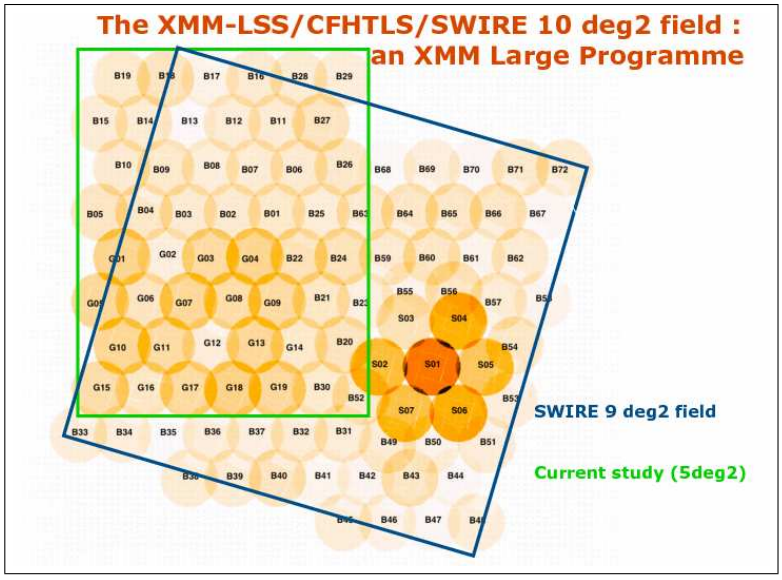


**FIG. 3.9**: *Disposition des 98 pointés du XMM-LSS, incluant le SXDS. L'échelle de couleur indique le temps d'exposition effectif, allant de 80 ks pour le pointé central du SXDS à ∼ 0.Le rectangle vert indique les données obtenues avant l'AO5 (sur lesquels se base cette thèse). Les observations supplémentaires obtenues au titre de l'AO5 permettent la couverture complète des 9 deg² du relevé SWIRE dans le champ XMM-LSS. Ce dernier est signalé par le carré bleu.*

- Radio : VLA, OCRA, AMIBA
- mm : APEX-SZ, SCUBA2
- IR : SWIRE, Herschel
- NIR : UKIDSS
- Optique : CFHTLS (W1), VIMOS (VVDS)
- UV : GALEX
- Gamma : INTEGRAL.

Le VLA, le CFHTLS et INTEGRAL couvrent l'ensemble du champ. Une couverture totale est aussi envisagée par Herschel (acceptée pour le temps garanti), SCUBA2 (programmée) et APEX-SZ. Enfin, chaque amas détecté en rayon X sera observé par OCRA.

Cette complémentarité permettra l'étude multi-longueur d'onde de la physique des amas ainsi que des galaxies d'amas. L'ensemble des observables connues pour les amas seront disponibles pour les détections du XMM-LSS : les lentilles gravitationnelles (à partir du CFHTLS), l'effet Sunyaev-Zel'dovich (OCRA, APEX-SZ, SCUBA2) et la photométrie des galaxies.

Une analyse détaillée des AGNs du XMM-LSS sera aussi possible.

# Chapitre 4
# Moyens mis en oeuvre

## Sommaire



## 4.1 Problématique

Pour le XMM-LSS, comme pour la plupart des relevés cosmologiques, une des tâches les plus ardues, mais fondamentale, consiste à définir un échantillon représentatif de la population réelle des sources étudiées (dans notre cas, les amas de galaxies et les AGNs). Malheureusement, de telles études sont toujours limitées par le potentiel technique des instruments, et il n'est généralement pas possible de prétendre à la détection de toute une classe d'objets sans introduire également de nombreuses fausses sources dans les catalogues. Les échantillons assemblés résultent alors d'un compromis entre complétude et contamination.

Dans le cas des données *XMM-Newton*, les difficultés liées à la détection de sources sont de natures multiples. Tout d'abord, l'inconvenient majeur est de devoir gérer le bruit de Poisson inhérent aux données : la majorité des sources ne se

traduisent sur les détecteurs que par une agglomération de quelques dizaines de photons noyés dans le fond. En vu de déterminer la nature des sources détectées (amas ou AGN), il faut être capable d'apprécier le caractère étendu ou ponctuel de la distribution de ces photons, dans des images où la sensibilité et la forme de la PSF varient fortement (voir section 3). Enfin, il arrive fréquemment que la détection des sources soient perturbée par la présence de trous entre les plaques CCD. Compte tenu du faible temps d'exposition des pointés, ces problèmes sont encore plus accentués dans le cas du XMM-LSS. Dans ces conditions, le développement, l'optimisation et l'évaluation du pipeline de traitement des données deviennent des tâches critiques. Cela nécessite notamment l'emploi de méthodes rigoureuses prenant en compte tous les effets instrumentaux.

Une fois l'échantillon constitué, son emploi en vue de contraindre les modèles cosmologiques ne sera possible que si les effets de sélection (complétude et contamination) associés aux algorithmes utilisés ont été précisement quantifiés. De plus, il convient de caractériser au mieux les propriétés des sources détectées afin de lever les dégénérescences intrinsèques aux modèles cosmologiques que l'on veut tester.

Par conséquent, une partie substantielle et nécessaire de ce travail de thèse a été le développement et la validation d'outils adaptés à ces objectifs. Le présent chapitre décrit et justifie les approches adoptées dans ce but.

## 4.2 Création d'images XMM simulées

En vue d'évaluer les performances des différents algorithmes, il est indispensable d'avoir recours à des observations simulées, afin de pouvoir comparer les résultats avec le contenu réel des données traitées. Pour ce faire, j'ai mis en place un système automatisé pour générer des images simulées dans la bande [0.5-2] keV (optimale pour la détection d'amas), sur la base du travail déjà effectué au CEA/Saclay par Ivan Valtchanov et René Gastaud.

### 4.2.1 Le logiciel `InstSimulation`

Les simulations sont créées par le programme `InstSimulation` (Valtchanov et al., 2001), développé en C++ dans le cadre des études de faisabilité du relevé XMM-LSS.

Ce logiciel requière en entrée un catalogue de sources à inclure dans l'observation (qui comprends la position, l'extension et le flux de chaque objet dans la bande considérée). Deux types de sources sont admis : des sources ponctuelles représentées par la PSF de l'instrument, et des sources étendues dont le profil est un modèle $\beta$ convolué avec la PSF.

Source par source, `InstSimulation` calcul le nombre de coups attendus sur chaque instrument EPIC en tenant compte de la surface effective dans l'axe optique des télescopes et d'une fonction de vignettage analytique[1]. Le simulateur incorpore

[1] le modèle utilisé est un ajustement polynomial d'ordre 4 de la surface effective en fonction de la distance hors axe. Pour les détecteurs MOS, on se contente de prendre en compte la moyenne azimuthale du vignettage.

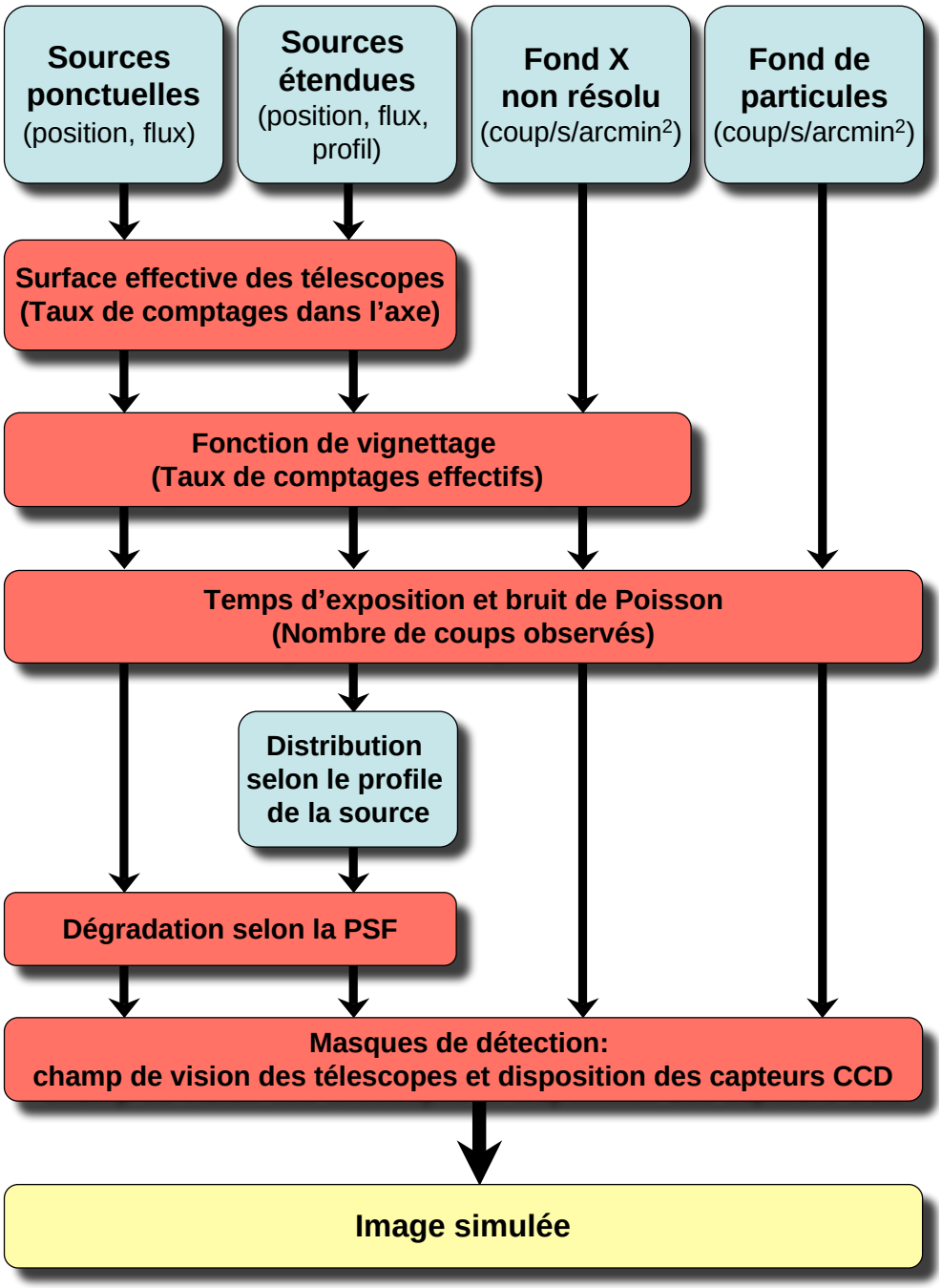


FIG. 4.1: *Fonctionnement schématique du logiciel* `InstSimulation`. *Les cases bleues indiquent les composantes astrophysique externes au programme (i.e celle qui sont fournies en entrée). Les cases rouges indiquent les effets instrumentaux pris en compte. Les flèches illustrent le déroulement du programme.*

**TAB. 4.1**: *Niveaux moyens de fond (en coups.$s^{-1}$.$arcmin^{-2}$) utilisé pour chaque détecteur EPIC lors de la production d'images simulées.*

| Détecteur | MOS1 | MOS2 | PN |
|---|---|---|---|
| **Fond diffus X (dans l'axe)** | $6.7 \times 10^{-4}$ | $7.6 \times 10^{-4}$ | $1.4 \times 10^{-3}$ |
| **Fond de particules** | $4.4 \times 10^{-4}$ | $4.2 \times 10^{-4}$ | $1.6 \times 10^{-3}$ |

alors à l'image une réalisation poissonienne de ce nombre de photons. Ces derniers sont redistribués un à un selon le profil de la source (pour les sources étendues), puis selon la PSF instrumentale[2]. `InstSimulation` ajoute par la suite un fond de rayon cosmique constant et le fond diffus X, dû aux sources non résolues et sujet au vignettage des télescopes. Pour finir, il applique les masques de détection des caméras EPIC.

Comparé aux autres logiciels de simulation disponibles pour l'obsevatoire XMM (`SciSim` et `QuickSim`), le champ d'application d'InstSimulation reste limité dans la mesure où il ne fonctionne que pour les détecteurs EPIC et génère directement des images et non des listes d'évènements (ceci le rends impropre aux simulations impliquant la spectroscopie ou l'analyse temporelle des photons). Cependant, il s'agit du logiciel le plus simple d'utilisation, et son exécution est quasi-instantanée.

Ma contribution personnelle à la version actuelle du programme tient à la mise à jour des modèles d'étalonnage d'XMM (PSF, vignettage, surface effective), à l'ajout de l'image du fond et de la carte d'exposition parmi les sorties possibles du programme, ainsi qu'à la création d'un mode de simulation '*in-situ*'. Ce dernier permet de spécifier en entrée les images et cartes d'exposition d'une observation réelle afin d'en extraire les informations spécifiques à ce pointé (axes optiques des télescopes, vignettage, ...) puis de simuler des sources directement dans les images pré-éxistentes. Le but est d'obtenir des simulations les plus réalistes possibles pour les amas de galaxies.

### 4.2.2 Modélisation du ciel en rayons X

Pour que les simulations soient réalistes, il faut que les données d'entrée du programme correspondent à des conditions physiques raisonnables. Lorsque l'utilisation du mode *in-situ* du simulateur se révèle peu éfficace (notamment pour évaluer la détection des sources ponctuelles), il sera nécessaire de disposer d'une description phénoménologique du ciel X bâtie sur des contraintes observationnelles.

Dans cette gamme spectrale et à haute lattitude galactique où la densité d'étoiles devient relativement faible, les AGNs représentent l'essentiel des sources (soit plus d'une centaine par pointé XMM en moyenne avec un flux supérieur à $10^{-15}$ erg.$s^{-1}$.$cm^{-2}$). Dans la bande [0.5-2] keV, leur densité et distribution en flux

[2]afin de pouvoir reproduire la forte asymétrie de la PSF, en bordure de champ, le modèle de PSF "MEDIUM", généré par le simulateur de trajectoires des photons `SciSim`, est utilisé pour toutes nos simulations

sont représentées fidèlement par la loi "*Log(N)-Log(S)*" de Moretti et al. (2003) :

$$N(>S) = \frac{1.09 \times 10^{-23}}{S^{1.82} + 1.34 \times 10^{-17}\ S^{0.6}}, \tag{4.1}$$

obtenue par l'analyse combinée d'observation profondes ROSAT, XMM et CHANDRA, où $N(>S)$ représente la densité projetée (par deg$^2$) de sources ponctuelles dont le flux excède la valeur S, et S est exprimée en erg.s$^{-1}$.cm$^{-2}$. De nombreuse études révèlent une correlation positive de la distribution spatiale des AGNs (e.g. Vikhlinin & Forman 1995, Carrera et al. 1998 et plus récemment Yang et al. 2003 et Basilakos et al. 2005). Cependant, l'amplitude de cette corrélation demeure suffisamment faible pour que l'on puisse la négliger en première approximation dans nos simulations. L'effet de la confusion des sources ponctuelles sera étudié indépendamment par la simulation de paires d'AGNs adjacents. Les autres populations de sources restent peu fréquentes et n'ont par conséquent pas d'influence significative sur la détection des amas et des AGNs.

Une seconde composante à considérer est le fond. Read & Ponman (2003) en ont effectué une étude détaillée pour XMM-Newton à partir d'observations archivées. Après l'élimination de toutes les sources de flux supérieur à $4\times10^{-15}$ erg.s$^{-1}$.cm$^{-2}$, ils ont constitué par empilement une carte profonde du fond diffus X et du fond de rayons cosmiques. Leur analyse a permis d'estimer le niveau moyen de ces deux composantes et de montrer qu'elles sont en première approximation homogènes dans la bande [0.5-2] keV (hormis la variation du fond astrophysique due au vignettage).

Néanmoins, la quasi-totalité du fond X diffus provient de sources ponctuelles non résolues (voir par exemple Worsley et al. 2004), ce qui implique de modéliser le fond et la population d'AGNs de manière cohérente. Cette dernière est simulée, dans le cas du XMM-LSS, jusqu'à un flux de $10^{-16}$ erg.s$^{-1}$.cm$^{-2}$(ce qui correspond pour une exposition de 10 ks à une moyenne de 1 photon détecté sur l'axe optique). Il faut donc corriger le niveau du fond diffus X de Read & Ponman (2003) de l'emission des AGNs de flux compris entre $10^{-16}$ et $4 \times 10^{-15}$ erg.s$^{-1}$.cm$^{-2}$. On estime leur contribution par intégration de la loi Log(N)-Log(S). Les valeurs de fond ainsi obtenues sont listées dans la Table 4.1.

En ce qui concerne les amas de galaxies, on empêchera la superposition des sources pour tenir compte de la faible densité de cette population (1 à 2 sources en moyenne par pointé XMM).

La Figure 4.2 présente une de ces simulations *ex nihilo*, accompagnée d'une observation réelle pour comparaison.

### 4.2.3 Mise en øeuvre du simulateur

Des scripts d'execution automatique permettent de générer à la volée un grand nombre d'images (soit totalement simulées, soit à partir d'une observation réelle), pour des configurations variées d'amas de galaxies (position, flux, extension). L'installation de ces procédures sur la ferme de PC du Centre de Calcul de l'IN2P3 à Lyon offre la possibilité de sonder, en quelques heures, l'ensemble de l'espace

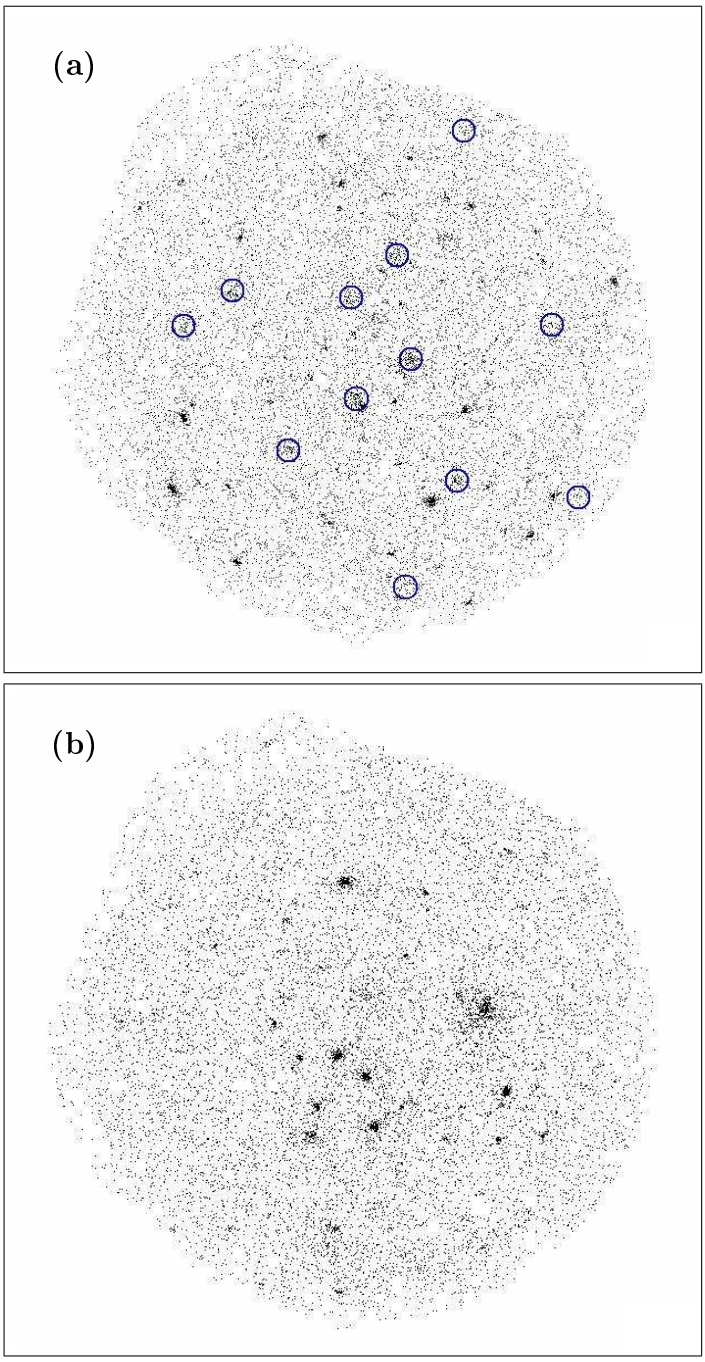


**FIG. 4.2**: *Exemples d'images XMM. (a) : Simulation synthétique. Les cercles bleus indiquent les amas injectés. Ceux-ci ont un taux de comptages de 0.02 cp/s, soit 200 photons détectés dans l'axe contre 30 à 50 en bordure de champ, lorsque l'amas tombe sur un trou des capteurs. Leur profil vérifie $r_c$=20" et $\beta$=2/3, et inclue 50% du flux total dans le rayon de 35" délimité par les cercles bleus. (b) : Observation réelle du XMM-LSS (dénommée B03 dans la Table 1 de Pierre et al. (2007)).*

des paramètres pour les amas dans une observation données. La mesure des performances et des effets de sélection associés à un algorithme donné devient ainsi possible dans un interval de temps raisonnable.

## 4.3 Construction des catalogues X

### 4.3.1 Sur la structure du pipeline de détection

Historiquement, la méthode la plus couramment utilisée pour la détection de sources dans des images X se base sur des algorithmes à cellules glissantes. Pixel par pixel, l'image est explorée à travers une boîte de taille fixe. Pour chaque position, on estime le fond local à partir des pixels environnant la boîte et le rapport signal sur bruit (S/N) d'une hypothétique source est calculé. Si le S/N excède un seuil fixé, on identifie le groupe de pixels correspondant comme une source. Une possibilité (employée notamment par la procédure `EBOXDETECT` développée pour XMM) est de considérer à la place du S/N la vraisemblance d'obtenir le flux mesuré dans la boite à partir d'une fluctuation de poisson du fond local.

Ces méthodes ont l'avantage d'être simples, rapides et robuste d'où leur utilisation très répandue en astronomie X. Cependant, à cause de l'estimation locale du fond, elles ne permettent de localiser que les sources dont la taille est proche de la largeur de la boîte utilisée. Cela les rend impropres à prendre en compte la large gamme d'échelles nécessaire à la détection de sources étendues dans des images contenant majoritairement des sources ponctuelles. En outre, les propriétés de l'instrument n'y sont pas rigoureusement modélisées, ce qui empêche une caractérisation et une classification adéquate des sources.

En vue de palier à ces défauts, les algorithmes récents effectuent désormais plusieurs itérations avec différentes tailles de boîtes, puis les profils des sources détectées sont ajustés à un modèle de PSF (éventuellement convoluée à un profil de source étendue) par la méthode du maximum de vraisemblance. Le pipeline officiel du satellite XMM fonctionne sur ce modèle : la procédure de détection par cellule glissante `EBOXDETECT` fournie une liste de source préliminaire puis le logiciel d'ajustement de profil `EMLDETECT` analyse en détails chaque détection. D'autres outils ne présentent pas les défauts des algorithmes de détection par cellule glissante : les plus utilisés sont les ondelettes (Rosati et al. 1995, Starck & Pierre 1998, Vikhlinin et al. 1998) et la tesselation de Voronoi suivie d'une percolation (Ebeling & Wiedenmann 1993, Scharf et al. 1997).

Dans la phase d'étude du XMM-LSS, Valtchanov et al. (2001) ont montré que les méthodes de filtrage d'images par ondelettes développées par Starck & Pierre (1998) donnaient de meilleurs résultats que les autres pour la détection de sources étendues dans les données XMM. En effet, elles incorporent naturellement le signal présent à toutes les échelles, contrairement aux algorithmes à cellule glissante, et elles ne tendent pas nécessairement à fusionner les sources voisines comme le font les algorithmes de percolations. Malgré cela, cette procédure ne prend pas en compte les caractéristiques instrumentales telles que la variation de la PSF dans le champ de vision des télescopes ce qui complique la classification des sources

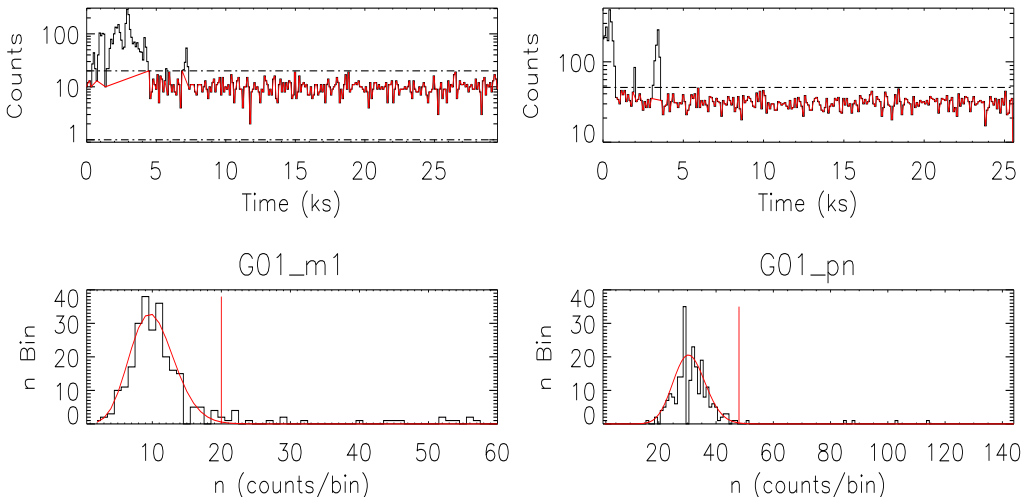


**FIG. 4.3**: *Filtrage des intervalles de temps exploitables du pointé G01 pour les détecteurs* `MOS1` *(gauche) et* `PN` *(droite). Bas : histogramme du nombre d'événements par intervalle de 104s ; la distribution de Poisson ajustée est tracée en rouge, et la coupure à* $3\sigma$ *correspondante est indiquée par le trait vertical. Haut : Courbe de lumière de l'observation par intervalle de 104s ; La coupure à* $3\sigma$ *est indiquée par le trait horizontal. La courbe rouge montre les périodes conservées après le filtrage.*

comme ponctuelles ou étendues.

En me basant sur ces résultats, j'ai constitué un pipeline reposant sur la sensibilité de l'algorithme ondelettes de Starck & Pierre (1998), mais qui inclut ensuite une procédure d'ajustement par maximisation de la vraisemblance pour la classification et la caractérisation des amas et AGNs. Le reste de cette section décrit en détail le fonctionnement et l'optimisation de ce pipeline.

### 4.3.2 Extraction des images

Les procédures inclues dans le XMM SCIENCE ANALYSIS SYSTEM (`XMM-SAS`) permettent d'obtenir directement des listes d'événements étalonnées à partir des fichiers de données brutes XMM. Celles-ci contiennent entre autre le temps d'arrivé, l'énergie, et la position sur les détecteurs de chaque particule observée. Des images et cartes d'expositions dans une bande d'énergie donnée peuvent ensuite être générées pour tout ou partie de chaque détecteur. Cependant, il est possible d'augmenter la qualité des images obtenues par un filtrage préalable des listes d'événements.

#### Sélection des intervalles de temps

Une source importante (et transitoire) de fond pour le satellite XMM-Newton provient de nuages de protons confinés dans la magnétosphère terrestre et qui s'engouffrent occasionnellement le long des miroirs des télescopes. Le signal reçu durant ces périodes est nettement plus élevé que la moyenne. Le niveau de fond

des images peut être considérablement réduit par l'identification et l'exclusion des intervalles de temps correspondants. Pour ce faire, le pipeline utilise la méthode présentée dans l'annexe A de Pratt & Arnaud (2002). Il s'agit de construire, pour chaque détecteur, l'histogramme du nombre d'événements enregistrés par intervalle de 104s dans les bandes dures ([10-12] keV pour les `MOS` et [12-14] keV pour `PN`) où la contribution relative des bouffées de protons est la plus importante. Ces histogrammes sont ensuite ajustés par une distribution de Poisson de moyenne $\lambda$. On considère les intervalles dans lesquels l'émission excède $\lambda + 3\sqrt{\lambda}$ comme contamminés (seuillage à $3\sigma$). La figure 4.3 montre un exemple de filtrage.

#### Sélection des motifs

Il arrive fréquemment qu'une même particule ou un même photon illumine plusieurs pixels adjacents sur les CCD. Selon le type de particule, les motifs caractéristiques obtenus sont différents. Bien que le fond de rayons cosmiques ne montre pas de fortes variations temporelles, on peut en exclure une fraction en supprimant les motifs qui ne proviennent généralement pas de photons. Dans le cas du XMM-LSS, on utilise seulement les événements simples et doubles pour le PN, tandis que l'on prend également en compte les événements triples et quadruples pour les MOS.

#### Extraction

La combinaison de la réponse instrumentale d'XMM avec la dépendance spectrale du fond X et la coupure exponentielle du spectre de brehmsstrahlung thermique, fait que le signal sur bruit pour la détection d'amas avec XMM est optimum dans les bandes d'énergies environnant 1 keV. Ainsi que le préconise Scharf (2002), on ne considerera que la bande [0.5-2] keV pour la détection d'amas. Il reste néanmoins utile d'analyser les images dans d'autres bandes pour l'étude des AGNs et pour construire des cartes de rapport de dureté. A partir des listes d'événements filtrées, des images sont donc extraites puis analysées dans chacune des bandes spectrales suivantes : [0.3-0.5] keV, [0.5-2] keV, [2-10] keV, [2-4.5] keV et [4.5-10] keV.

### 4.3.3 Ondelettes et bruit de Poisson

L'algorithme de détection de source est basé sur une prise en compte rigoureuse du bruit de Poisson dans l'espace des ondelettes. Dans l'article décrivant cette méthode, Starck & Pierre (1998) testent différentes techniques de filtrage sur des images X et démontrent la supériorité de leur approche dite par "autoconvolution d'histogramme" pour la reconstruction d'images débruitées.

#### Transformation en ondelettes

Les ondelettes ont été originellement introduites dans le but de permettre l'analyse temps-fréquence du signal. En effet, la transformation de Fourier, uti-

lisée généralement pour retranscrire l'amplitude d'un signal depuis l'espace des temps vers celui des fréquences, ne conserve pas d'information directe sur la localisaton temporelle des différentes composantes fréquentielles. Elle ne permet donc pas aisément l'étude des signaux non-stationaires. Dans le cas de l'imagerie astronomique, où l'on observe une grande dynamique d'échelles, pouvoir considérer simultanément la position et la dimension caractéristique des objet est crucial, d'où l'interet croissant pour les ondelettes dans ce domaine.

L'analyse en ondelettes consiste à remplacer le concept de fréquence par celui d'échelle ou résolution, et à associer, pour un signal donné $f(x) \in \mathcal{L}^2(\mathbb{R})$, une amplitude (le coefficient d'ondelette) à chaque point de l'espace position-échelle. Mathématiquement, une ondelette est une fonction $\psi$ de carré sommable (dite 'ondelette mère') dont les translations-dilatations forment une base de $\mathcal{L}^2(\mathbb{R})$. Le coefficient correspondant à une ondelette $\psi : \mathbb{R} \to \mathbb{R}$ pour l'échelle $a$ et la position $b$ est un simple produit scalaire (Morlet et al., 1982), soit dans le cas d'une transformation continue à 1 dimension :

$$W(a,b) = \frac{1}{\sqrt{a}} \int f(x)\psi\left(\frac{x-b}{a}\right) dx, \tag{4.2}$$

où l'intégrale porte sur le recouvrement des supports de l'ondelette $\psi(x)$ et de la fonction $f(x)$. La transformation en ondelette (tout comme la transformée de Fourrier) se réduit donc à une décomposition du signal sur une base présentant des propriétés particulières : $\{\psi(\frac{x-b}{a})\}_{(b,a)\in\mathbb{R}\times\mathbb{R}^+}$. Le coefficient en ondelette ainsi défini fournit une mesure du degré de conformité entre la forme locale du signal au point $b$, et l'ondelette dilatée à l'échelle $a$.

Une propriété primodiale de cette transformation est son inversibilité. Le signal d'origine s'exprime en effet à l'aide des coefficients d'ondelettes comme :

$$f(x) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{a}} W(a,b)\psi\left(\frac{x-b}{a}\right) \frac{da\ db}{a^2} \tag{4.3}$$

où $C_\psi = \int_0^{+\infty} \frac{\hat{\psi}^*(\nu)\hat{\psi}(\nu)}{\nu} d\nu$ et $\hat{\psi}$ est la transformée de Fourier de l'ondelette $\psi$. La formule d'inversion existe si et seulement si $C_\psi < \infty$, ce qui conduit à utiliser ce critère comme définition des fonctions d'ondelettes. Une condition nécessaire pour cela (et presque suffisante) s'exprime comme $\int \psi(x)dx = 0$. Il s'agit donc d'une fonction oscillante, d'où son nom d'ondelette.

En pratique, l'utilisation d'une transformation continue s'adapte mal à l'analyse d'images (dans lesquels le signal est déjà échantillonné). De plus, la formule de reconstruction 4.3 contient une somme d'informations très redondantes. On peut cependant construire des fonctions d'ondelettes dont les translations/dilatations discrètes génèrent une base de $\mathcal{L}^2(\mathbb{R})$. Les plus employées sont les ondelettes dyadiques :

$$\{2^{-j/2}\phi(2^{-j}x - k)\}_{j,k\in\mathbb{Z}}, \tag{4.4}$$

auxquelles nous nous limitons dans ce mémoire.

### Modélisation du bruit de Poisson

Pour filtrer des images dans l'espace ondelettes, on cherche à déterminer parmi les coefficients d'ondelette les fluctuations locales du signal qui sont significatives. Dans ce but on définit sur la transformée en ondelettes des données un test de signification statistique basé sur la probabilité P d'obtenir les coefficients d'ondelettes observés en l'absence de signal (i.e. pour une distribution spatiale uniforme) :

$$P \quad \begin{cases} W(a,b) > 0 : & P = Prob(W_N > W(a,b)) \\ W(a,b) < 0 : & P = Prob(W_N < W(a,b)). \end{cases} \tag{4.5}$$

Si $P$ dépasse une valeur seuil $\varepsilon$, on considère que l'obtention des données par une simple fluctuation du fond ne peut être exclue. La valeur des coefficients d'ondelettes correspondants est donc annulée. On reconstruit ensuite itérativement l'image débruitée à partir des coefficients significatifs.

La densité de probabilité associée, dans l'espace des ondelettes, à une variable aléatoire de distribution spatiale uniforme découle directement de l'équation 4.2 : il s'agit de l'ondelette elle-même. Le calcul exact de la probabilité P dépend par contre du type de bruit inhérent aux données.

Dans le cas des images X, les N photons inclus dans le support de l'ondelette à une position et une échelle donnée peuvent être considérés comme la réalisation d'une somme de N variables aléatoires de distribution identique. La densité de probabilité correspondante s'écrit alors comme la convolution des densités de probabilité individuelles. Pour illustrer ce point, considérons deux variables aléatoires indépendantes $x$ et $y$ prenant des valeurs entières ; la probabilité pour que $x = m$ et $y = n$ simultanément est donc $p(m,n) = p_x(m)p_y(n)$. La probabilité associée à la variable somme $z = x + y$ s'obtient alors par sommation sur l'ensemble des couples $(m,n)$ correspondants :

$$p_z(k) = \sum_{m=-\infty}^{\infty} p_x(m)p_y(k-m).$$

Ce résultat s'étend au cas des variables continues et la somme devient une convolution.

L'autoconvolution de l'ondelette utilisée pour la transformation permet donc de déterminer la distribution attendue $W_N$ des coefficents d'ondelette dans le cas où les structures observées sont dues au bruit. Notons que cette distribution tend vers une loi normale quand le nombre de photon dans le support de l'ondelette devient grand.

### Mise en øeuvre

En pratique, le pipeline XMM-LSS utilise la procédure `mr_filter` de la bibliothèque MR/1 (Starck et al., 1998). Ce programme utilise l'algorithme 'à trou' pour la transformation en ondelette. Afin d'augmenter le rapport signal sur bruit, et limiter les artefacts dus aux trous entre les CCD (non pris en compte par le logiciel), le filtrage s'oppére sur une image co-additionnée des trois detecteurs MOS1,

MOS2 et PN. On fixe le seuil de filtrage $\varepsilon$ du bruit à $10^{-3}$, ce qui correspondrait pour un signal gaussien à ne conserver que les détections excedant $\sim 3.1\sigma$. Plus de détails sur les ondelettes et l'algorithme utilisé sont donnés dans l'Annexe A.

### 4.3.4 Extraction des sources

La procédure `mr_filter` permet un filtrage rigoureux des images, mais n'en extrait pas les sources. Afin d'obtenir un catalogue de détections préliminaires, on execute alors le logiciel `SExtractor` sur les images débruitées. Celui-ci fonctionne essentiellement en 4 étapes :

- Il génère tout d'abord une carte de fond. Pour cela, une image à basse résolution est estimée sur un ensemble de cellules par réjection itérée à $3\sigma$ autour de la valeur médiane. Un filtrage médian peut ensuite être appliqué, puis la carte finale s'obtient par interpolation bilinéaire entre le centre des cellules.
- Les groupes de pixels adjacents qui excèdent un certain seuil au-delà du fond sont alors identifiés. Le seuil utilisé correspond à un facteur constant par rapport au R.M.S de la carte de fond.
- Si nécessaire, `SExtractor` fragmente ces détections par l'identification de maxima locaux dont la contribution au flux total dépassent une fraction prédéfinie.
- Enfin, le logiciel effectue une analyse géométrique et photométrique des sources.

`SExtractor` a été développé pour traiter des images optique dans lesquelles la quantité d'information est bien supérieure aux données X (statistique plus importante qui recouvre toute l'image). On ne peut donc l'utiliser directement sur les images X brutes qui contiennent très peu de photons. Cependant, il est tout à fait adéquat pour analyser les images filtrées en ondelette, à condition d'utiliser des réglages appropriés. La mesure du fond notamment doit tenir compte à la fois de la densité de sources, de la largeur de la PSF ainsi que des structures propres au fond lui-même. Pour cela, on fixe la taille des cellules de fond à une valeur intermediaire de 64 pixels de côté (i.e.$\sim 2.7'$) pour ne pas être biaisée par la présence de sources ponctuelles, mais rendre compte des variations à plus grande échelles du fond. Par la suite, on applique un filtrage median sur 4x4 cellules pour éviter d'inclure dans le fond une fraction du flux des amas (étendus et parfois brillants). En, ce qui concerne la détection, la quasi absence de bruit dans les images filtrées impose un seuil de détection élevée ($6\sigma$). De même le seuil en flux pour la segmentation des sources n'est que de 0.3% car le filtrage a déjà permis de déterminer que les sous-structures résiduelles étaient significatives.

La figure 4.4 montre les détections `SExtractor` superposées sur les images filtrées en ondelettes pour les simulations de la figure 4.2.

### 4.3.5 Le logiciel Xamin

Afin de prendre en compte en détail les effets instrumentaux, chaque source détectée par `SExtractor` est ensuite examinée par une procédure d'ajustement au

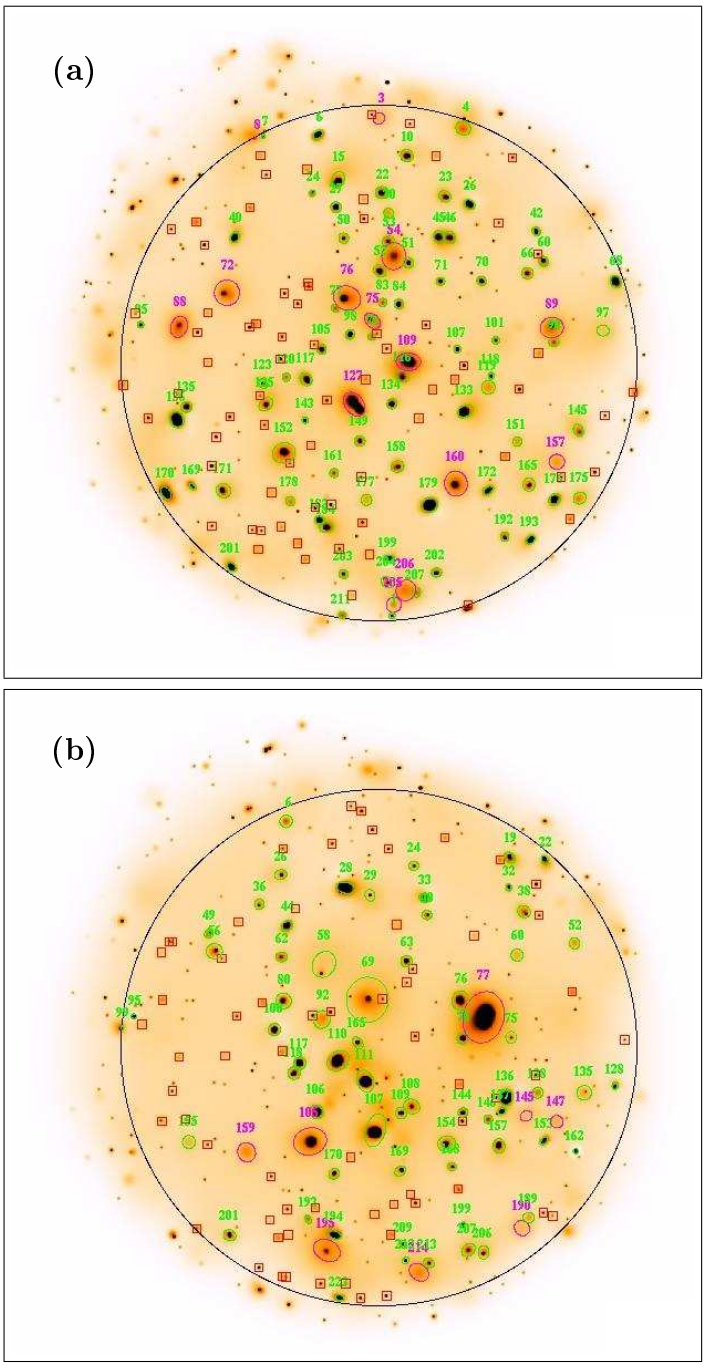


**FIG. 4.4**: *Filtrage en ondelette et détections SExtractor sur les images de la Fig. 4.2. Les sources les plus succeptibles d'être étendues, en vertu de l'extension et de l'indice de stellarité estimé par SExtractor, sont indiquées en magenta. Les carrés rouges correspondent aux détections non-physiques (largeur à mi-hauteur inférieure à 2.7″, soit moins de la moitié de la PSF d'XMM) ; celles-ci sont éliminées avant l'ajustement du profil des sources décrit Section 4.3.5. Le cercle noir dénote les limites du masque utilisé pour la détection (13′ de l'axe optique moyen des 3 télescopes).*

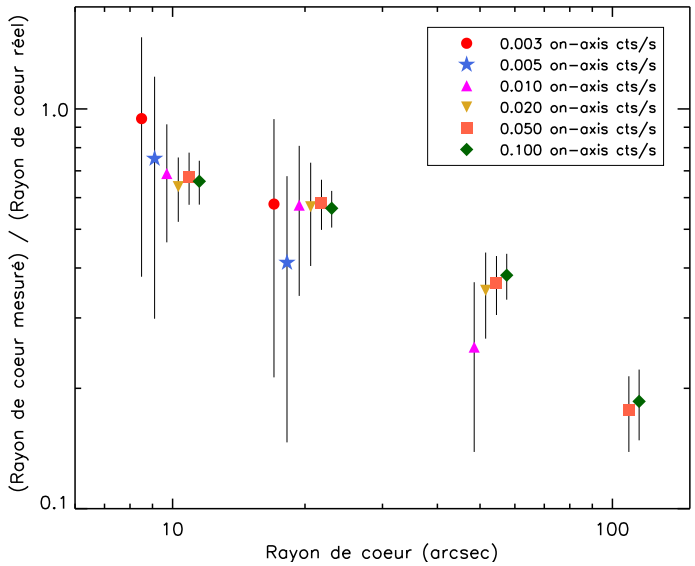


**FIG. 4.5**: *Caractérisation par* `EMLDETECT` *de la taille de sources étendues simulées. Le rayon de coeur est systematiquement sous-estimé, en particulier pour les sources les plus étendues. Ce résultat reste inchangé même en fournisssant comme paramêtre d'entrée au logiciel la carte exacte du fond simulée, ainsi que les valeurs réelles du flux et du rayon de coeur des sources comme valeurs de départ pour l'ajustement.*

maximum de vraisemblance. On pourrait pour cela utiliser le programme `EMLDETECT`, développé spécifiquement pour XMM. Nos premières simulations ont cependant montré que la version 4.27 disponible alors donnait de très mauvais résultats pour l'analyse des sources étendues (voir à ce propos la figure 4.5). J'ai donc dû développer une nouvelle procédure d'ajustement au maximum de vraisemblance : `Xamin`.

### Modèle de vraisemblance

Les photons observés pour un source X donnée correspondent à une réalisation de Poisson de son profil d'emission. La statistique habituellement utilisée pour ajuster ce type de données est la statistique de Cash (où C-statistique, Cash 1979) :

$$C = 2 \sum_{i=1}^{N_{pix}} \left( M_i - Y_i \ln M_i \right), \tag{4.6}$$

où $Y_i$ et $M_i$ sont respectivement le nombre de photons observés dans le pixel $i$ et le nombre prédit par le modèle de source testé. Il s'agit (à une constante additive et un facteur -2 près) du logarithme de la vraisemblance obtenue pour une distribution de Poisson dans chaque pixel de l'image. La C-statistique s'identifie donc à un $\chi^2$ lorsque le nombre de photons par pixel devient grand.

Un des objectifs essentiels d'`Xamin` est de donner des résultats non biaisés pour les sources étendues faibles, afin de permettre une classification propre de nos détections. Il apparaissait donc nécessaire à la fois d'inclure le niveau de fond dans les composantes libres, mais aussi de réduire au maximum le nombre de paramètres ajustés. Dans ce but, une version modifiée de la C-statistique a été introduite. Dans le cas d'une image, le modele d'émission $M_i$ peut être paramétré comme le produit d'une normalisation $N_{mod} = \sum_{i=1}^{N_{pix}} M_i$ et d'une distribution spatiale $d_i = M_i/N_{mod}$ indépendantes. En introduisant cette notation dans la C-stastique et en minimisant analytiquement sur la normalisation du modèle, on obtient la valeur ajustée $N_{mod} = \sum_{i=1}^{N_{pix}} Y_i$. La nouvelle statistique pour $d_i$ s'écrit alors simplement :

$$E = -2 \sum_{i=1}^{N} (Y_i \ln d_i) \,, \tag{4.7}$$

Ce formalisme permet d'ajuster à la fois le flux de la source et le niveau du fond en ne considérant qu'un seul paramètre libre : leur contribution relative au flux total observé.

### Modèle de source

Le modèle de source utilisé par `Xamin` a 5 paramètres libres : sa position dans l'image (R.A.,Dec.), son extension ($R_c$), son taux de comptage sur les detecteurs MOS ($CR_{MOS}$ et son taux de comptage sur le PN ($CR_{PN}$). Le choix d'identifier les taux de comptages des deux MOS a encore pour but de minimiser le nombre de paramètres libres de l'ajustement. Ceci est possible en première approximation dans la mesure où les deux capteurs sont de fabrication identiques et les tests d'étalonage ont révélé des performances quasi-similaires. En revanche, il n'était pas possible de lier de manière générale les taux de comptages MOS et PN : leur réponse spectrale est sensiblement différente et, pour un flux donné, le rapport entre les deux taux de comptage dépend du spectre de la source observée.

A partir de ces paramètres, `Xamin` crée une carte d'émission de la source en convoluant avec la PSF un modèle $\beta$ de rayon de coeur $R_c$ et de pente $\beta = 2/3$. Dans le cas des sources ponctuelles où $R_c = 0$, le model se réduit au profil de la PSF. Pour chaque détecteur, le modèle d'émission est ensuite normalisé aux taux de comptages intégrés que l'on veut tester . Pour cela, la fraction du flux située en dehors de la région ajustée est prise en compte. Pour finir, chaque carte d'émission est multipliée par la carte d'exposition correspondante. Les cartes de photons ainsi générées prennent en compte les principaux effet instrumentaux (PSF, vignettage, masques de détection, pixels mort ou chaud, efficacité quantique des CCD et transmission des RGAs).

Afin d'injecter ce modèle de source dans la statistique $E$, on ajoute un fond uniforme de manière à satisfaire la condition $N_{mod} = \sum_{i=1}^{N_{pix}} Y_i$. L'image finale est alors normalisée pour obtenir la distribution $d_i$.

### Mise en øeuvre

L'analyse photométrique et isophotale de `SExtractor` permet d'avoir une estimation préliminaire du centre, de l'extension et du flux des détections pour démarrer l'ajustement.

Source par source, `Xamin` extrait des données brutes un sous-échantillon dans une cellule entourant la position initiale. Les régions contaminées par d'autres sources sont masquées à l'aide de la carte de segmentation produite par `SExtractor`. Des tests préliminaires sur nos images simulées ont mis en évidence la necessité d'adapter la taille de boîte utilisée au profil de chaque source. La taille de la cellule est alors fixée à 6 fois l'extension isophotale mesurée par `SExtractor`.

L'ajustement du modèle de source aux données utilise l'algorithme de minimisation `AMOEBA` (Press et al., 1992). Bien qu'il ne soit pas parmi les plus rapides, celui-ci à l'avantage d'être peu sensible aux minima locaux et donc plus robuste dans les cas où la surface de vraisemblance est irrégulière (une propriété particulièrement souhaitable pour traiter des données poissoniennes à bas signal sur bruit). La procédure commence par fixer l'extension de la source à 0 pour ajuster un distribution ponctuelle. Puis la vraisemblance est minimisée pour le modèle complet, en utilisant l'extension estimée par `SExtractor` comme valeur de départ. La différence des minima de log-vraisemblance entre les deux ajustements permet de définir un rapport de vraisemblance (ou plus exactement, son logarithme) EXT_LIKE caractérisant la probabilité que la source considérée soit étendue. De la même manière, on obtient pour chaque détection un rapport de vraisemblance DET_ML en comparaison le meilleur ajustement avec un modèle de fond uniforme (i.e. sans source).

L'ensemble des paramètres mesurés pour les analyses ponctuelles et étendues sont inclus dans le catalogue de sortie. Celui-ci comprend également des informations complémentaires comme : le temps d'exposition effectif à la position de la source, la distance du trou de CCD le plus proche et celle du plus proche voisin.

### Performances

La validation puis l'optimisation des différentes composantes algorithmiques d'`Xamin` se fonde sur les simulations générées avec `InstSimulation`. Ces réglages furent facilités par la grande quantité d'information fournie dans les catalogues de sortie du logiciel. Pour comparaison, `EMLDETECT` se base sur une valeur seuil de vraisemblance d'extension pour ne retourner que les paramètres correspondant soit au meilleur model de source ponctuelle, soit à celui de source étendue.

Comme le montre la figure 4.6, `Xamin` retourne une estimation fiable des caractéristiques des sources étendues. L'expérience aquise lors du développement de ce logiciel permet d'identifier trois points critiques dans l'algorithme d'`EMLDETECT` qui semblent expliquer les mauvais résultats de la figure 4.5 :

– L'analyse précise du profil d'émission des sources étendues nécessite une cellule suffisament grande pour inclure les régions externes, dominées par le fond. De ce point de vue, l'utilisation d'une taille de boîte fixe est mal adaptée à l'analyse d'image comprenant une grande dynamique d'échelle.

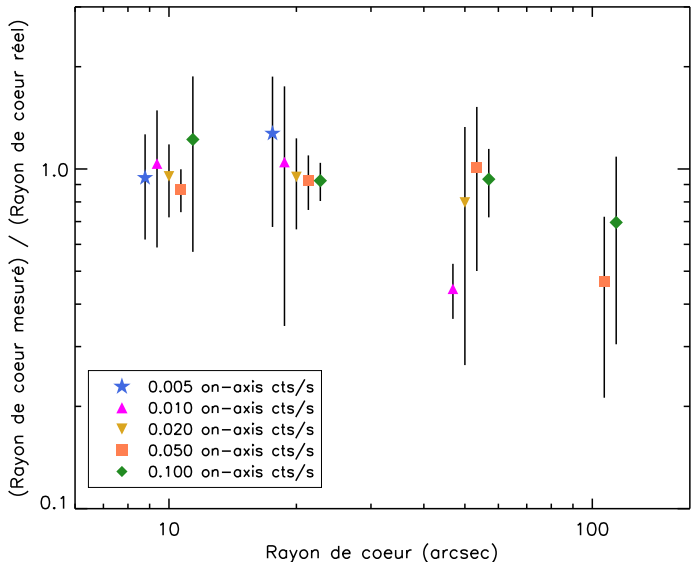


**FIG. 4.6**: *Caractérisation par* `Xamin` *de la taille de sources étendues simulées. Le biais inhérent à* `EMLDETECT` *semble ne pas affecter* `Xamin`*. Seul les sources très étendues (à partir d'un rayon de coeur de 100") voient leur extension légèrement sous-estimée.*

L'utilisation de larges boîtes pour l'ensemble des sources, en vue d'optimiser les résultats pour les sources étendues, crée en effet des problèmes de confusion lors de l'analyse des sources ponctuelles. Cela engendre de plus un temps d'exécution prohibitif.

- Pour gérer la confusion, `EMLDETECT` cherche à fragmenter la source qu'il reçoit en entrée et, dans le cas où l'ajustement s'en trouve amélioré, ajoute de nouvelles détections au catalogue. Les sources étendues sont régulièrement divisées par ce processus. Cela est d'autant plus vrai si le catalogue d'entré fourni à `EMLDETECT` provient de l'algorithme de cellule glissante `EBOXDETECT` qui a lui même tendance à fragmenter les sources étendues. L'algorithme d'`EMLDETECT` permet également, en principe, de fusionner des sources se situant à l'intérieur de la même boîte. Cependant, l'expérience montre que l'opération inverse est généralement favorisée.
- L'estimation du fond pose toujours problème dans l'étude des sources étendues. Les défauts actuels du logiciel `EMLDETECT` ne semblent pas provenir principalement du fond utilisé, dans la mesure où les résultats sur nos images simulées restent biaisées même lorsqu'on lui fourni la carte exacte du fond. Néanmoins, les mesures restent fortement dépendantes de ce paramètre d'entrée et il faudrait donc l'ajuster simultanément.

Le pipeline développé pour le relevé XMM-LSS a en comparaison l'avantage de reposer sur l'analyse multi-résolution pour déterminer l'échelle pertinente à l'étude de chaque source. Celle-ci est également utilisée pour gérer la confusion via l'utilisation du masque de segmentation. Le recour aux ondelettes permet donc de

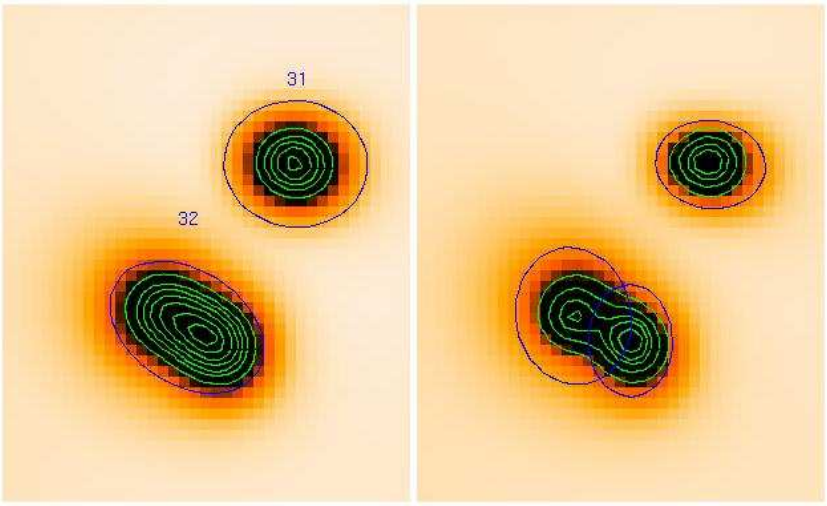


**FIG. 4.7**: *Illustration des effets de confusion des sources ponctuelles. L'image de gauche et les contours verts montrent le résultat du filtrage en ondelettes sur des images XMM simulées, pour un champ contenant un AGN seul, et une paire de sources ponctuelles séparées par 20″. Le seuil de fitrage est fixé à $\varepsilon = 10^{-4}$, et l'échelle 2 est exclue. Les détections successives de* `SExtractor` *sont indiquées par les ellipses bleues. L'image de droite montre l'analyse du même champ en utilisant l'échelle 2 et $\varepsilon = 10^{-3}$. Les deux sources ponctuelles sont maintenant correctement séparées.*

corriger l'essentiel des défauts apparents d'`EMLDETECT`. Notons malgré tout qu'une telle utilisation de la carte de segmentation n'est envisageable que pour les champs peu denses. Dans le cas d'observations très profondes, la confusion rend nécessaire l'incorporation d'un ajustement multi-sources et donc une confrontation directe aux problèmes que nous venons de mentioner.

Enfin, l'algorithme d'`Xamin` incorpore l'ajustement du niveau de fond tout en utilisant un modèle statistique qui limite les dégénérescences associées.

### 4.3.6 Validation du pipeline

#### Confusion et contamination

Il se révèle souvent difficile de prévoir l'efficacité d'un algorithme complexe sur la base des éléments qui le composent. Par conséquent, les paramètres libres du pipeline ont été réglés, in fine, en considérant le résultat de l'ensemble de la chaîne de détection.

Un exemple intéressant de ce type d'optimisation concerne la confusion des sources ponctuelles. Comme `Xamin` ne peut pas fragmenter les sources, il est critique de vérifier que `SExtractor` puisse traiter correctement les problèmes de confusion de sources ponctuelles. Pour ce faire, j'ai simulé des pointés ne contenant que des paires d'AGNs avec des séparations allant de 10" à 30". Comme l'illustre

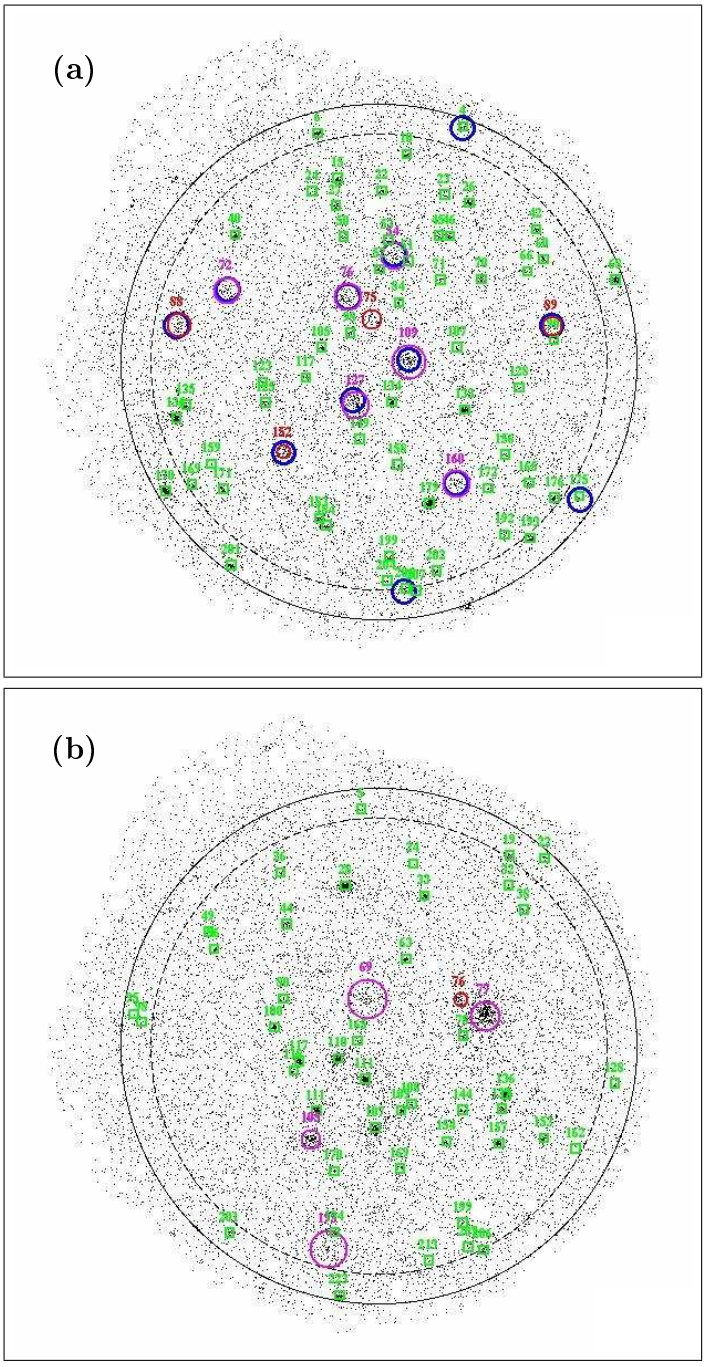


**FIG. 4.8**: *Detections Xamin pour les images de la Fig.4.2, superposées sur les images photons. Les carrés verts correspondent aux sources ponctuelles détectées, tandis que les cercles magentas et rouges indiques les respectivement les amas C1 et C2. Les cercles bleus rappellent la position d'entrée des amas simulés. Enfin les cercles noirs pleins et en tiret délimitent les distances hors-axe maximales autorisées pour la détection par SExtractor, et la sélection des sources étendues avec Xamin. Les sources étendues détectées dans le pointés B03 sont inspectées plus en détails dans la Fig. 4.9.*

la figure 4.7, ces simulations ont clairement montré la nécessité, lors du filtrage en ondelettes, d'utiliser à la fois un faible seuil de détection ($\varepsilon = 10^{-3}$) et d'inclure les plus petites échelles dans le filtrage en ondelette (échelle 2 contenant le signal entre 2 et 4 pixels, soit 5 à 10″). Ce choix de paramètres génère de nombreuses structures indésirables à petite échelle, dues uniquement au bruit de Poisson. Il ne s'agit donc pas d'un réglage optimal si l'on s'en tient aux images filtrées. Cependant, toutes ces fluctuations sont par la suite éliminées sans problème par Xamin et le résultat final en est finalement amélioré.

### Construction des échantillons

Les paramètres pertinents pour la construction d'échantillons propres à `Xamin` sont : le flux et la vraisemblance de détection (pour l'élimination des fausses détections), ainsi que l'extension et la vraisemblence d'extension (afin de distinguer les sources étendues et ponctuelles).

Dans le cas du XMM-LSS, où on decide de travailler à bas signal sur bruit, la détermination du flux réel d'une source devient difficile. L'utilisation d'une limite en flux observé n'a alors plus vraiment de sens physique, en particulier à cause du biais d'Eddington (Eddington 1913, parfois improprement appelé biais de Malmquist). La vraisemblance de détection calculé par `Xamin` est par ailleurs fortement corrélée au flux réel des sources simulées (cf. Fig. 4.10) ce qui permet de penser qu'elle est un meilleur indicateur de la qualité d'une détection. Ces critères nous conduisent à ne considérer que la vraisemblance de détection et non le flux pour éliminer les fausses détections. Pour cela, on détermine une valeur seuil à partir des simulations contenant uns distribution réalistes d'AGNs - comme la densité d'amas est faible, ceux-ci peuvent être négligés dans ce contexte. Un seuil de 15 en vraisemblance de détection assure un taux de fausses détections de l'ordre de une par pointé (à comparer avec 45-50 AGNS détectés par pointé en moyenne).

D'une manière générale, on ne tentera de détecter des sources étendues que pour des distances hors-axe inférieures à 11.5′. En effet, les distortions de la PSF en bord de champ compliquent la classification des sources, alors que le recouvrement des pointés du XMM-LSS permet en général d'ignorer complètement cette partie des pointés. Par ailleurs, à moins de se restreindre aux détections les plus significatives, il n'est également pas possible de définir une séparation claire entre les sources étendues et ponctuelles dans l'espace de paramètre d'`Xamin` (DET_ML, EXT_ML, EXT). Pour cette raison, j'ai introduit un système de classes de sources étendues correspondant à différents taux de contamination (essentiellement par des sources ponctuelles).

- La classe 1 (C1) est définie comme la région maximale de l'espace des paramètres totalement dépourvue de contamination. Elle répond aux critères DET_ML>32, EXT_ML>33, et EXT>5". Comme le caractère étendue de ces sources est sans ambiguité, il s'agit en général des amas de galaxies (et galaxies proches) les plus brillants.
- La classe 2 (C2) correspond à une contamination controlée. Il s'agit de la région maximale pour laquelle la proportion de fausses détections est infé-

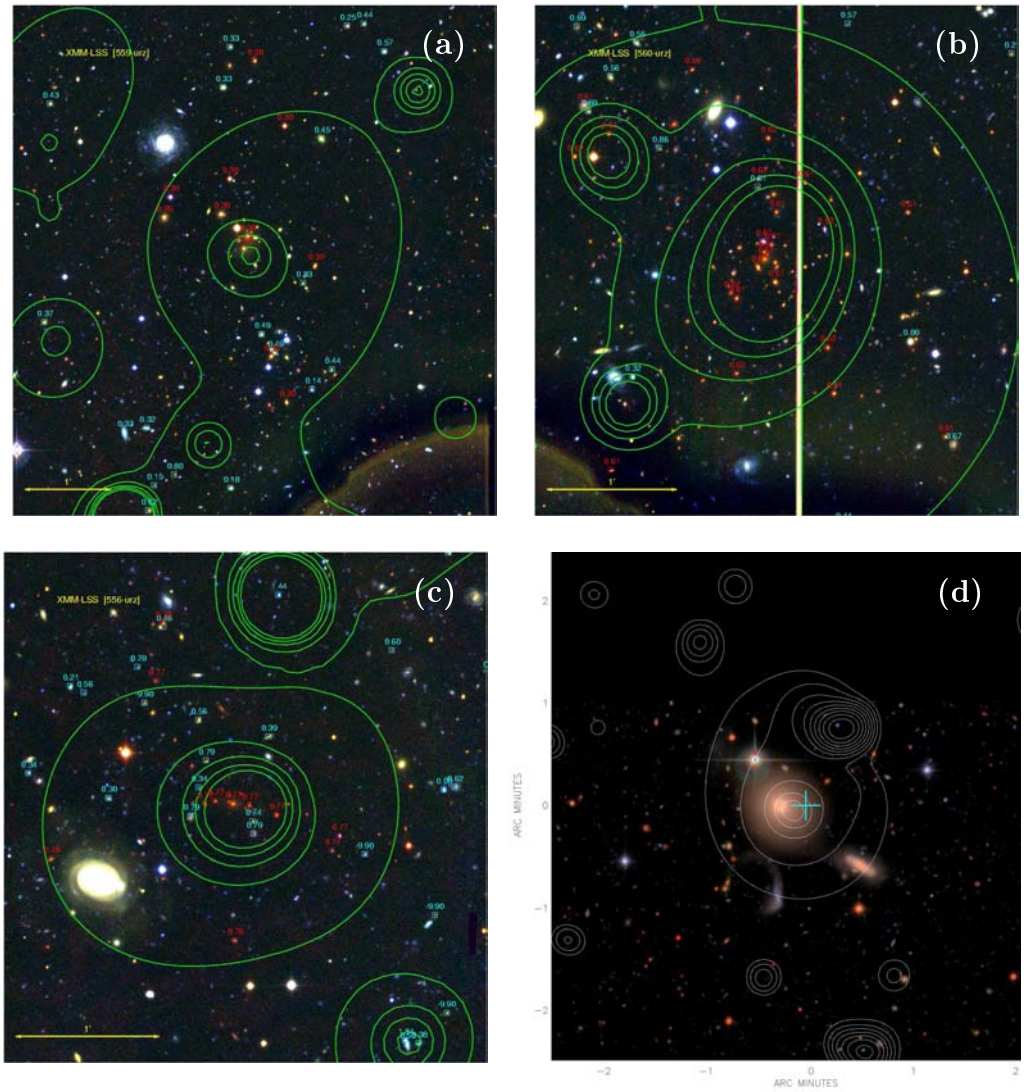


**Fig. 4.9**: *Détections C1 du pointé B03. Il s'agit d'une observation particulièrement riche en amas : on compte en moyenne moins d'une détection C1 par pointé XMM. Les contours de l'image filtrée en ondelette sont superposés à des images trois couleurs du CFHTLS. Les redshift mesurés spectroscopiquement dans chaque champ sont indiqués par des cercles rouges pour les galaxies membre de l'amas et des carrés bleus pour les autres. (a) : XLSSC 008 à z=0.30. (b) : XLSSC 001 à z=0.61. L'unique détection C2 du pointé correspond à la sous-structure située au nord de l'amas. Compte tenu de l'absence d'une contrepartie optique évidente, il est difficile d'affirmer si l'émission provient de l'amas ou d'un AGN dans la ligne de visée. (c) : XLSSC 002 à z=0.77. (d) : Galaxie proche en intéraction. Bien qu'il ne s'agisse pas d'un amas, le caractère étendu de l'émission X est indiscutable ce qui justifie sa classification comme C1.*

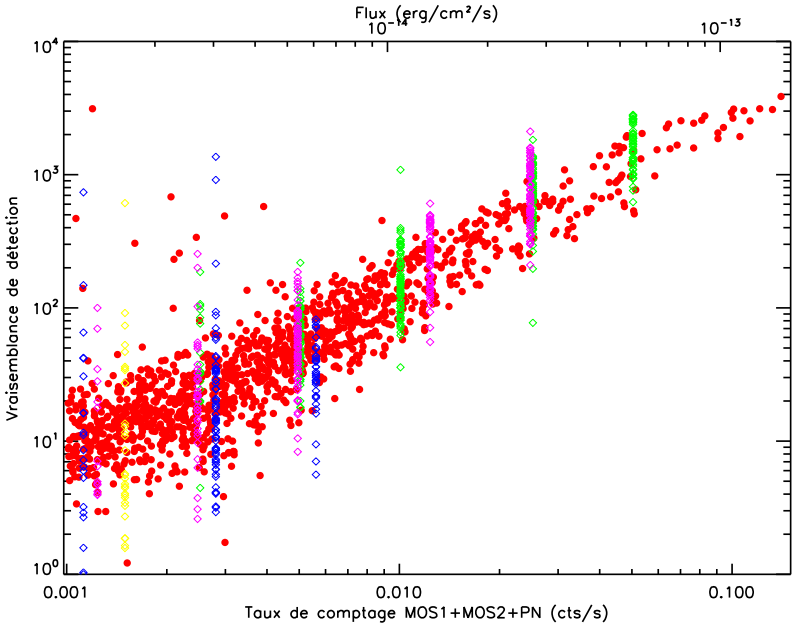


**FIG. 4.10**: *Corrélation entre la vraisemblance de détection estimée par* `Xamin` *pour l'ajustement de source ponctuelle et le flux réel dans une boite de* $35'' \times 35''$. *Les sources ponctuelles sont représentées par des ronds pleins ; Les losanges correspondent à des sources étendues avec* $\beta$*=2/3 et différents rayons de coeur :* $10''$ *(vert),* $20''$ *(magenta),* $50''$ *(bleu) et* $100''$ *(jaune).*

rieure à 50%, définie par EXT_ML>15 et EXT>5″. Le nettoyage de la contamination résiduelle nécessite alors une analyse ultérieure des sources, généralement par la confirmation de l'amas en optique (soit spectroscopiquement, soit à l'aide de redshifts photométriques). La constitution de cet échantillon demande donc plus de travail, de ressources et de temps.

– La classe 3 (C3) contient tous les amas confirmés en optique et associés à de l'émission X qui ne remplissent pas les critères C1 ou C2. Cet échantillon ne sera pas cosmologiquement exploitable dans la mesure où sa fonction de sélection reste indéterminée. Il rassemble cependant des sources potentiellement interessantes (telles que des amas à grand redshift) et permet de classifier certains amas étudiés dans les phases préliminaires du XMM-LSS, avant que les critères de sélection ne soient rigoureusement définis.

## 4.4 Analyse détaillée des amas de l'échantillon

Pour la détermination des paramètres cosmologiques comme pour l'étude de la physique des amas, les contraintes apportées par les grands relevés dépendent directement du nombre de sources considérées. L'information obtenu sera néanmoins d'autant plus complète que les propriétés des structures détectées auront été rigoureusement évaluées. De plus, l'existence de tels échantillons permet de

tester l'impact des effets de sélection sur la compréhension de ces objets.

Cette section traite donc des différentes méthodes utilisées afin de déterminer les propriétés physiques des amas du XMM-LSS.

### 4.4.1 Mesure des redshifts

Une étape majeure pour caractériser les amas de notre échantillon est d'estimer leur redshift.

Pour ce faire, la séquence rouge de chaque amas a été mesurée quantitativement à l'aide d'image CTIO et estimée qualitativement sur des images du CFHTLS. La couleur de la séquence permet une estimation préliminaire du décalage spectrale, ainsi que la sélection des galaxies membres les plus probables.

Ces candidats sont ensuite observés par spectroscopie multi-fente. Compte tenu des contraintes de positionnement des fentes et de la forte densité de galaxies près du centre des amas, cela nécessite en général deux ou trois masques. Les fentes restantes furent utilisées, lorsque cela était possible, pour observer la contrepartie optique des AGNs X présent dans le champ. Le redshift des galaxies observées a été mesuré par corrélation croisée avec des spectres types de galaxies elliptiques d'amas. Ces résultats ont par ailleurs été vérifiés manuellement lorsque des raies d'absorption ou d'émission non-ambigues permettaient une estimation évidente du décalage spectral.

Le critère utilisé pour la validation d'un amas est la mesure de redshifts concordants pour trois galaxies situées à moins de 500 kpc du centre de l'émission X. La seule exception à cette règle, XLSSC 052 avec 2 galaxies membres seulement, est discutée dans l'article Pacaud et al. (2007). Les 70 sources vaidées à ce jour comportent de 3 à 50 galaxies membres confirmées par amas.

Compte tenu de la disponibilité de données photométriques de qualités sur le champ XMM-LSS (CFHTS), nous considérons actuellement la possibilité de valider les nouveaux canddidats à l'aide de redshifts photométriques. Une telle procédure pourrait être validée sur la base de l'échantillon spectroscopique déjà constitué.

### 4.4.2 Mesure des températures

La température du gaz est une donnée clef pour l'étude des amas de galaxies. Elle permet non-seulement d'estimer la masse des structures, mais fournit également une information riche sur la thermodynamique du gaz. Une bonne mesure de la température nécessite une statistique importante et n'est donc généralement pas accessible dans le cadre des grands relevés qui considèrent des sources sélectionnées près de la limite de détection de l'instrument.

#### Méthode

Dans le cas du XMM-LSS, l'émission détectable des amas est en général peu étendue. L'estimer du fond à la position de la source à partir d'un anneau l'en-

tourant reste donc possible. L'étude spectroscopique de nos détections s'en trouve facilitée.

Dans ce but, on commence par extraire un profil radial de la source dans la bande [0.5-2] keV et déterminer le plus petit rayon au delà duquel son émission semble négligeable (quelques rayons de coeur). On extrait ensuite le spectre du fond dans l'anneau le moins étendu possible au delà de ce rayon et permettant de recuellir au moins 3000 coups.

A partir du niveau moyen de fond, le rayon utilisé pour extraire le spectre de l'amas est ensuite calculé de manière à optimiser le rapport signal sur bruit dans la bande [0.5-2] keV (la figure 4.11 montre un exemple de région d'extraction).

Les spectres de source pour les trois détecteurs sont ensuite ajustés à un model d'emission thermique APEC sur la bande [0.3-10] keV par le logiciel `Xspec`. On utilise pour cela les matrices de réponse appropriées à la région d'extraction. L'ajustement tient compte des contraintes issues des spectres de fond, correctement renormalisés. La bande [7.5-8.5] keV est ignorée pour le détecteur PN, à cause des raies fortement inhomogènes du fond qui ne peuvent être correctement modélisées par une estimation locale. Enfin, l'abondance du plasma est fixée dans un rapport de 0.3 aux abondances solaires.

Une méthode couramment utilisé pour effectuer l'ajustement consiste à rééchantilloner le spectre afin de disposer d'au moins 20 ou 30 photons par point de donné, et de pouvoir utiliser la statistique du $\chi^2$. Pour des amas aussi faibles que ceux du XMM-LSS, cela constituerai une perte considérable de résolution, et il ne serait plus possible de contraindre la température. A nouveau, la C-statistique est donc préconisée.

### Validation

Lors de l'étude de l'échantillon préliminaire à bas redshift présenté dans l'article Willis et al. (2005), nous avons cherché à déterminer sous quelles conditions, i.e. pour quel type de sources, il nous était possible d'obtenir une température fiable.

Malgré la faible statistique, la méthode précédemment décrite conduisait lors des premières mesures à des résultats très précis : des températures constamement froides (de 0.5 à 2 keV) avec des barres d'erreur de l'ordre de 10%. Pour vérifier ces résultats, j'ai tenté de reproduire les valeurs obtenues à l'aide de simulations Monte-Carlo. Pour cela, on génére des réalisations de Poisson de spectres de fond issus des listes d'évènement de champs vides assemblées par Read & Ponman (2003), normalisées aux nombres de coups moyens observés dans nos données. On ajoute ensuite, pour différents flux, l'emission thermique d'une source à l'aide de la commande de simulation FAKEIT d'`Xspec`. Ceci permet d'assurer la cohérence entre le modèle simulé et celui qu'`Xspec` utilise pour l'ajustement. Les spectres (fond+source) et fond seul sont finalement ajustés par la même méthodes que pour les données réelles. Cela a permis d'identifier un fort biais vers les basses températures dans `Xspec`, dans le cas où l'on utilise la C-statistique pour de bas flux (c.f. Willis et al. 2005). Les faibles barres d'erreurs estimées sont également artificielles.

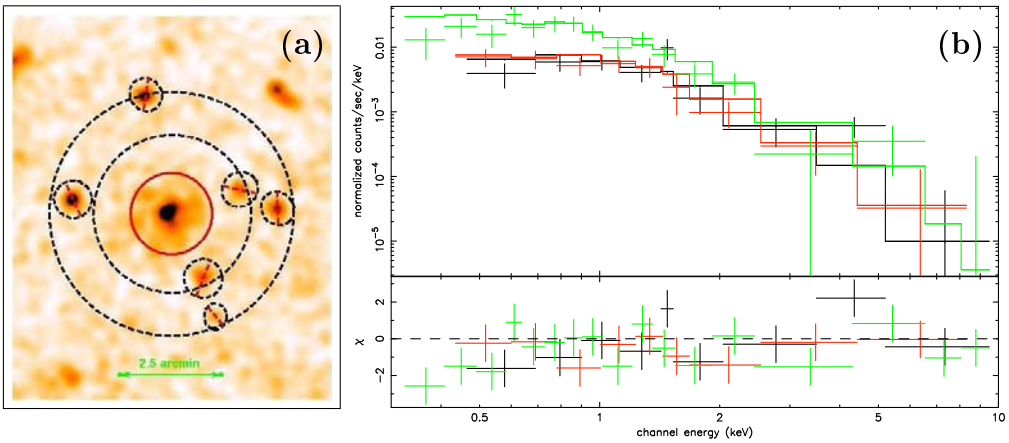


**FIG. 4.11**: *Mesure de temperature pour la source XLSSC 010 de l'échantillon C1. (a) : image co-additionné MOS1+MOS2+PN après convolution par une gaussienne d'écart-type 3 pixels. Le cercle rouge indique la région d'extraction du spectre de la source ; le spectre de fond est extrait dans la région délimitée par les traits pointillés noirs. Les ellipses en pointillé traversées par un trait rouge correspondent aux sources ponctuelles exclues pour l'estimation du fond. (b) : spectre de la source. Les points de données ainsi que leur barres d'erreurs sont indiqués par les croix (noir : MOS1, rouge : MOS2, vert : PN), le modèle ajusté est indiqué par les trois histogrammes. Pour plus de visibilité, les données ont été ré-échantillonnées pour que chaque point corresponde à un S/N supérieur à 3. Les résidus (en terme de $\chi^2$) sont montrés dans le panneau du bas. La température estimée pour cette source est* $T = 2.4^{+0.5}_{-0.4}$ *keV.*

Une inspection détaillée du modèle statistique utilisé par `Xspec` a par la suite révélé un comportement étrange et instable lorsque le spectre de fond contenait des valeurs nulles. Cela nous a amené à remarquer la disparition du biais et des effets associés lorsque les spectres sont ré-échantillonés de manière à supprimer les valeurs nulles. Les barres d'erreurs estimées sont alors proches de la dispersion statistique des températures mesurées (voir à ce propos l'annexe de Willis et al. (2005)). Pour nos données, nous avons utilisés un nombres minimum de 3 photons de fond dans chaque intervalle d'énergie. La figure 4.11 illustre le résultat de cette procédure pour la source XLSSC 010.

En se basant sur les 5 deg$^2$ du XMM-LSS déjà analysés, il semble possible d'obtenir une température par cette méthode pour tous les amas C1, avec une erreur de l'ordre de 20 à 30% en moyenne.

Cette possibilité ouvre des perspectives supplémentaires pour l'étude de l'échantillon, et constitue une première comparé aux grands relevés des précédentes générations. Ce résultat doit être imputé à la bonne résolution spectrale d'XMM, et au fait que la coupure exponentielle du brehmsstrahlung thermique pour des spectres de groupes et amas à 1-4 keV décalés vers le rouge tombe systématiquement dans la bande de sensibilité maximum du télescope.

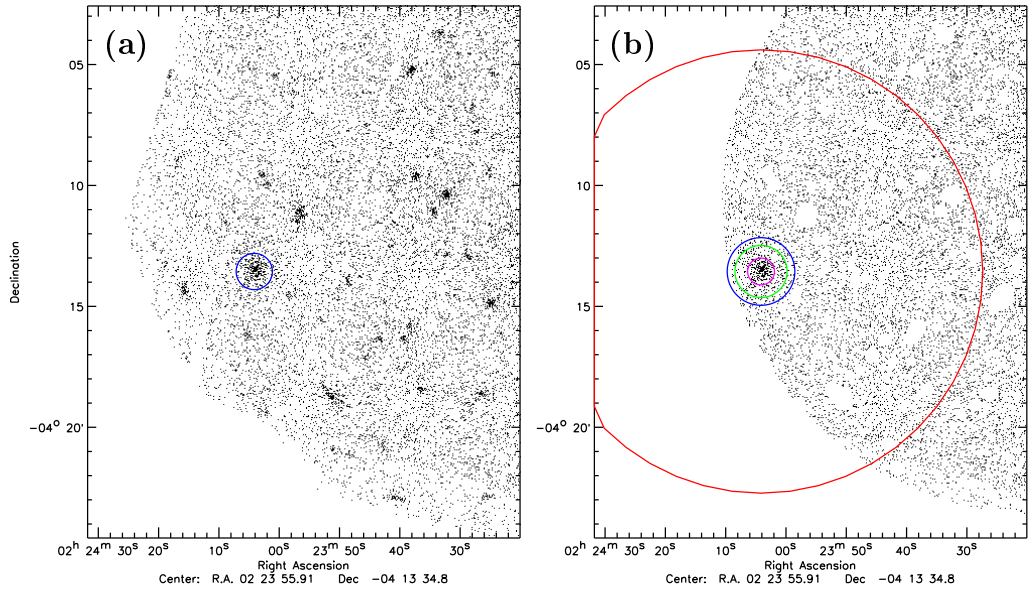


**Fig. 4.12**: *Extraction du profil de l'amas XLSSC 029 à z=1.05. La température moyenne estimée pour cette source est* $T$=$4.1^{+0.9}_{-0.7}$ *keV. (a) Les images XMM brutes co-additionnées ; la position de la source correspond au rond bleu indique. (b) Images nettoyées et co-additionnées. Les cercles indiquent différent rayons d'intérêt : en rouge le rayon de 500″ au-delà duquel le fond est mesuré, en bleu le rayon finalement utilisé pour l'ajustement du profil (dans lequel un S/N de 3 est obtenu), en magenta le rayon à l'intérieur duquel les spectres de la source ont été extraits afin d'obtenir sa température et en vert le rayon* $r_{500}$ *déduit de la température de la source (voir texte).*

### 4.4.3 Mesure des luminosités

Bien que la luminosité des amas soit un moins bon traceur de la masse totale que la température, il s'agit d'un paramètre critique pour l'analyse des relevés : la capacité de détection des sources en est une fonction quasi-exclusive. De plus, estimer simultanément le flux récolté et la température permet de remonter à la densité de gaz dans l'amas.

Afin d'utiliser une définition physique cohérente de la luminosité, il faut la mesurer dans un rayon physique tenant compte du caractère auto-similaire des amas (i.e. défini par rapport à un contraste de densité fixe). Pour cela, il faut déterminer précisément le profil de brillance de surface pour chaque source.

#### Méthode

Tout comme pour la détection des sources, le profil d'émission des amas est ajusté dans la bande [0.5-2] keV de manière à optimiser le rapport S/N.

Dans un premier temps, nous avons pratiqué une analyse à deux dimensions du profil d'émission, permettant de prendre en compte une éventuelle ellipticité des amas analysés. Cependant les résultats obtenus alors étaient très instables à

cause du faible signal disponible. La méthode finalement retenue consiste donc à ajuster un modèle $\beta$ sphérique au profil radial des sources.

Rasia et al. (2006) ont montré sur des amas simulés que cette supposition entraine des biais importants dans l'estimation de la masse des amas, et que les résidus de l'ajustement en terme de $\sigma$ sont statistiquement significatifs. Cependant, la déviation par rapport au profil d'émission réel des amas reste inférieure à 10% sur l'intervalle [0.1-1] $r_{vir}$ à condition que l'amas étudié ne subisse pas une fusion majeure, et soit donc dominé par une unique concentration de gaz. Cette précision semble suffisante si l'on se cantonne à estimer la luminosité des sources.

L'extraction du profil procède par 4 étapes :

- On opère d'abord un nettoyage des images X brutes. Pour cela, on supprime les régions distantes de plus de 13′de l'axe optique moyen. Puis, on exclut les pixels associés à d'autres sources que l'amas considéré, tel qu'indiqués par la carte de segmentation construite par `SExtractor`. Les masques correspondants sont finalement générés. La figure 4.12 montre un exemple d'image et le résultat du nettoyage.
- Les niveaux du fond de particule et du fond vignetté dans l'axe sont ajustés (en utilisant la C-statistique) sur les région non-masquées situées à plus de 500″de l'amas.
- Dans ce rayon de 500″, on extrait alors un premier profil en coups dans des anneaux concentriques de 3″autour de la position `SExtractor` de l'amas. Celui-ci est ensuite ré-échantilloné afin d'obtenir un S/N minimum de 3 dans chaque anneau. On ignore finalement les parties externes qui ne permettent pas de satisfaire ce critère.
- En vue d'obtenir le profil de brillance de surface, on soustrait le fond de particule mesuré précédemment. Puis les données sont corrigées du temps d'exposition et de la surface de chaque anneau. Enfin, on soustrait la contribution du fond vignetté

Le problème crucial pour ajuster un profil à une dimension est de prendre en compte rigoureusement la PSF. En effet il n'existe pas d'opération mathémathique générale qui permette d'obtenir directement le profile radial à une dimension généré par la convolution de deux distributions à deux dimensions (même si celles-ci vérifient la symétrie sphérique). Pour ce faire, nous construisons une matrice de redistribution de la PSF, similaire à celles qui sont utilisées pour la spectroscopie. On commence par considérer un système d'anneau très fin autour de la source. La fraction du flux redistribuée par la PSF depuis chacun de ces anneaux vers les couronnes dans lesquels le profil a été extrait est obtenue par le calcul d'une intégrale de recouvrement. Pour le modèle de PSF de Ghizzardi 2001 (MOS) et Ghizzardi 2002 (PN), on peut évaluer l'intégrale analytiquement. Lors de l'ajustement, on calcul le flux du modèle $\beta$ dans chacun des anneaux de départ et une multiplication matriciel fournit le profil observé. Compte tenu du nombre important de photons dans chaque anneau, la statistique du $\chi^2$ est utilisée, avec des erreurs de Poisson sur les points de données. La figure 4.13 montre le profil mesuré pour l'amas XLSSC 029.

Une fois le profile de brillance de surface déterminé, on le convertit en des

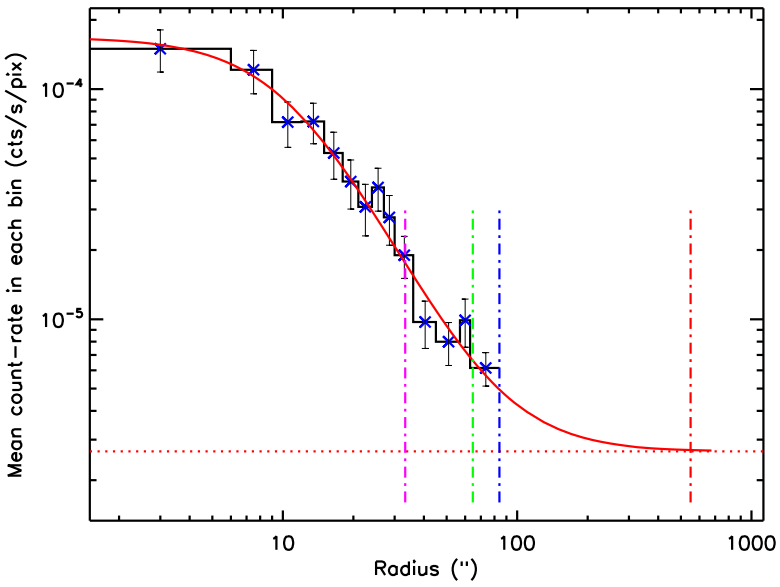


**FIG. 4.13**: *Profil de l'amas XLSSC 029. Les points de données sont indiqués par les croix bleues avec les barres d'erreurs associées, l'histogramme délimite les anneaux dans lesquels chaque point a été mesuré. La ligne pleine rouge montre le profil ajusté, et les lignes verticales indiquent différents rayons caractéristique avec le même code de couleurs que dans la figure 4.12. Les paramêtres ajustés pour cette source sont* $\beta = 0.52$ *et* $R_c = 6.4$".

profils de flux et de luminosité bolométrique projetés. Les facteurs de conversion correspondants sont dérivés du modèle spectral précédemment ajusté. Pour les amas du XMM-LSS, on intègre la luminosité dans une ouverture angulaire équivalente à $r_{500}$. Cependant, notre hypothèse d'isothermalité ainsi que l'utilisation d'un modèle $\beta$ ne permettent pas d'obtenir un profil de masse crédible, comme l'ont démontré Rasia et al. (2006) et Vikhlinin et al. (2006). C'est pourquoi on estime la valeur de $r_{500}$ à partir de la relation $M_{500} - T$ locale de Finoguenov et al. (2001), que nous avons réévaluée dans le cadre du modèle de concordance cosmologique, associée à une évolution auto-similaire :

$$r_{500} = 0.391\, T_X^{0.63}\; h_{70}(z)^{-1} Mpc \tag{4.8}$$

### Validation

L'ensemble de la procédure fut encore une fois testée sur des images simulées. Ainsi que discuté dans l'annexe de Pacaud et al. (2007), les erreurs individuelles obtenues sur $R_c$ et $\beta$ sont grandes, et les valeurs mesurées très bruitées. Par conséquent on ne peut se fier à ces paramètres pour obtenir une information sur le gaz d'un amas donné. Cependant, les profils extraits reproduisent fidèlement l'emission réelle des sources simulées, et on peut déterminer des flux précis et peu biaisés, ce

qui était le but recherché. L'explication probable de cet apparent paradoxe tient au fait que le faible temps d'exposition pour nos amas ne nous permet pas d'extraire des profils à suffisament grand rayon ; hors la valeur de $\beta$ dépend essentiellement de la pente dans les régions externes et donc ne peut être correctement déterminée. La valeur du rayon de coeur est également impactée par cette indétermination du fait de la dégénérescence $\beta$-$R_c$ pour les partie centrales du profil. Cela implique que nos mesures doivent être considérées avec précautions lorsque l'on tente d'extrapoler les mesures au dela du rayon sur lequel l'ajustement a été effectué. Pour la grande majorité de nos sources, $r_{500}$ se trouve en deçà de ce rayon et le problème ne se pose donc pas.

## 4.5 Discussion

Les méthodes présentées dans cette section permettent d'obtenir, en se basant uniquement sur les listes d'évènements fournies par le télescope, un catalogue complet, contenant l'ensemble des informations disponibles compte tenu de nos données. Une attention particulière a été portée sur la validation de chacune des étapes par l'intermédiaire de simulations massives et systématiques. Cela nous a permis entre autre d'identifier dans chaque cas les limitations intrinsèques aux techniques utilisées et d'en tenir compte par la suite.

En ce qui concerne la détection des amas, les ondelettes semblent à ce jour la méthode la plus complète pour traiter proprement la présence à la fois d'AGN et d'amas, et générer des listes de sources fiables. De plus le logiciel `Xamin` permet une classification rigoureuse des sources étendues tout en rivalisant avec le logiciel officiel `EMLDETECT` pour la détection et la photométrie des sources ponctuelles.

L'analyse individuelle des amas détectés a de plus révélé qu'il était possible d'obtenir une estimation de température pour une fraction significative de nos sources, et notamment pour l'ensemble de l'échantillon C1. Associé à une mesure de luminosité crédible, cela ouvre la possibilité d'étudier la relation $L_X$-$T$ pour un échantillon couvrant un intervalle de masse suffisant et dont la fonction de sélection peut être précisément mesurée.

Il existe d'autres paramètres intéressants et potentiellement mesurables à partir des données X seules. En particulier, le profil d'entropie et le profil de masse (de gaz et totale). Ces quantités restent toutefois largement inaccessibles dans avec nos données car elles nécessiteraient la mesure du profil de température et des contraintes précises sur la brillance de surface dans les régions externes des amas.

# Chapitre 5
# Résultats scientifiques

## Sommaire



Ce chapitre se concentre sur l'exploitation des données XMM observées dans le cadre du relevé XMM-LSS entre 2000 et 2003 (à l'exclusion du Subaru Deep Survey). Celles-ci comprennent donc les 18 pointés utilisables du XMDS (20ks), ainsi que 27 observations d'une durée nominale de 10ks. La Table 1 de l'article Pierre et al. (2007, inséré section 7.4) fournit des informations détaillées sur chacun de ces pointés (identifiant, position, temps d'exposition utilisables pour les 3 détecteurs EPIC).

La majorité des résultats que nous allons mentionner ont fait l'objet de publications dans des revues à comité de lecture, inclues dans la partie II de ce mémoire. Il n'est donc pas question ici de répéter l'intégralité des travaux développées dans chaque article, mais plutôt d'en extraire les résultats essentiels et d'y ajouter quelques compléments afin de faire apparaître la cohérence globale et la démarche adoptée pour l'ensemble des travaux présentés. Nous invitons le lecteur à consulter la partie II en parallèle pour des informations plus détaillées.

Le chapitre se conclura sur un aperçu des catalogues extraits des observations plus récentes.

## 5.1 L'échantillon d'AGN

### 5.1.1 Descriptions

Bien que l'étude des amas de galaxies soit présentée comme le but principal du XMM-LSS, environ 95% des sources visibles dans les données XMM hors du plan Galactique sont des AGNs. Le relevé fournit ainsi le plus large échantillon jamais construit à cette profondeur sur une surface contigüe.

L'analyse des 45 premiers pointés, associée au critère de sélection général `DET_ML`>15 présenté dans le chapitre précédent, permet d'identifier 2980 sources en bande [0.5-2] keV, et 1255 en bande [2-10] keV. Seul 2% des sources dans la bande molle sont classifiées comme potentiellement étendues au sens des critères C1/C2. L'identification croisée des sources entre les deux bandes et sur différents pointés (avec une distance de corrélation de $6''$) permet de combiner l'ensemble de ces détections en 3385 sources distinctes.

Pierre et al. (2007, inséré section 7.4) présente plus en détail les propriétés de ce catalogue qui vient d'être rendu public via une base de données située à Milan[1]. Un liste des détections optique du CFHTLS associées à chaque source y est également disponible.

### 5.1.2 Effets de sélection

#### Complétude et contamination

Sur la base du niveau de contamination observé dans les simulations [$\sim 0.5$ source par pointé de 10ks d'après Pacaud et al. (2006, inséré section 7.2)], le catalogue de sources ponctuelles susmentionné doit contenir au plus 30 fausses détections, soit moins de 1%. Il se révèle donc relativement propre.

Les simulations d'AGNs générées pour tester et évaluer le pipeline permettent également de calculer la fonction de sélection associée au seuillage en vraisemblance. L'article Pacaud et al. (2006, inséré section 7.2), fournit par exemple la probabilité de détection moyenne en fonction du flux des AGNs dans les $10'$ centrales d'un pointé de 10ks. Dans ces conditions, le relevé atteindrait une complétude supérieure à 80% à partir d'un flux de $4 \times 10^{-15}$erg.s$^{-1}$.cm$^{-2}$.

[1] `URL : http ://cosmosdb.iasf-milano.inaf.it/XMM-LSS/`

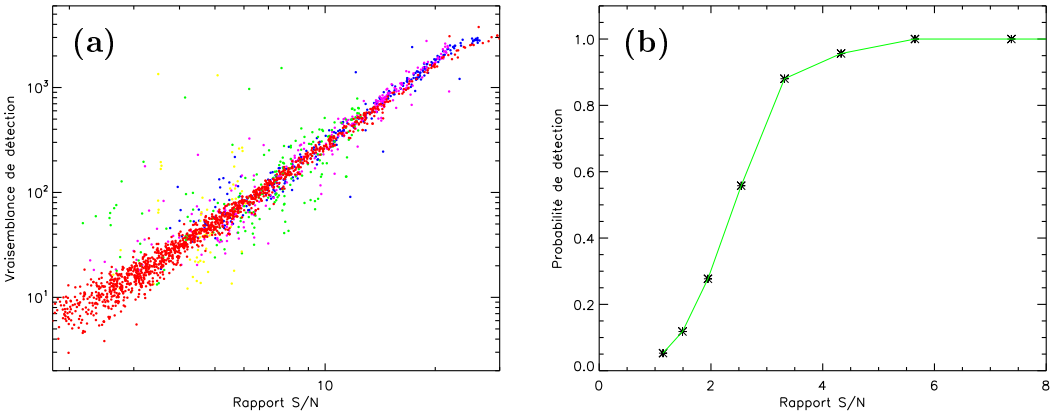


**FIG. 5.1**: *Modélisation de la fonction de sélection en terme de S/N.* **(a) :** *Corrélation entre la vraisemblance de détection mesurée par* `Xamin` *et le rapport S/N des sources dans une boite de* $65'' \times 65''$*. Le code de couleurs est le même que pour la Fig. 4.10.* **(b) :** *fonction de sélection des sources ponctuelles en fonction du signal sur bruit dans une ouverture circulaire de* $10''$ *de rayon.*

Un recours systématique à ce type de simulations semblerait la méthode la plus rigoureuse et robuste pour quantifier les effets de sélection dans les conditions d'observation propres à chaque pointé (temps d'exposition, modèle de fond). Cela peut cependant se révéler difficile à mettre en oeuvre dans la pratique, car il faudrait renouveler l'ensemble des simulations chaque fois que l'on réévalue des paramètres susceptibles d'influencer la détection (niveau de fond, distribution spatiale et en flux des AGNs, modèle de vignettage des télescopes).

Une méthode plus efficace pour modéliser les effets de sélection (au sens du temps de calcul et de la flexibilité) consiste à paramétrer chaque effet en terme de rapport S/N des détections : comme le montre la Fig. 5.1a, ce dernier affiche une corrélation prononcé avec la vraisemblance de détection qui nous sert de critère discriminant. En pratique, le S/N estimé dans une ouverture circulaire de $10''$ de rayon semble une paramétrisation acceptable de l'ensemble des effets associés aux sources ponctuelles. Son lien avec la probabilité de détection pour les AGNs issus des simulations est par exemple illustré par la Fig. 5.1b.

Pierre et al. (2007, inséré section 7.4) applique cette modélisation pour estimer la surface effective du relevé pour les sources ponctuelles en fonction de leur flux. Le catalogue atteint 80% de complétude vers des flux respectifs de $3.5 \times 10^{-15}/2 \times 10^{-14}$erg.s$^{-1}$.cm$^{-2}$ dans les bandes molle/dure.

### Photométrie et astrométrie

L'analyse des simulations a révélé un fort biais photométrique positif à bas flux. Celui-ci peut s'interpréter comme un biais d'Eddington [Eddington (1913)] : la photométrie des sources les plus faibles devient très bruitée et la limite de

vraisemblance imposée tend à sélectionner préférentiellement les détections dont le flux a été surévalué. Cette fois encore, l'amplitude moyenne (en %) du biais est très bien corrélée au rapport S/N des sources et peut s'écrire :

$$b = 20 \times \exp\left(-S/N\right)^{0.2}. \tag{5.1}$$

La dispersion autour de cette tendance générale (correspondant à une loi log-normale) décroît avec le S/N :

$$\sigma_b = 1.25 \times [S/N]^{-1.3} \tag{5.2}$$

En pratique, cette augmentation du biais et de la dispersion empêche d'établir facilement une loi bijective entre flux observé et flux réel et donc de déterminer une correction photométrique pour les sources dont le S/N devient plus petit que 3. Un seuil comparable ($S/N > 4$) fut estimé de manière théorique par Hogg & Turner (1998) en considérant le biais d'Eddington pour un bruit de mesure gaussien.

Cela rend également impossible à bas flux la correction du nombre de source observé par la fonction de sélection. Comme on peut le vérifier sur la Fig. 5.1a, un rapport S/N de 3 correspond encore à une probabilité de détecter les sources de l'ordre de 75%. Compte tenu du nombre important d'AGNs à bas flux dans le ciel (c.f. Eq.4.1), cela signifie qu'un nombre conséquent de sources ne peuvent être utilisées pour des études statistiques dès que les effets de sélection doivent être pris en compte : environ 20% des détections dans la bande molle sont concernées.

Dans les conditions d'observation nominales du XMM-LSS (expositions de 10ks et niveau de fond moyen observé dans les premiers pointés), cette limite fondamentale du pipeline correspond à des flux en bande [0.5-2] keV de $2.1 \times 10^{-15}/3.9 \times 10^{-15}$ erg.s$^{-1}$.cm$^{-2}$ respectivement pour une source dans l'axe / à $10'$ de l'axe optique.

Une relation entre l'incertitude sur l'astrométrie des AGNs du catalogue et le rapport S/N peut également être formulée. Celle-ci a permis d'estimer les erreurs présentées dans Pierre et al. (2007, inséré section 7.4).

### Distribution en Flux

La distribution en flux des sources X extragalactiques est désormais bien contrainte jusqu'à des niveaux avoisinant respectivement $10^{-17}$ et $10^{-16}$erg.s$^{-1}$.cm$^{-2}$ dans les bandes molle et dure (voir par exemple Moretti et al. 2003). Il s'agit d'un sujet d'investigation encore actif pour les relevés ultra-profonds qui tentent de résoudre la majorité du fond diffus X en sources discrètes, tel que le prédisent les modèles actuels (e.g. Gilli et al. 2001). Dans le cas du XMM-LSS, dont les observations sont moins profondes, la fraction du fond diffus résolu en sources ponctuelles est de l'ordre de 30% [voir Gandhi et al. (2006, inséré section 7.6)]. Reproduire la distribution en flux des sources constitue alors un test de validation plutôt qu'un nouveau résultat.

Cette mise à l'épreuve fut entreprise dans Pierre et al. (2007, inséré section 7.4). Après une correction moyenne du biais photométrique, la densité de source

en fonction du flux est calculée puis corrigée de la fonction de sélection. Les courbes Log(N)-Log(S) ainsi obtenues se trouvent, pour les deux bandes, dans les marges d'erreurs des mesures déduites des relevés profond XMM et Chandra (ainsi que des précédentes estimations utilisant ROSAT).

### 5.1.3 Fonction de corrélation

En se basant sur une modélisation en S/N des effets de sélection, l'autocorrélation angulaire (ACF) des AGNs jusqu'à des échelles atteignant $100''$ a également été étudiée pour l'article Gandhi et al. (2006, inséré section 7.6). Dans ce but, des catalogues de distribution spatiale uniforme sont générés aléatoirement à partir de la Log(N)-Log(S) observée, et assujettis au critère S/N>3. La distribution spatiale de ces échantillons simulés sert alors de référence pour quantifier l'excès de corrélation présent dans nos catalogues (soumis au préalable au même critère SN>3).

On mesure une faible corrélation en bande [0.5-2] keV, compatible avec les travaux antérieurs utilisant ROSAT (Vikhlinin & Forman, 1995) et XMM (Basilakos et al., 2005). Contrairement aux résultats de Basilakos et al. (2004) et Yang et al. (2003), aucune corrélation significative n'a par contre été observée pour la bande dure. La différence avec ces derniers provient probablement de l'échantillon utilisé : leur étude se base sur des données Chandra plus profondes que le XMM-LSS et disposant d'une meilleure résolution spatiale. Les travaux de Plionis et al. (2008) exposent en effet plusieurs indications en faveur d'une variation de la fonction de corrélation observée avec le flux limite de l'échantillon considéré. Le désaccord avec Basilakos et al. (2004) semble, en revanche, plus complexe. Il pourrait provenir de la variance cosmique (leur analyse utilise une surface nettement plus petite). Il subsiste aussi de subtiles différences dans l'estimateur de l'ACF utilisé (voir l'article pour plus de détails).

Une corrélation à un niveau de $2/3\sigma$ est tout de même observée dans la bande [2-10] keV sur des échelles supérieures à $50''$ lorsque l'on sélectionne seulement les AGNs qui présente les spectres les plus durs. Ceux-ci sont définis dans notre cas par un rapport de dureté $\frac{D-M}{D+M}$ >-0.2, où D est le flux en bande dure et M en bande molle. Ce type de spectre ne peut s'expliquer que pour des AGNs dont l'émission est fortement obscurcie par le materiel environnant (nécessitant des densités de colonne $N_H > 10^{-22}\,\mathrm{cm}^{-2}$). Si ce résultat préliminaire venait à être confirmé, cela impliquerait que les AGNs obscurcis se trouvent préférentiellement associés aux zones de densité élevé. On pourrait alors en conclure que le milieu dense présent dans ces régions contribue non seulement à la formation des AGNs, mais aussi à leur enfouissement. Il convient cependant de remarquer que le sous-échantillon d'AGNs obscurcis correspond en moyenne à des flux moins élevés. A la lumière des travaux de Plionis et al. (2008), la différence de corrélation angulaire observée pourrait donc également trouver une explication indépendante de l'obscurcissement dans le cas où elle dépendrait du flux considéré.

### 5.1.4 Autres résultats

Le catalogue a également servi de base à l'étude des propriétés physiques des noyaux actifs de galaxies. Un sous-échantillon représentatif de 99 AGNs, extrait du catalogue en bande [2-10] keV et comportant des données en spectroscopie optique, a été étudié dans Garcet et al. (2007). L'article classifie les sources en fonction de leur obscuration optique et de leur absorption X qui devraient avoir la même origine dans la version la plus simple du schéma unifié (Antonucci, 1993). Cependant la corrélation modérée observée pour l'échantillon s'avère difficile à réconcilier avec ce modèle (1/3 des sources ont des classifications différentes entre X et optique). La majorité des AGNs discordants étant obscurcis uniquement dans les bandes optique, l'article démontre que l'absence de corrélation est probablement dû à des effets de dillution du flux optique de l'objet central dans l'émission de la galaxie hôte.

## 5.2 Les amas du XMM-LSS

Les résultats présentés par la suite dans ce mémoire se rapporte tous aux amas de galaxies extraits du catalogue de Pierre et al. (2007, inséré section 7.4).

### 5.2.1 L'échantillon C1

L'application des critères C1 aux 45 pointés valides permet de sélectionnner 33 sources, dont 28 amas (tous confirmés par spectroscopie optique) et 5 galaxies proches, soit une densité surfacique d'amas supérieure à 5 deg$^{-2}$. Cet échantillon a été présenté et étudié dans Pacaud et al. (2007, inséré section 7.5). Les 28 amas ont des redshifts distribués entre z=0.05 et 1.05 avec un pic prononcé vers z$\sim$0.3. Les températures mesurées pour ces sources varient entre 0.5 et 5 keV et sont fortement corrélées à leur redshift : les 3 systèmes identifiés à $z < 0.1$ ont des températures de l'ordre de 0.5 keV, tandis que 6 des 7 sources avec $z > 0.6$ ont des températures supérieures à 2.5 keV. Cette propriété peut s'interpréter par la combinaison d'un biais de Malmquist avec la géométrie du relevé : les amas froids, donc peu lumineux ne peuvent être détectés à grand redshift, tandis que le volume couvert à bas redshift n'est pas suffisant pour contenir des amas massifs.

La distribution en brillance de surface de ces sources semble indiquer des morphologies très diverses. De nombreux groupes sont apparemment relaxés, présentant même des signes d'un coeur froid (leur profile central est très piqué). D'autres se révèlent perturbés ou en cours de formation.

On peut également estimer des masses approximatives pour ces sources, en les décrivant par un modèle $\beta$ isotherme. Ces masses (dans $r_{500}$) s'échelonnent entre $8\times10^{12}$ et $3\times10^{14}\,M_\odot$ et correspondent bien à des groupes et amas peu massifs. De telles mesures sont cependant affectées par plusieurs biais. Vikhlinin et al. (2006) montre par exemple qu'utiliser un modèle $\beta$ tend à sous-estimer les masses d'environ 30% pour une température de 5 keV en raison d'une augmentation brusque de la pente du profile de densité dans les régions externes des groupes, un chiffre qui augmente à mesure que la température diminue. Rasia et al. (2006) ont confirmé

l'existence de ce biais sur des amas issus de simulations numériques, et estiment une déviation de 40%.

Par ailleurs, Balogh et al. (1999) montre que les groupes avec $T \leq 0.8$ keV ne vérifient vraisemblablement pas l'équilibre hydrostatique. L'émission des 3 groupes en deçà de $z = 0.1$ et $T = 1$ keV paraît de toute manière ambigüe car une grande partie du signal provient des régions entourant la galaxie centrale. Il devient alors irréalisable de dissocier la fraction du flux qui provient du MIA de l'émission propre à la galaxie. Ceci implique que la fonction de sélection issue des simulations est inapte à décrire proprement ces systèmes et ils sont par conséquent ignorés dans toutes les études statistiques. L'émission de l'amas XLSSC 018 semble, par ailleurs, contaminée par un AGN et celui-ci sera donc également écarté.

Pour finir, un amas supplémentaire (XLSSC 050) est détecté sur un pointé contaminé par des particules. Bien que cette source ne soit pas non plus inclue dans les analyses statistiques présentées par la suite, elle vient grossir l'échantillon. Il s'agit en effet du cas interessant de deux sous-groupes en fusion à z=0.14. Vus de face, ils présentent un décalage entre l'émission X et la distribution de galaxie rappelant, dans un cas moins extrême, le célèbre amas du boulet, 1E 0657-56 (Markevitch et al., 2002; Clowe et al., 2006).

### 5.2.2 Autres échantillons

#### L'échantillon C2

Celui-ci comprend 38 détections (auxquelles s'ajoutent 2 galaxies prôches, un amas déjà classé comme C1 à partir d'un autre pointé et 3 sous-structures associées à des amas C1). Si le taux de contamination d'environ 50% attendu d'après les simulations s'avère exact, près de 20 amas devraient être identifiés parmi ces sources. Le suivi spectroscopique a permis d'en confirmer 9 à ce jour, tandis que la simple superposition des contours X sur les images optique du CFHTLS permet d'identifier au moins une dizaine de fausses détections. L'analyse des spectres n'est pas encore terminée, mais il est peu probable que plus de 15 sources soient finalement assimilées à des amas. Il serait risqué dans l'état actuel de nos connaissances de se prononcer sur l'origine de cette apparente surcontamination de l'échantillon. Il pourrait s'agir simplement de variance cosmique.

Les redshifts des sources déjà identifiées s'étalent entre $0.3 < z < 1.22$. Parmi celles-ci se trouve XLSSC 046, l'amas le plus distant du XMM-LSS à ce jour, et l'un des 10 amas X les plus distants publiés jusqu'à présent. L'analyse effectué pour Bremer et al. (2006) tend de plus à montrer qu'il s'agirait de l'amas au plus bas flux détecté à $z > 1$.

#### L'échantillon C3

Plus d'une dizaine d'amas associés avec une source X n'ayant pas satisfait aux critères C1/C2 ont été spectroscopiquement validés [dont 4 publiés dans Willis et al. (2005, inséré section 7.1) et 4 dans Pierre et al. (2006, inséré section 7.3)].

Les données obtenues au titre du suivi spectroscopique comprennent plusieurs dizaines d'autres candidats qui attendent confirmation.

Une compilation des amas C2 et C3 sur l'ensemble des premiers $5\mathrm{deg}^2$ est en cours (Adami et al., en préparation).

### 5.2.3 Modéliser la fonction de sélection

Les programmes de simulation ont permis de déterminer la fonction de sélection C1/C2 des amas, dans l'espace (Flux,Extension) pour des pointés de 10 ks, en supposant un niveau de fond moyen représentatif. Les courbes correspondantes sont disponibles dans Pacaud et al. (2006, inséré section 7.2). Les probabilité mesurées dépendent fortement de l'extension des sources, et ne peuvent donc être modélisées par une limite en flux ou une surface effective dépendant du flux, contrairement aux modèles utilisés dans la plupart des relevés antérieurs [e.g. RDCS (Rosati et al., 1998), BCS (Ebeling et al. 1998 et Ebeling et al. 2000), NORAS (Böhringer et al., 2000), REFLEX (Böhringer et al., 2004)]. A notre connaissance, les seules études dans lesquels la fonction de sélection tient explicitement compte du profil des sources sont Vikhlinin et al. (1998) et Adami et al. (2000).

Du fait de cette dimension supplémentaire, simuler la fonction de sélection sur l'ensemble des pointés du relevé devient encore plus épineux pour les sources étendues que pour les AGNs. En outre, ces simulations seraient attachées à un profil de source particulier et ne seraient pas nécessairement valables si l'on en changeait. Le temps de calcul correspondant serait, par ailleurs, d'autant plus élevé que la densité de source étendues dans les observations simulées est plus faible. Tout ces arguments conduisent de nouveau à privilégier un modèle plus général, basé sur le rapport S/N.

Cependant, dans le cas des sources étendues, cette grandeur n'est plus définie de manière univoque en raison de la variété de profils possibles, mais son interprétation dépend du rayon dans lequel on la calcule. Dans ce but, nous avons recherché, de manière empirique, un rayon dans lequel il serait possible de définir une relation stable entre S/N et probabilité de détection pour l'échantillon C1, indépendante de l'extension de la source. Comme le montre la Fig. 5.2 une telle propriété est obtenue pour une ouverture circulaire de $45''$ de rayon. Sur un plan purement conceptuel, cette mesure dans un rayon de 3 fois la FWHM de la PSF semble d'ailleurs pertinente afin de refléter la capacité de l'algorithme à décider du caractère étendu d'une source. La relation moyenne obtenue permet d'évaluer immédiatement la probabilité de classifier un amas dans l'échantillon C1 pour un temps d'exposition et un niveau de fond donné, en fonction de son flux et de son profil.

La surface effective exposée dans le papier Pacaud et al. (2007, inséré section 7.5) et utilisée dans l'étude de l'échantillon C1 provient de ce modèle.

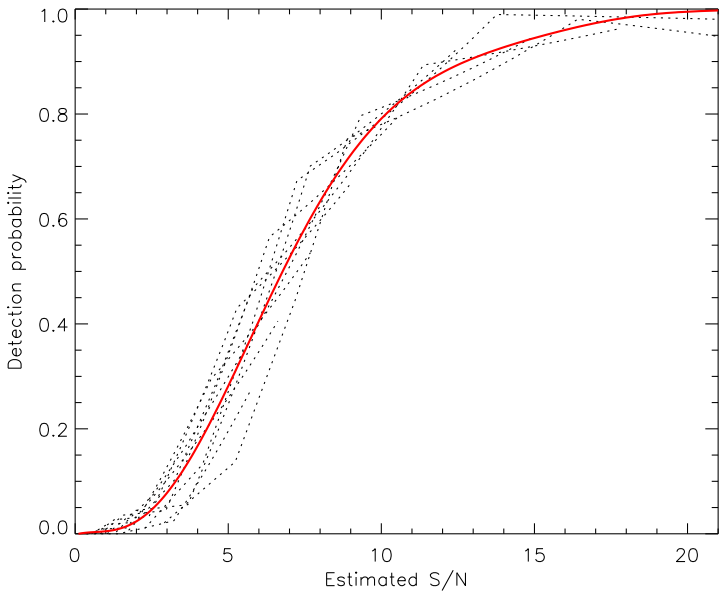


**FIG. 5.2**: *Fonction de sélection C1. Les courbes en pointillés proviennent de simulations et montrent la probabilité de détection en fonction du rapport S/N généré dans 45″ de rayon pour des amas ayant différents rayons de coeur (allant de 10 à 100″) et $\beta = 2/3$. La courbe rouge est un ajustement polynomial à l'ensemble de ces données.*

## 5.3 Evolution de la relation $L_X$-$T$ avec les amas C1

L'étude détaillée de la corrélation $L_X - T$ de l'échantillon C1 est présentée dans Pacaud et al. (2007, inséré section 7.5). Nous en retraçons ici les grandes lignes en insistant sur certains points importants.

### 5.3.1 Vers une modélisation plus rigoureuse

#### Etat des lieux

La relation $L_X$-$T$ locale des amas semble bien contrainte (e.g. Markevitch (1998), Arnaud & Evrard (1999)) depuis maintenant une dizaine d'année. Il apparaît désormais clairement que cette loi d'échelle évolue négativement au cours du temps, i.e. associait de par le passé une plus grande luminosité à une température donnée. Cependant, nombres de résultats incompatibles ont été rapportés quant à l'amplitude de cette évolution [voir entre autre Ettori et al. (2004), Vikhlinin et al. (2002) et Novicki et al. (2002)]. Une telle divergence conduit à une incertitude sur la luminosité des amas d'un facteur 2/3/4 respectivement à des redshifts de 0.6/1.1/1.5. De plus, si les évolutions les plus légères parmi ces mesures (e.g. Ettori et al. (2004) ou Maughan et al. (2006)) sont au premier ordre compatibles avec les prévisions du modèle auto-similaires, il n'existe pas à l'heure actuelle de modèle capable de prédire la forte évolution observée dans la plupart des cas.

Plusieurs explications plausibles peuvent être avancées pour expliquer ces incohérences :

- Tout d'abord, la distribution auto-similaire de la masse [Pointecouteau et al. (2005) et Vikhlinin et al. (2006)] et du gaz [Vikhlinin et al. (2005), Pratt et al. (2006) et Pratt et al. (2007)] dans les amas implique que les propriétés physiques de ces systèmes doivent être mesurées dans un rayon représentatif de la densité critique de l'univers (typiquement $r_{500}$). Or toutes ces études n'utilisent pas la même convention pour définir le rayon d'intégration. Pire encore, le rayon adopté pour établir les références locales de Markevitch (1998) et Arnaud & Evrard (1999) n'a jamais été explicité. L'impact de cette imprécision reste cependant limité dans la mesure où la luminosité totale varie lentement avec le rayon dans les parties externes.
- Les travaux actuels supposent une evolution de la forme $L(z,T)/L(z=0,T) \propto (1+z)^{\alpha}$. Cela pourrait engendrer des résultats divergents pour différentes couvertures en redshift dans le cas où cette fonctionnelle serait une mauvaise représentation de l'évolution réelle.
- Finalement, les échantillons étudiés pour faire ces mesures sont sélectionnés de manière relativement arbitraire. L'impact de possible effets de sélection n'a jamais été envisagée (et pour cause : il ne serait tout simplement pas possible de les modéliser).

Dans ce contexte, la possibilité de mesurer des températures pour l'échantillon C1 dont les effets de sélection sont bien connus constitue une opportunité unique de jeter quelque lumière sur ce problème. Pour cela, nous supposons que seul la normalisation de la relation évolue dans le temps, la dégénérescence $z$-$T$ de l'échantillon ne nous permettant pas de contraindre la pente à différentes époques (il s'agit toutefois d'une approche fréquente dans ce type d'études).

### Impact de l'incomplétude

La méthode usuelle pour caractériser les relations d'échelle consiste en une regression de type $\chi^2$ dans l'espace $\log - \log$. Bien que très employé en astronomie, ce procédé ne permet pas de prendre en compte des effets subtils tel que l'incomplétude du catalogue utilisé.

Pour palier à ce défaut, une possibilité est de considérer explicitement la distribution des sources dans l'espace $L_X$-$T$ pour un modèle donné, et de la normaliser pour en déduire une vraisemblance associée à chaque amas. Il suffit alors d'inclure le processus de sélection des sources dans la distribution $L_X$-$T$ pour corriger d'éventuels biais.

Une méthode de ce type est développée dans Pacaud et al. (2007, inséré section 7.5) en se basant sur une distribution Log-normale de la luminosité à une température fixe (d'après Stanek et al. 2006), et en supposant un rayon de coeur de 180 kpc pour tous les amas afin d'estimer la fonction de sélection. Un biais de Malmquist apparait alors à basse température dans une région qui, en raison de la pente importante de la fonction de masse, englobe une large fraction des amas C1.

Du fait de la généralité du modèle exposé, il est probables que nombres d'échantillons passés, construits près de la limite de détection des instruments,souffrent du même type de biais. Cela pourrait donc s'appliquer aux échantillons d'amas locaux massifs, pour lesquels la variation en taille des sources (et donc la dillution du flux) est succeptible d'influencer la probabilité de détection.

### 5.3.2 Résultats préliminaires

L'essentiel des conclusions relatives à la relation $L_X$-$T$ peuvent se résumer à deux observations, illustrées par les figures 7a et 7b de Pacaud et al. (2007, inséré section 7.5).

Tout d'abord, les échantillons utilisés dans les études précédentes se rapportent à différentes coupes en redshift. Contrairement à l'incohérence manifeste des résultats antérieurs, les contraintes apportées pour chacune de ces bandes sur le facteur d'évolution $F = L(z,T)/L(z=0,T)$ se révèlent en fait compatibles entre elles. Par contre, la tendance générale avec le redshift ne peut être retranscrite par la simple loi de puissance $F = (1+z)^\alpha$ généralement utilisée, ce qui engendre l'apparente contradiction. L'image globale qui en émerge montre plutôt une forte évolution négative depuis z$\sim$0.3 précédée par une évolution quasi nulle ou légèrement positive. Un résultat similaire a été obtenu simultanément et indépendamment par Branchesi et al. (2007). Si l'on ignore tout effet de sélection, les amas de l'échantillon C1 semblent se conformer au même type d'évolution. Néanmoins aucun modèle actuel ne permet de décrire une évolution aussi forte (4 à 6 fois plus importante que pour le modèle auto-similaire) depuis z 0.5[2].

Si l'on inclue maintenant la fonction de sélection dans le modèle de vraisemblance, l'évolution observée diminue drastiquement, du fait de la correction du biais de Malmquist. La tendance pour les amas C1 devient alors statistiquement indissociable (pour nos données limitées) du model auto-similaire.

### 5.3.3 Limitations

De nombreuses dégénérescences viennent nuancer ces conclusions.

Tout d'abord, l'amplitude du biais de Malmquist dépend fortement de la valeur de dispersion supposée autour de la relation $L_X$-$T$. Or ce paramètre est mal contraint. Si la mesure obtenue par Stanek et al. (2006) implique une évolution auto-similaire, d'autres valeurs autorisées indiqueraient une évolution nulle ou à l'inverse un impact limité des effets de sélection. Pour assurer la cohérence de l'ajustement, il faudrait déterminer la valeur de la dispersion directement avec l'échantillon utilisé, ce qui pourrait devenir possible avec les futurs échantillons C1 sur une surface plus importante.

D'autre part, les amas C1 ne permettent pas d'étudier l'évolution de la pente de la relation $L_X$-$T$. Bien que d'autres études (comme Lumb et al. 2004) ne trouvent

[2]On notera tout de même que, selon l'article récent de Nusser & Silk (2007), l'effet de la porosité du MIA induite par les cavité radio permettrait d'aller dans ce sens.

pas de variation de pente significative avec le redshift, un tel phénomène pourrait aisément imiter une évolution en normalisation pour notre échantillon.

Le profile fixe de ($R_c = 180$ kpc, $\beta = 2/3$) utilisé pour appliquer la fonction de sélection constitue aussi une hypothèse très simplificatrice. Une étude détaillée du profil des groupes à redshift intermédiaire permettrait d'utiliser un modèle plus réaliste et deviendra également possible avec les échantillons futurs. Une première analyse de ce type est actuellement en cours : il s'agit d'empiler les données (préalablement ré-échelonnée) issues des systèmes de températures et redshifts similaires parmi les amas C1 (Alshino et al, en préparation).

Pour finir, il convient de remarquer que toute incertitude sur les échantillons de référence locaux (rayon utilisé pour définir la luminosité, biais de sélection) se répercute de manière directe sur l'analyse de l'évolution.

### 5.3.4 Perspectives d'amélioration

La première nécessité pour dépasser le présent résultat serait d'augmenter le nombre d'amas de l'échantillon. Bien qu'il s'agisse d'une remarque triviale (quel observateur ne veut pas améliorer sa statistique ?), elle prend tout son sens dans le cas des relations d'échelle d'amas. Pour la plupart des spécialistes du domaine, il semble en effet primordial pour limiter les erreurs systématiques d'avoir recours à des observations profondes et détaillées de chaque amas. Cela a pour effet indirect de réduire la taille des échantillons considérés. Or, nos résultats montrent que les effets de sélection ont potentiellement un impact tout aussi important, et ne peuvent être corrigés sans une mesure précise de la dispersion - et donc un échantillonnage important du plan $L_X$-$T$ à différents redshifts.

Dans le même ordre d'idée, il est coutumier pour l'analyse des relations d'échelle de s'en tenir à des amas relaxés. La modélisation des effets de sélection associés à ce tri a postriori s'avère difficile et peut donc corrompre la correction du biais de Malmquist.

Récemment, Nord et al. (2008) ont montré que la corrélation entre les dispersions individuelles des relations $M$-$T$ et $M$-$L_X$ joue aussi un rôle dans le biais des échantillons limités en flux à travers un effet subtil qui met en jeu la fonction de masse des halos. Notre modèle de vraissemblance doit donc être étendu en conséquence.

Pour résumé, un échantillon idéal pour étudier les relations d'échelles devrait en principe :

- être associé à une fonction de sélection bien définie. Afin de ne pas corrompre le processus de sélection, cela implique d'effectuer une détection en aveugle des sources sur des champs aléatoirement sélectionnés.
- comprendre assez de sources pour contraindre la dispersion et la corrélation associées aux lois d'échelles.
- se baser sur des données suffisamment profondes, ou un suivi a posteriori pour pouvoir au minimum mesurer les masses d'un sous-échantillon représentatif et rigoureusement sélectionné.
- idéalement, inclure un nombre suffisant de sources dans l'univers proche pour

ré-évaluer les relations d'échelles locales dans un rayon clairement définis et en tenant compte des biais de sélection. Compte tenu de la faible densité d'amas massifs dans l'univers local, cela implique un relevé à très grande échelle (au moins 1/4 du ciel).

Il semble important de mentionner qu'aucun projet de ce type n'est prévu avec la génération actuelle d'instrument. En effet, les projets visant spécifiquement à contraindre les relations d'échelle comportent trop peu de sources pour mesurer de manière précise les dispersions et leur corrélation autour des loi moyennes. De plus, la manière dont ces échatillons sont sélectionnés ne permet généralement pas de modéliser les effets de sélection. A l'inverse, les relevés les plus volumineux, basés sur l'exploitation des archives, engloberont assez de sources, mais ne permetteront pas d'assembler facilement des échantillons locaux de manière controllée : la plupart des amas de l'univers proche sont déjà bien connus et observés volontairement avec les instruments récents. L'échantillon local correspondant devrait donc (pour estimer les effets de sélection) être extrait de relevés complémentaires. Dans cette perspective, les catalogues construits avec ROSAT [e.g. HIFLUGCS (Reiprich & Böhringer, 2002), 400 deg$^2$ (Burenin et al., 2007), REFLEX (Böhringer et al., 2004), RASS-SDSS (Popesso et al., 2004)] constituent une référence, mais la plupart ne seront pas réobservés systématiquement par les nouveaux instruments (à l'exception notable de HIFLUGCS). L'unique altenatives seraient alors d'attendre des futurs missions comme eROSITA (Predehl et al., 2006) qui prévoit un relevé englobant tout le ciel.

## 5.4 Analyse multi-longueur d'onde

De plus amples informations sur les propriétés des sources détectées ainsi que les biais de sélection associés peuvent émerger par l'analyse de données multi-longueur d'onde. Parmi les divers possibilités mentionnées dans la section 2.3, l'existence des données du CFHTLS sur la majeure partie du champ XMM-LSS permet d'étudier en premier lieu la distribution des galaxies, ainsi que les effets de lentilles gravitationnelles qu'elles subissent.

### 5.4.1 Les amas en optique

Pour l'article Mazure et al. (2007), un catalogue de 44 candidats amas fut constitué sur la base des données D1 du CFHTLS (1 $deg^2$), par l'identification de surdensités spatiales dans des tranches de redshift photométrique entre $0.1 < z < 1.25$. Les mesures spectroscopiques du VVDS disponibles dans cette région démontrent la fiabilité du procédé adopté : une surdensité dans la même tranche de redshift a été identifiée pour 15 des 16 amas pour lesquels des données sont disponibles.

La comparaison de ce catalogue avec les détections X que nous présentons dans Pierre et al. (2006, inséré section 7.3) est excellente : 2 des amas C1 sont détectés à un niveau dépassant $3\sigma$, et les 6 autres à plus que $4\sigma$. La moitié des amas C3 sont détectés à $3\sigma$. Parmi les 3 sources restantes, une se situe proche

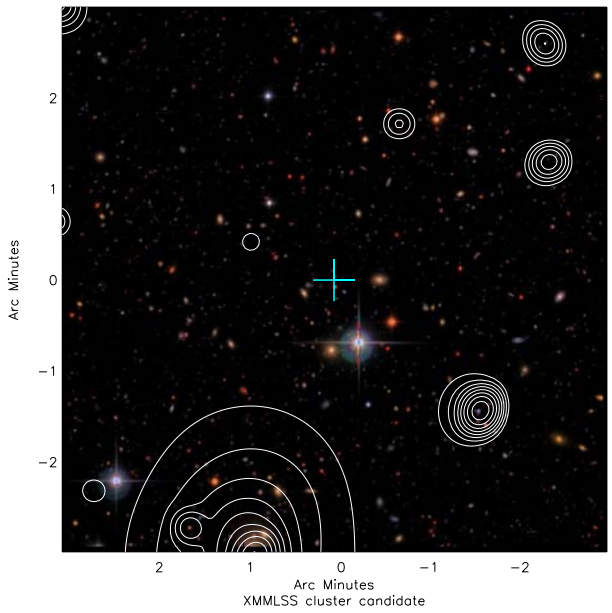


**FIG. 5.3**: *La détection WL03 de Bergé et al. (2008, inséré section 7.7). Les contours X issus du filtrage en ondelette sont superposé à l'image en vrai couleur dérivée des bandes (g,r,i) du CFHTLS. La croix bleu indique la position de la détection lensing.*

d'une région masquée, et les deux autres coincident avec des surdensités atteignant $1.5\sigma$. L'analyse optique permet de réveler 11 autres détections dont 5 semblent confirmées par les données VVDS. Ces résultats sont compatibles avec la meilleure sensibilité aux systèmes de faible masse que l'on pourrait attendre d'un relevé optique profond.

D'autres systèmes optique, non détectés en X, ont été identifiées en dehors du champ D1 via leur séquence rouge [Andreon et al. (2004), Andreon et al. (2005), Andreon et al. (2008)]. Tous semblent suffisamment peu riches (comparé à leur redshift) pour ne pas remettre en cause notre estimation de la complétude associée aux catalogues X.

### 5.4.2 L'analyse de cisaillement gravitationnel

Notons, pour commencer, qu'aucun amas du catalogue XMM-LSS actuel ne présente d'arc gravitationnel. Un tel phénomène est en effet peu probable, compte tenu de la distribution en masse de nos amas.

Dans l'article, Bergé et al. (2008, inséré section 7.7), nous présentons les détections de lentilles gravitationnels effectuées en aveugle sur 4 $deg^2$ du CFHTLS en bande $i'$. La région considérée inclue le champs D1 et recouvre la majeure partie du catalogue X. Il s'agit du premier échantillon d'amas rendu public dont la construction s'appuie sur la méthode des shapelets [Refregier & Bacon (2003),

Massey & Refregier (2005)].

### Echantillon et effets de sélection

L'analyse révèle 6 détections dans le champ 'Deep' contre aucune dans les 4 $deg^2$ du 'wide' pour un seuil de $2.5\sigma$. Parmi celles-ci, 3 sont des amas C1 présentés dans Pacaud et al. (2007, inséré section 7.5). Un quatrième amas est classifié comme C1 à partir des données XMM de l'AO5 (le pointé correspondant était contaminé lors des précédentes observations). Les deux autres ne présentent par contre aucune contrepartie X.

La différence de taux de détection entre les champs 'Deep' et 'Wide' résulte directement de la profondeur relative des deux relevés, et peut être justifiée à l'aide d'un modèle analytique présenté dans l'article. Ce même modèle permet d'expliquer la majorité des amas C1 non détectés, les quelques sources restantes se situant près de régions masquées ou en bordure de champ.

L'existence des deux détections lensing sans contrepartie X, WL02 et WL03, semble plus complexe à interpréter. Compte tenu du faible seuil en S/N, il pourrait s'agir de fausses détections. WL03 coincide cependant avec une source à $4\sigma$ identifiée par Gavazzi & Soucail (2007) à partir d'une analyse des données par la méthode KSB. Le signal qu'ils détectent est suffisament significatif pour associer à la source un redshift tomographique d'environ 0.3. Ce cas mérite donc une étude plus détaillée.

Comme l'indique la Fig. 5.3, le candidat amas n'est associé à aucune surdensité évidente de galaxies dans les image CFHTLS. Par contre, l'amas XLSSC 041 se trouve à moins de $4'$ et une étoile saturée à $\sim 1'$. Il reste donc plausible que le signal observé soit artificiel. Dans le cas contraire, la détection pourrait provenir d'effets de projection, à moins qu'il ne s'agisse d'un amas en cours d'éffondrement, associé aux galaxies éparses visibles sur l'image optique, dont le gaz n'a pas enocre pu atteindre une température suffisante pour émettre en X.

### Relation M-T

On observe généralement un grande dispersion entre les masses lensing et les propriétés du gaz dans les amas. La majorité des études disponibles à ce jour se concentrent sur un intervalle de masse restreint. L'information sur la pente se trouve alors noyée dans le bruit [e.g. Bardeau et al. (2007), Smith et al. (2005)]. La combinaison de nos 4 quatre groupes avec l'échantillon d'amas massifs de Bardeau et al. (2007) permet de couvrir l'intervalle de temperaure [1-10] keV et pour la première fois d'apporter de fortes contraintes sur la relation $M_{200}$-$T_X$ provenant du lensing. En supposant une évolution auto-similaire, on obtient la relation locale :

$$\frac{M_{200}}{10^{14}h^{-1}M_{\odot}} = 2.71^{+0.79}_{-0.61}\left(\frac{T}{4\,keV}\right)^{1.60\pm0.44} \tag{5.3}$$

Cette loi est remarquablement similaire à celles obtenues par Arnaud et al. (2005) et Vikhlinin et al. (2006) à partir de données X seules. Les contraintes sur la pente

dérive essentiellement des 4 nouveaux amas peu massifs situé entre $0.14 < z < 0.50$. Ce résultat est donc quelque peu dégénéré avec l'évolution auto-similaire supposée.

## 5.5 Implications cosmologiques des amas C1

Compte tenu des dégénérescences déjà mentionnées, la taille de l'échantillon C1 actuellement disponible ne semble pas permettre d'établir des contraintes cosmologiques robustes. Il s'avère néanmoins instructif de comparer les amas observés avec les prévisions des modèles disponibles, reflétant la compréhension actuelle dont nous disposons sur ces objets, leur formation et leur évolution. Pour celà, on se restreindra à des univers de types $\Lambda CDM$ plats dans lesquels l'énergie sombre correspond à une constante cosmologique.

### 5.5.1 Modèle phénoménologique

Pour déterminer la population d'amas C1 attendue, on utilise la combinaison d'un modèle de halos semi-analytique de type Press-Schechter, dont chaque composante fut validée par des simulations numériques, avec des relations d'échelles observationnelles. Une description détaillée est présentée dans Pacaud et al. (2007, inséré section 7.5), mais nous en listons ici les principales étapes :

- La forme du spectre de puissance linéaire à z=0 est d'abord calculée pour un indice spectral primordial, $n_s$, donné en supposant une fonction de transfer (Bardeen et al. 1986 ;(Sugiyama, 1995)). Puis on le normalise à la valeur adoptée pour $\sigma_8$.
- On en déduit ensuite la fonction de masse au redshift considéré via l'expression analytique de Jenkins et al. (2001), en faisant évoluer le spectre de puissance local suivant le facteur de croissance linéaire de Carroll et al. (1992) .
- Un profil NFW, dont la concentration est dérivée du modèle de Bullock et al. (2001) permet de convertir la masse des halos en $M_{500}$
- Puis les relations $M_{500}$-$T$ de Arnaud et al. (2005) et $L_X$-$T$ de Arnaud & Evrard (1999) sont appliquées, en supposant une évolution auto-similaire. Pour tenir compte de l'augmentation de pente à basse température observée par Arnaud et al. (2005), on introduit artificiellement une cassure à 3 keV. En deçà, la masse varie comme $T^{1.9}$.
- Pour tenir compte de la dispersion des relations d'échelle, on redistribue la luminosité des halos selon une loi log-normale compatible avec Stanek et al. (2006) autour de la valeur moyenne du modèle.
- Finalement, le flux correspondant est estimé. On peut prendre en compte la fonction de sélection en supposant un rayon de coeur fixe de 180 kpc pour tous les amas[3].

[3]On pourra trouver une approximation similaire dans Adami et al. (2000), où ils fixent $R_c = 250\ kpc$ dans un univers Einstein - de Sitter

| $\Omega_m$ | $100\Omega_b h^2$ | $h$ | $n_s$ | $\sigma_8$ |
|---|---|---|---|---|
| $0.237 \pm 0.034$ | $2.23 \pm 0.075$ | $0.735 \pm 0.032$ | $0.951 \pm 0.016$ | $0.74 \pm 0.05$ |

**Tab. 5.1**: *Contraintes individuelles sur les paramètres cosmologiques provenant uniquement de l'analyse des données CMB du satellite WMAP (selon la Table 5 de Spergel et al. 2007). L'espace est supposé plat, ce qui implique* $\Omega_\Lambda = 1 - \Omega_m$ *pour les redshift considérés dans cette thèse (z <2).*

On obtient ainsi la distribution des halos dans l'espace ($z$,$L_X$,$T$). La surface actuellement disponible ne permettant pas d'estimer la fonction de corrélation des amas, ce modèle suffit à décrire toutes les propriétés observables des sources C1.

### 5.5.2 Application aux amas C1

Nous comparons ici la distribution des amas C1 avec les prévisions associées au jeu de paramètres cosmologiques déterminés par WMAP [Spergel et al. (2007), voir Table 5.1].

#### Distribution en redshift

Les densités totales ainsi que les distributions en redshift se révèlent compatibles, dans les erreurs associées au bruit de Poisson. Seul un déficit d'amas autour de $z \sim 0.5$ requiert de prendre en compte la variance cosmique pour en trouver une explication.

Ces résultats semblent très sensibles aux différentes suppositions sur les relations d'échelles. En particulier, l'incertitude sur la niveau de dispersion autorisé autour des relations d'échelle moyennes a un impact important : La faire varier dans les marges d'erreur à $1\sigma$ de Stanek et al. (2006) modifie de plus de 50% la densité totale prédite.

Dans la plupart des travaux qui présentent des contraintes cosmologiques basées sur les amas X, l'impact de cette dégénérescence se révèle une des principales sources d'erreur [e.g. Schuecker et al. (2003), Pierpaoli et al. (2003)]. L'effet est généralement estimé '*pour l'exemple*' en considérant, comme nous venons de le faire, les valeurs extrême permises pour la dispersion ; mais le résultat est rarement inclu dans l'incertitude de mesure cumulée. De plus, il n'existe à ce jour aucune tentative visant à contraindre simultanément la dispersion, les relations d'échelles et les paramètres cosmologiques de manière auto-cohérente, tel qu'il faudrait le faire dans le cadre des relevés auto-calibrés [Majumdar & Mohr (2003), Hu (2003)].

Une autre dégénérescence importante concerne l'incertitude sur la physique des groupes. Compte tenu de la sensibilité croissantes des instruments et de la forme de la fonction de masse, les prochains relevés seront, tout comme le XMM-LSS, dominés par des systèmes peu massifs. Or les relations d'échelle dans le régime des groupes restent mal déterinées. Dans le cas de notre model, ne pas introduire la coupure à 3 keV de la loi $M$-$T$(ce qui revient à adopter la relation

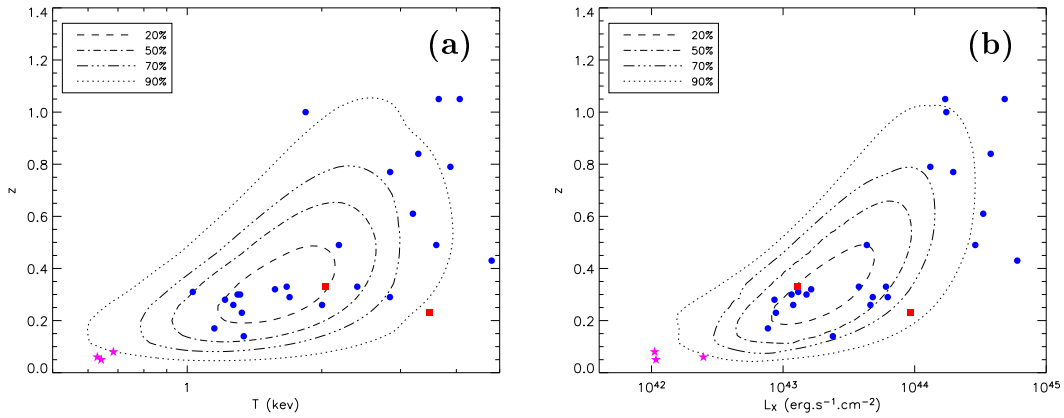


**FIG. 5.4**: *Distribution des amas C1 dans l'espace $L_X/T/z$. Les carrés rouges représentent des amas dont l'émission X est potentiellement contaminée, et les étoiles roses correspondent aux groupes locaux. Nos prévisions dans le cadre du modèle WMAP3 sont illustrées par des contours d'isodensité englobant les fractions de la population indiquées dans la légende.5.5.1.*

de Vikhlinin et al. 2006) diminue d'un facteur deux le nombre d'amas attendus. La sensibilité des comptages aux processus non-gravitationnels qui influencent les groupes constitue un argument supplémentaire pour l'étude de ces systèmes, à moins d'ignorer délibérément plus de la moitié des sources détectées dans les relevées X à venir.

#### Distribution dans le plan z-L-T

La Fig. 5.4 montre la distribution complète de nos sources dans l'espace des paramètres mesurables pour les amas C1. La distribution globale des amas en température et luminosité est au premier ordre comparable avec les contours de densité issu du modèle. On observe cependant un excès d'amas avec $L_X > 4{\times}10^{44}$erg.s$^{-1}$et $T > 3.5$ keV. Celui-ci semble marginalement significatif : la probabilité qu'une telle surdensité provienne du bruit de Poisson est de 20% sur la température, et 9% sur la luminostité. Il semble difficile de déterminer l'origine de cette déviation avec le peu de sources disponibles. Si l'on se fie aux paramètres estimés pour WMAP3, des cause probables seraient une déviation des relations d'échelle à grand redshif (par rapport au prédiction auto-similaires utilisées ici), ou une modélisation inprécise du profil des sources en vue d'appliquer la fonction de sélection. Tester ces deux possibilités nécessiterait des observations complémentaires.

### 5.5.3 Contraindre $\sigma_8$ et $\Omega_m$

Pour finir, considérons en tant que cas d'école les contraintes sur les paramètres cosmologiques fournies par la distribution des amas C1. Comme les 25 sources de

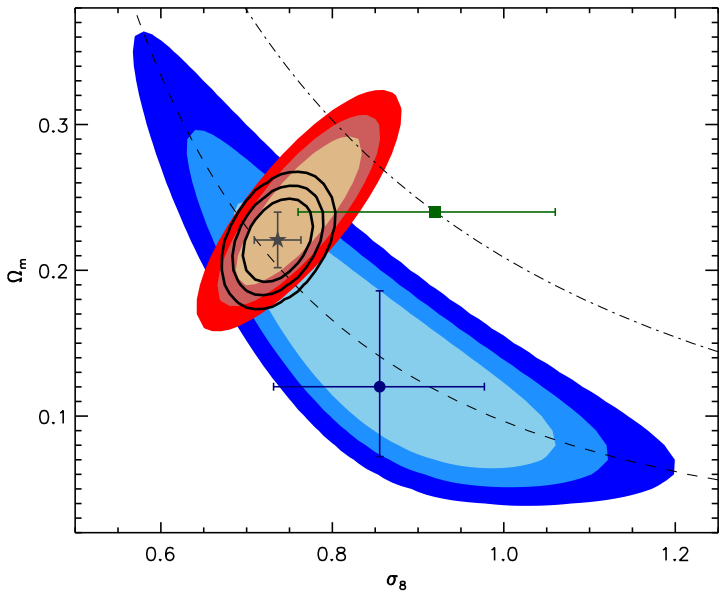


**FIG. 5.5**: *Contraintes dans le plan $\sigma_8$-$\Omega_m$. Les ellipses bleu indiquent les contraintes jointes issues de la distribution en redshift des amas C1, tandis que le rond bleu et les barres d'erreurs associées montrent les contraintes individuelles à $1\sigma$. Les ellipses rouges correspondent à la mesure de WMAP et les contours noirs à la combinaison C1+WMAP. L'étoile grise et les barres d'erreurs associées montrent l'incertitude finale sur chacun des paramètres. Le carré vert indique la mesure de $\sigma_8$ par Bergé et al. (2008).*

l'échantillon avec $z > 0.1$ ne sont pas assez nombreuses pour contraindre l'ensemble des paramètres libres, on se contentera de mesurer $\sigma_8$ et $\Omega_m$ en supposant tous les autres paramètres connus. Puisque nous fixons les relations d'échelles des amas, il y aurait peu de sens à prendre en compte la distribution complète en température/luminosité des sources, avec toutes les dégénérescences que celà implique. On ne considerera donc que la distribution en redshift des amas C1.

La probabilité d'observer N sources à des redshift $z_1, ..., z_N$ pour une distribution $\frac{dn}{dz}(z)$ donnée s'écrit alors :

$$P = \mathrm{Poi}(N, N_{tot}) \times \prod_{i=1}^{N} d(z_i) \tag{5.4}$$

où $N_{tot} = \int_{z_{min}}^{z_{max}} \frac{dn}{dz} dz$ est le nombre total d'amas prédit, $d(z) = \frac{1}{N_{tot}} \frac{dn}{dz}(z)$ leur distribution, et $\mathrm{Poi}(x, y)$ une loi de Poisson de paramètre $y$.

On en déduit la vraisemblance pour une réalisation de Poisson :

$$\mathcal{L} = \exp\left(-N_{tot}\right) \times \prod_{i=0}^{N_c} \frac{dn}{dz}(z_i) \tag{5.5}$$

En pratique, on utilise $z_{min}$=0.1 (compte tenu de l'exclusion des groupes locaux de l'échantillon) et $z_{max}$=1.4 car la densité d'amas attendue devient alors négli-

| Parameter | C1 | C1 + WMAP3 |
|---|---|---|
| $\sigma_8$ ..... | $0.85 \pm 0.12$ | $0.736 \pm 0.027$ |
| $\Omega_m$ ..... | $0.12^{+0.07}_{-0.05}$ | $0.22 \pm 0.02$ |

**TAB. 5.2**: *Contraintes individuelles ($1\sigma$, après marginalisation) sur $\sigma_8/\Omega_m$ obtenues avec la distribution en redshift des amas C1 et les données WMAP. Les autres paramètres du modèle sont fixés aux valeurs de la Table 5.1.*

geable. La valeur de $\Omega_m$ ainsi obtenue est basse mais les contraintes à un niveau de confiance de 68% sont compatibles avec le modèle WMAP3 (voir Figure 5.5 et Table 5.2).

La dégénérescence $\sigma_8$-$\Omega_m$ obtenue (illustrée par la ligne en tirait sur la Fig. 5.5) suit approximativement la relation :

$$\sigma_8 \times (\Omega_m/0.24)^{0.41} = 0.68 \tag{5.6}$$

Cette dépendance fonctionnelle diffère de la loi $\sigma_8 \propto \Omega_m^{-0.6}$ fréquemment utilisée pour modéliser les contraintes provenant des amas. Cette dernière relation provient à l'origine de la densité d'amas optique riches dans l'univers local (White et al., 1993). Bien qu'elle soit souvent mentionnée dans les études en bande X, il semble important de rappeler qu'elle n'a aucune raison a priori de s'appliquer à des échantillons sélectionnés par d'autres méthodes et sur des intervalles de redshift plus larges (contrairement aux suppositions de nombreux articles sur le sujet, e.g. Pierpaoli et al. 2003).

Afin d'améliorer ces résultats, on peut aussi combiner la vraisemblance issue des amas C1 avec les contraintes WMAP. Pour cela, on modelise la vraisemblance issue des données WMAP par une loi gaussienne bivariée avec un coefficient de corrélation de 0.8 [4]. L'échantillon C1 permet ainsi de diviser d'un facteur proche de 2 les incertitudes provenant de WMAP (Table 5.2).

Ce dernier résultat reste conditionel à la validité des relations d'échelles utilisées, et aux paramètres $H_0$, $\Omega_b$ et $n_s$ sur lesquelles les contraintes n'ont pas été margnalisées. En outre, la présente analyse ne tient pas compte de la variance cosmique qui engendre des fluctutations d'amplitude similaire au bruit de Poisson pour notre échantillon.

### 5.5.4 Données supplémentaires

Nombre de directions peuvent être envisagées afin d'améliorer les résultats présentés dans cette section. Celles-ci sont généralement tributaire de l'acquisition de nouvelles données permettant soit d'accroître l'échantillon, soit de mieux connaître les amas déjà détectés. Je conclurai donc ce chapitre par une description des données aditionnelles disponibles ou en cours d'aquisition.

[4] combiné aux paramètre de la Table 5.1, cela permet de reproduire qualitativement les contours de vraisemblance réels tracés par exemple dans Lesgourgues et al. (2007)

### Les observations XMM de l'AO5

L'observation des 46 nouveaux pointés obtenus lors de l'AO5 a pris fin depuis plusieurs mois et les données ont depuis été réduites. Celles-ci permettent de doubler la surface disponible et donc d'augmenter significativement le nombre de source.

La contamination de ces pointés par des particules de la magnétosphère se révèle beaucoup plus élevée que dans les observations précédentes. Le temps d'exposition moyen diminue de 20% par rapport aux observations de l'AO1 et AO2. Le nombre de sources ponctuelles détectées diminue en conséquence d'environ 65 par pointé à $\sim$50.

Même après le filtrage, le niveau moyen de fond dans les données restantes est plus élevé.Cela se répercute directement sur les détections d'amas. Sur une surface équivalente, on ne détécte que 15 (resp. 29) source C1 (resp. C2). De plus, la classe C1 comprend 4 détections qui semblent résulter de la confusion de paires de sources ponctuelles, tandis qu'aucune fausse détection n'avait été identifiée précédemment. Le comité d'allocation de temps d'XMM vient d'accepter au titre de l'AO7 la ré-observation des 19 pointés les plus contaminés (ceux pour lesquels le temps d'exposition utilisable pour le PN était inférieur à 7ks). Cela permettra de mieux comprendre les détections préliminaires, ainsi que de rendre plus homogène la couverture du relevé.

Les données du SXDS ont elles aussi été traitées, permettant la détection de 2 autres amas C1 et 3 candidats C2. Bien qu'elles soient plus profondes que les observations du XMM-LSS, la densité de sources atteinte y est similaire. En effet, la version actuelle du pipeline XMM-LSS ne peut traiter la confusion dans les champs profonds et il fut nécessaire de découper les observations en tranches de 20 ks.

### Observations X profondes

Dans un article récemment soumis à MNRAS (Maughan et al., soumis), l'amas XLSSC 029 a été ré-analysé à partir d'observations profondes XMM (80 ks) et Chandra (130 ks). Ce travail a notamment permis d'obtenir le premier profil de température pour un amas à z=1. La température moyenne mesurée est maintenant de T=$3.7^{+0.2}_{-0.1}$ keV pour une luminosité $L_{500} = (4.9 \pm 0.3)^{44}$ erg.s$^{-1}$, ce qui apporte une crédibilité certaine à notre précédent résultat de T=$4.1^{0.9}_{0.7}$ keV, $L_{500} = (4.8 \pm 0.4)^{44}$ erg.s$^{-1}$basé sur les donnés du relevé et, par extension, à l'ensemble de nos mesures.

D'autres ré-observations profondes d'amas détectés par le XMM-LSS ont été proposées, sans succès jusqu'à présent. Ces demandes de temps s'axent sur deux thématiques : la physique des groupes à grand redshift (Chandra), ainsi qu'une étude approfondie avec XMM des amas les plus massifs pour lesquels des suivis radio/mm par effet SZ sont actuellement organisés.

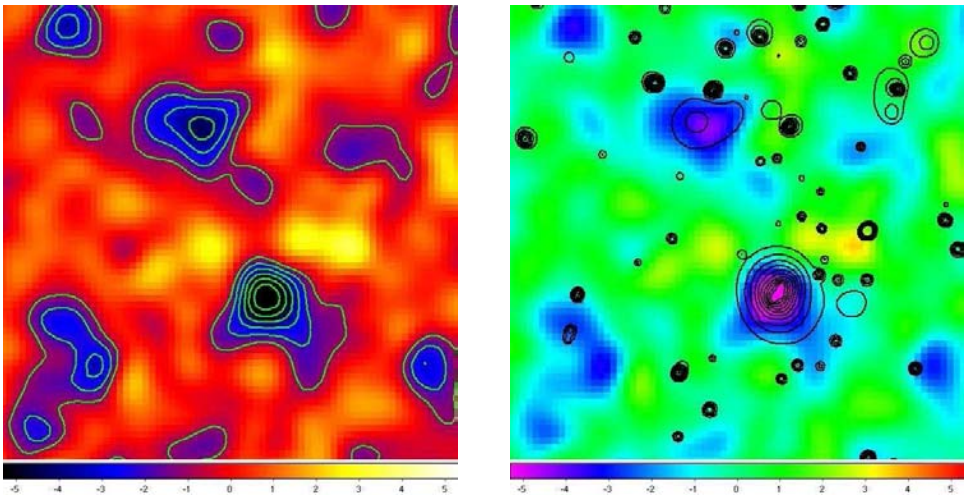

**FIG. 5.6**: *Cartes de rapport signal sur bruit extraites des observations APEX-SZ. Gauche : les contours indiquent des niveaux de S/N allant de -1 à -6 par saut de 1. Le décrément le plus significatif est XLSSC 006. L'autre décrément significatif au nord correspond à un amas détecté en optique via sa séquence rouge, RzCS 059, et publié dans Andreon et al. (2005). Ce dernier a un redshift photométrique* $z_{phot} \sim 0.9$. *Droite : l'émission X filtrée en ondelette est indiquée par les contours noir espacés de manière logarithmique.*

### Autres longueurs d'onde

Plusieurs projets d'observation de l'effet Sunyaev-Zel'dovich associés au XMM-LSS viennent de débuter. En particulier, OCRA acquiert actuellement des données à 30 GHz sur les amas les plus massifs du relevé tandis que APEX-SZ a commencé à cartographier le champs W1 à 150 GHz. Les observations APEX-SZ couvrent à ce jour un champ de 0.7 $\deg^2$ avec une sensibilité de 15$\mu$K rms, centré sur l'amas C1 le plus massif, XLSSC 006. La Fig. 5.6 donne un aperçu de ces nouvelles données. Cela permettra prochainement des études conjointes d'amas dans plusieurs longueurs d'onde (X, SZ, lentilles gravitationnelles). La mesure de température du gaz par spectroscopie X pourra par exemple être comparée aux résultats d'une déprojection simultannée des données X et SZ.

Par ailleurs, les observations XMM recouvrent maintenant la totalité du champ SWIRE de 10 $deg^2$. Un catalogue combiné SWIRE/CFHTLS vient d'être assemblé afin de travailler sur les propriétés des galaxies et AGNs dans le champ. Il deviendra ainsi possible d'étudier plus précisémment l'effet du milieu intra-amas sur les galaxies et d'augmenter la qualité des redshifts photométriques. En ce qui concerne la recherche d'amas, l'identification des amas à haut redshift sera facilitée par la corrélation des sources SWIRE avec l'émission X (voir par exemple Bremer et al. 2006). La confirmation des candidats amas à l'aide de redshifts photométiques constitue une solution de remplacement peu onéreuse aux observations spectroscopiques.

# Chapitre 6
# Conclusions et perspectives

Au cours de ce travail de thèse, nous avons établi et validé un ensemble de méthodes et de programmes permettant de maitriser de bout en bout l'exploitation cosmologique d'un relevé X profond tel que le XMM-LSS : depuis la réception des données jusqu'à l'extraction de contraintes cosmologiques, en passant par la détection et la caractérisation des sources. Une attention particulière a été portée sur la validation à l'aide de simulations de chaque démarche adoptée, ainsi que sur le contrôle des effets de sélection et leur impact sur toute mesure ultérieure. Il en résulte un large échantillon d'AGNs utilisables pour des études statistiques.

Ce travail à permi d'assembler le seul échantillon d'amas XMM public à ce jour dont la fonction de sélection est connue. Il s'agit de plus du premier dont la température ait été mesurée pour toutes les sources directement à partir des données du relevé. La modélisation cosmologique de cet échantillon a révélé d'importantes sources de biais, négligées jusqu'à présent, dans l'analyse de la relation $L_X$-$T$ et nous avons décrit un formalisme permettant de s'en affranchir.

L'analyse conjointe des relations d'échelle, de la densité et de la distribution des sources constitue une première étapes sur le chemin des contraintes cosmologiques déduites des amas par la méthode d'auto-calibration des relevés.

Nos résultats mettent également en relief les nombreuses dégénérescences restant à lever dans le but d'utiliser les amas comme sonde cosmologique de précision. Ces limitations nous permettent de formuler un ensemble de prescriptions nécessaires à l'obtention de contraintes auto-suffisantes avec les amas de galaxies.

Les effets de sélections inhérent à la construction de larges échantillons ont généralement des conséquences subtiles. Pour comprendre les biais qui peuvent en résulter, il faut tenir compte non-seulement les relations d'échelles des amas, mais également de la dispersion autour de ces tendances moyennes, et parfois même de la corrélation entre les dispersions sur ces différentes caractéristiques (c.f. Nord et al., 2008). Il en résulte que l'auto-calibration des relevés nécessite un échantillonnage important de l'espace des paramêtres décrivant les amas, afin de contraindre la distribution complète des sources. Cette donnée se trouve encore accentuée si l'on ajoute les effets de variance cosmique intervenant dans le processus de formation des structures. Il faut donc de nombreuses sources distribuées sur un volume important pour qu'un échantillon puisse être considéré comme représentatif.

Cette vision semble entrer en conflit avec la tendance actuelle à préférer des

observations profondes sur peu de sources pour mieux comprendre la physique du gaz en diminuant les erreurs systématiques. Il s'agit en fait de deux démarches complémentaires et nécessaires en vue d'une compréhension globale des objets étudiés. Sur ce point, l'exploitation des archives XMM et Chandra ainsi que les futurs missions prévoyant un relevé profond de tout le ciel apporteront une avancée significative.

Dans un compte-rendu de conférence récent soumis à rapporteur (Pierre et al., 2008, inséré section 7.8), nous montrons qu'un large relevé XMM a déjà le potentiel de répondre à certaines questions fondamentales : en particulier l'évolution de la relation $L_X$-$T$ et la confirmation de paramètres importants du modèle de concordance de par la densité et la distribution des amas à grand redshift.

Une autre approche essentielle consiste en la combinaison des observables. Si des observations profondes et détaillées dans une longueur d'onde permettent de mieux comprendre la physique des amas, il est indispensable d'en formuler une description cohérente compatible avec toutes les observations disponibles. Dans le cas des grands relevés, la combinaison de différentes observables fournit une méthode alternative pour quantifer et comprendre les systématiques et l'impact des effets de sélection.

Plusieurs travaux allant dans ce sens ont été mentionnés : la détection d'amas par redshift photométrique Mazure et al. (2007), la détection et mesure de masse par les effets de lentilles gravitationnelles Bergé et al. (2008), l'étude des propriétés optique et infra-rouge des galaxies d'amas (SWIRE/CFHTLS) ainsi que la mesure de l'effet Sunyaevv-Zel'dovich. D'un point de vue personnel, je m'engage dans cette dernière voie en rejoignant la collaboration APEX-SZ pour mon premier contrat Post-doctoral. La combinaison unique sur le champ XMM-LSS d'échantillons sélectionnés en X, SZ et optique devrait permettre d'inclure pour la première fois un test d'auto-cohérence sur les effets de sélections des différentes longueurs d'ondes et la physique des amas dans la dérivation de contraintes cosmologiques avec les amas de galaxies.

# Bibliographie


Adami, C., Slezak, E., Durret, F., Conselice, C. J., Cuillandre, J. C., Gallagher, J. S., Mazure, A., Pelló, R., Picat, J. P. & Ulmer, M. P., "Large scale diffuse light in the Coma cluster : A multi-scale approach", A&A, **429**, 39 (2005).

Adami, C., Ulmer, M. P., Romer, A. K., Nichol, R. C., Holden, B. P. & Pildis, R. A., "The Bright SHARC Survey : The Selection Function and Its Impact on the Cluster X-Ray Luminosity Function", ApJS, **131**, 391 (2000).

Allen, S. W., Rapetti, D. A., Schmidt, R. W., Ebeling, H., Morris, R. G. & Fabian, A. C., "Improved constraints on dark energy from Chandra X-ray observations of the largest relaxed galaxy clusters", accepté par MNRAS, `ArXiv :0706.0033` (2007).

Ameglio, S., Borgani, S., Pierpaoli, E. & Dolag, K., "Joint deprojection of Sunyaev-Zeldovich and X-ray images of galaxy clusters", MNRAS, **382**, 397 (2007).

Andreon, S., De Propris, R., Puddu, E., Giordano, L. & Quintana, H., "Scaling relations of the colour-detected cluster RzCS 052 at z=1.016 and of some other high redshift clusters", accepté par MNRAS, `ArXiv :0710.1023v1` (2008).

Andreon, S., Valtchanov, I., Jones, L. R., Altieri, B., Bremer, M., Willis, J., Pierre, M. & Quintana, H., "Batch discovery of nine $z\tilde{1}$ clusters using X-ray and K or R, z' images", MNRAS, **359**, 1250 (2005).

Andreon, S., Willis, J., Quintana, H., Valtchanov, I., Pierre, M. & Pacaud, F., "Galaxy evolution in clusters up to z= 1.0", MNRAS, **353**, 353 (2004).

Antonucci, R., "Unified models for active galactic nuclei and quasars", ARA&A, **31**, 473 (1993).

Arnaud, M., "X-ray observations of clusters of galaxies", dans F. Melchiorri & Y. Rephaeli (rédacteurs), *Background Microwave Radiation and Intracluster Cosmology*, p. 77–+ (2005).

Arnaud, M., Aghanim, N., Gastaud, R., Neumann, D. M., Lumb, D., Briel, U., Altieri, B., Ghizzardi, S., Mittaz, J., Sasseen, T. P. et al., "XMM-Newton observation of the Coma Galaxy cluster. The temperature structure in the central region", A&A, **365**, L67 (2001).

Arnaud, M. & Evrard, A. E., "The L_X-T relation and intracluster gas fractions of X-ray clusters", MNRAS, **305**, 631 (1999).

Arnaud, M., Pointecouteau, E. & Pratt, G. W., "The structural and scaling properties of nearby galaxy clusters. II. The M-T relation", A&A, **441**, 893 (2005).

Astier, P., Guy, J., Regnault, N., Pain, R., Aubourg, E., Balam, D., Basa, S., Carlberg, R. G., Fabbro, S., Fouchez, D. et al., "The Supernova Legacy Survey : measurement of $\Omega_M$, $\Omega_\Lambda$ and w from the first year data set", A&A, **447**, 31 (2006).

Balogh, M. L., Babul, A. & Patton, D. R., "Pre-heated isentropic gas in groups of galaxies", MNRAS, **307**, 463 (1999).

Barcons, X., Carrera, F. J. & Ceballos, M. T., "H1320+551 : a type 1.8/1.9 Seyfert galaxy with an unabsorbed X-ray spectrum", MNRAS, **339**, 757 (2003).

Barcons, X., Fabian, A. C. & Rees, M. J., "The physical state of the intergalactic medium", Nature, **350**, 685 (1991).

Bardeau, S., Soucail, G., Kneib, J.-P., Czoske, O., Ebeling, H., Hudelot, P., Smail, I. & Smith, G. P., "A CFH12k lensing survey of X-ray luminous galaxy clusters. II. Weak lensing analysis and global correlations", A&A, **470**, 449 (2007).

Bardeen, J. M., Bond, J. R., Kaiser, N. & Szalay, A. S., "The statistics of peaks of Gaussian random fields", ApJ, **304**, 15 (1986).

Barger, A. J., Cowie, L. L., Mushotzky, R. F., Yang, Y., Wang, W.-H., Steffen, A. T. & Capak, P., "The Cosmic Evolution of Hard X-Ray-selected Active Galactic Nuclei", AJ, **129**, 578 (2005).

Basilakos, S., Georgakakis, A., Plionis, M. & Georgantopoulos, I., "The Clustering of XMM-Newton Hard X-Ray Sources", ApJ, **607**, L79 (2004).

Basilakos, S. & Plionis, M., "Cosmological constraints from X-ray AGN clustering and Type Ia supernova data", MNRAS, **360**, L35 (2005).

Basilakos, S. & Plionis, M., "Constraining the Cold Dark Matter Spectrum Normalization in Flat Dark Energy Cosmologies", ApJ, **650**, L1 (2006).

Basilakos, S., Plionis, M., Georgakakis, A. & Georgantopoulos, I., "The XMM-Newton/2dF Survey - VI. Clustering and bias of the soft X-ray point sources", MNRAS, **356**, 183 (2005).

Baugh, C. M. & Efstathiou, G., "The Three-Dimensional Power Spectrum Measured from the APM Galaxy Survey - Part One - Use of the Angular Correlation Function", MNRAS, **265**, 145 (1993).

Becker, M. R., McKay, T. A., Koester, B., Wechsler, R. H., Rozo, E., Evrard, A., Johnston, D., Sheldon, E., Annis, J., Lau, E. et al., "The Mean and Scatter of the Velocity Dispersion-Optical Richness Relation for maxBCG Galaxy Clusters", ApJ, **669**, 905 (2007).

Bélanger, G., Goldwurm, A., Melia, F., Ferrando, P., Grosso, N., Porquet, D., Warwick, R. & Yusef-Zadeh, F., "Repeated X-Ray Flaring Activity in Sagittarius A*", ApJ, **635**, 1095 (2005).

Benjamin, J., Heymans, C., Semboloni, E., van Waerbeke, L., Hoekstra, H., Erben, T., Gladders, M. D., Hetterscheidt, M., Mellier, Y. & Yee, H. K. C., "Cosmological constraints from the 100-deg$^2$ weak-lensing survey", MNRAS, **381**, 702 (2007).

Benoît, A., Ade, P., Amblard, A., Ansari, R., Aubourg, É., Bargot, S., Bartlett, J. G., Bernard, J.-P., Bhatia, R. S., Blanchard, A. et al., "Cosmological constraints from Archeops", A&A, **399**, L25 (2003).

Bergé, J., Pacaud, F., Réfrégier, A., Massey, R., Pierre, M., Amara, A., Birkinshaw, M., Paulin-Henriksson, S., Smith, G. P. & Willis, J., "Combined analysis of weak lensing and X-ray blind surveys", accepté par MNRAS, `ArXiv :0712.3293` (2008).

Bernardeau, F., "Weak lensing detection in CMB maps.", A&A, **324**, 15 (1997).

Bertschinger, E., "Self-similar secondary infall and accretion in an Einstein-de Sitter universe", ApJS, **58**, 39 (1985).

Böhringer, H., Collins, C. A., Guzzo, L., Schuecker, P., Voges, W., Neumann, D. M., Schindler, S., Chincarini, G., De Grandi, S., Cruddace, R. G. et al., "The ROSAT-ESO Flux-limited X-Ray (REFLEX) Galaxy Cluster Survey. IV. The X-Ray Luminosity Function", ApJ, **566**, 93 (2002).

Böhringer, H., Schuecker, P., Guzzo, L., Collins, C. A., Voges, W., Cruddace, R. G., Ortiz-Gil, A., Chincarini, G., De Grandi, S., Edge, A. C. et al., "The ROSAT-ESO Flux Limited X-ray (REFLEX) Galaxy cluster survey. V. The cluster catalogue", A&A, **425**, 367 (2004).

Böhringer, H., Schuecker, P., Guzzo, L., Collins, C. A., Voges, W., Schindler, S., Neumann, D. M., Cruddace, R. G., De Grandi, S., Chincarini, G. et al., "The ROSAT-ESO flux limited X-ray (REFLEX) galaxy cluster survey. I. The construction of the cluster sample", A&A, **369**, 826 (2001).

Böhringer, H., Voges, W., Huchra, J. P., McLean, B., Giacconi, R., Rosati, P., Burg, R., Mader, J., Schuecker, P., Simiç, D. et al., "The Northern ROSAT All-Sky (NORAS) Galaxy Cluster Survey. I. X-Ray Properties of Clusters Detected as Extended X-Ray Sources", ApJS, **129**, 435 (2000).

Bonamente, M., Joy, M. K., LaRoque, S. J., Carlstrom, J. E., Reese, E. D. & Dawson, K. S., "Determination of the Cosmic Distance Scale from Sunyaev-Zel'dovich Effect and Chandra X-Ray Measurements of High-Redshift Galaxy Clusters", ApJ, **647**, 25 (2006).

Bond, J. R., Cole, S., Efstathiou, G. & Kaiser, N., "Excursion set mass functions for hierarchical Gaussian fluctuations", ApJ, **379**, 440 (1991).

Borgani, S., Murante, G., Springel, V., Diaferio, A., Dolag, K., Moscardini, L., Tormen, G., Tornatore, L. & Tozzi, P., "X-ray properties of galaxy clusters and groups from a cosmological hydrodynamical simulation", MNRAS, **348**, 1078 (2004).

Borgani, S., Rosati, P., Tozzi, P., Stanford, S. A., Eisenhardt, P. R., Lidman, C., Holden, B., Della Ceca, R., Norman, C. & Squires, G., "Measuring $\Omega_m$ with the ROSAT Deep Cluster Survey", ApJ, **561**, 13 (2001).

Branchesi, M., Gioia, I. M., Fanti, C. & Fanti, R., "High redshift X-ray galaxy clusters. II. The $L_X$-$T$ relationship revisited", A&A, **472**, 739 (2007).

Bremer, M. N., Valtchanov, I., Willis, J., Altieri, B., Andreon, S., Duc, P. A., Fang, F., Jean, C., Lonsdale, C., Pacaud, F. et al., "XMM-LSS discovery of a z = 1.22 galaxy cluster", MNRAS, **371**, 1427 (2006).

Bryan, G. L. & Norman, M. L., "Statistical Properties of X-Ray Clusters : Analytic and Numerical Comparisons", ApJ, **495**, 80 (1998).

Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R. & Dekel, A., "Profiles of dark haloes : evolution, scatter and environment", MNRAS, **321**, 559 (2001).

Burenin, R. A., Vikhlinin, A., Hornstrup, A., Ebeling, H., Quintana, H. & Mescheryakov, A., "The 400 Square Degree ROSAT PSPC Galaxy Cluster Survey : Catalog and Statistical Calibration", ApJS, **172**, 561 (2007).

Burwitz, V., Haberl, F., Neuhäuser, R., Predehl, P., Trümper, J. & Zavlin, V. E., "The thermal radiation of the isolated neutron star RX J1856.5-3754 observed with Chandra and XMM-Newton", A&A, **399**, 1109 (2003).

Carrera, F. J., Barcons, X., Fabian, A. C., Hasinger, G., Mason, K. O., McMahon, R. G., Mittaz, J. P. D. & Page, M. J., "Clustering of X-ray selected active galactic nuclei", MNRAS, **299**, 229 (1998).

Carrera, F. J., Ebrero, J., Mateos, S., Ceballos, M. T., Corral, A., Barcons, X., Page, M. J., Rosen, S. R., Watson, M. G., Tedds, J. A. et al., "The XMM-Newton serendipitous survey. III. The AXIS X-ray source counts and angular clustering", A&A, **469**, 27 (2007).

Carroll, S. M., Press, W. H. & Turner, E. L., "The cosmological constant", ARA&A, **30**, 499 (1992).

Carter, J. A. & Read, A. M., "The XMM-Newton EPIC background and the production of background blank sky event files", A&A, **464**, 1155 (2007).

Cash, W., "Parameter estimation in astronomy through application of the likelihood ratio", ApJ, **228**, 939 (1979).

Cavaliere, A. & Lapi, A., "Missing Baryons, from Clusters to Groups of Galaxies", ApJ, **673**, L5 (2008).

Clowe, D., Bradač, M., Gonzalez, A. H., Markevitch, M., Randall, S. W., Jones, C. & Zaritsky, D., "A Direct Empirical Proof of the Existence of Dark Matter", ApJ, **648**, L109 (2006).

Collins, C. A., Guzzo, L., Böhringer, H., Schuecker, P., Chincarini, G., Cruddace, R., De Grandi, S., MacGillivray, H. T., Neumann, D. M., Schindler, S. et al., "The ROSAT-ESO Flux-Limited X-ray (REFLEX) galaxy cluster survey - II. The spatial correlation function", MNRAS, **319**, 939 (2000).

Decourchelle, A., Sauvageot, J. L., Audard, M., Aschenbach, B., Sembay, S., Rothenflug, R., Ballet, J., Stadlbauer, T. & West, R. G., "XMM-Newton observation of the Tycho supernova remnant", A&A, **365**, L218 (2001).

Dicke, R. H., Peebles, P. J. E., Roll, P. G. & Wilkinson, D. T., "Cosmic Black-Body Radiation.", ApJ, **142**, 414 (1965).

Ebeling, H., Edge, A. C., Allen, S. W., Crawford, C. S., Fabian, A. C. & Huchra, J. P., "The ROSAT Brightest Cluster Sample - IV. The extended sample", MNRAS, **318**, 333 (2000).

Ebeling, H., Edge, A. C., Bohringer, H., Allen, S. W., Crawford, C. S., Fabian, A. C., Voges, W. & Huchra, J. P., "The ROSAT Brightest Cluster Sample - I. The compilation of the sample and the cluster log N-log S distribution", MNRAS, **301**, 881 (1998).

Ebeling, H. & Wiedenmann, G., "Detecting structure in two dimensions combining Voronoi tessellation and percolation", Phys. Rev. E, **47**, 704 (1993).

Eddington, A. S., "On a formula for correcting statistics for the effects of a known error of observation", MNRAS, **73**, 359 (1913).

Eisenstein, D. J. & Hu, W., "Baryonic Features in the Matter Transfer Function", ApJ, **496**, 605 (1998).

Eke, V. R., Baugh, C. M., Cole, S., Frenk, C. S. & Navarro, J. F., "Galaxy groups in the 2dF Galaxy Redshift Survey : the number density of groups", MNRAS, **370**, 1147 (2006).

Eke, V. R., Navarro, J. F. & Steinmetz, M., "The Power Spectrum Dependence of Dark Matter Halo Concentrations", ApJ, **554**, 114 (2001).

Ettori, S., "Note on a polytropic $\beta$-model to fit the X-ray surface brightness of clusters of galaxies", MNRAS, **311**, 313 (2000).

Ettori, S., Tozzi, P., Borgani, S. & Rosati, P., "Scaling laws in X-ray galaxy clusters at redshift between 0.4 and 1.3", A&A, **417**, 13 (2004).

Fabian, A. C., Allen, S. W., Crawford, C. S., Johnstone, R. M., Morris, R. G., Sanders, J. S. & Schmidt, R. W., "The missing soft X-ray luminosity in cluster cooling flows", MNRAS, **332**, L50 (2002).

Favata, F., Giardino, G., Micela, G., Sciortino, S. & Damiani, F., "An XMM-Newton-based X-ray survey of pre-main sequence stellar emission in the L1551 star-forming complex", A&A, **403**, 187 (2003).

Finoguenov, A., Böhringer, H. & Zhang, Y.-Y., "XMM-Newton study of the two-dimensional structure of the REFLEX-DXL galaxy clusters", A&A, **442**, 827 (2005).

Finoguenov, A., Reiprich, T. H. & Böhringer, H., "Details of the mass-temperature relation for clusters of galaxies", A&A, **368**, 749 (2001).

Forman, W., Jones, C., Cominsky, L., Julien, P., Murray, S., Peters, G., Tananbaum, H. & Giacconi, R., "The fourth Uhuru catalog of X-ray sources.", ApJS, **38**, 357 (1978).

Gandhi, P., Garcet, O., Disseau, L., Pacaud, F., Pierre, M., Gueguen, A., Alloin, D., Chiappetti, L., Gosset, E., Maccagni, D. et al., "The XMM large scale structure survey : properties and two-point angular correlations of point-like sources", A&A, **457**, 393 (2006).

Garcet, O., Gandhi, P., Gosset, E., Sprimont, P. G., Surdej, J., Borkowski, V., Tajer, M., Pacaud, F., Pierre, M., Chiappetti, L. et al., "The XMM large scale structure survey : optical vs. X-ray classifications of active galactic nuclei and the unified scheme", A&A, **474**, 473 (2007).

Gavazzi, R. & Soucail, G., "Weak lensing survey of galaxy clusters in the CFHTLS Deep", A&A, **462**, 459 (2007).

Ghizzardi, S., "In flight calibration of the PSF for the MOS1 and MOS2 cameras", XMM-SOC-CAL-TN-0022 (EPIC-MCT-TN-011), Available at http ://xmm.vilspa.esa.es/docs/documents/ (2001).

Ghizzardi, S., "In flight calibration of the PSF for the PN camera", XMM-SOC-CAL-TN-0029 (EPIC-MCT-TN-012), Available at http ://xmm.vilspa.esa.es/docs/documents/ (2002).

Gilli, R., Daddi, E., Zamorani, G., Tozzi, P., Borgani, S., Bergeron, J., Giacconi, R., Hasinger, G., Mainieri, V., Norman, C. et al., "The spatial clustering of X-ray selected AGN and galaxies in the Chandra Deep Field South and North", A&A, **430**, 811 (2005).

Gilli, R., Salvati, M. & Hasinger, G., "Testing current synthesis models of the X-ray background", A&A, **366**, 407 (2001).

Gioia, I. M., Maccacaro, T., Schild, R. E., Wolter, A., Stocke, J. T., Morris, S. L. & Henry, J. P., "The Einstein Observatory Extended Medium-Sensitivity Survey. I - X-ray data and analysis", ApJS, **72**, 567 (1990).

Giovannini, G., Tordi, M. & Feretti, L., "Radio halo and relic candidates from the NRAO VLA Sky Survey", New Astronomy, **4**, 141 (1999).

Gitti, M., Brunetti, G., Feretti, L. & Setti, G., "Particle acceleration in cooling flow clusters of galaxies : The case of Abell 2626", A&A, **417**, 1 (2004).

Gitti, M., Brunetti, G. & Setti, G., "Modeling the interaction between ICM and relativistic plasma in cooling flows : The case of the Perseus cluster", A&A, **386**, 456 (2002).

Gladders, M. D. & Yee, H. K. C., "A New Method For Galaxy Cluster Detection. I. The Algorithm", AJ, **120**, 2148 (2000).

Güdel, M., Audard, M., Briggs, K., Haberl, F., Magee, H., Maggio, A., Mewe, R., Pallavicini, R. & Pye, J., "The XMM-Newton view of stellar coronae : X-ray spectroscopy of the corona of <ASTROBJ>AB Doradus</ASTROBJ>", A&A, **365**, L336 (2001).

Gunn, J. E., "Massive galactic halos. I - Formation and evolution", ApJ, **218**, 592 (1977).

Gunn, J. E. & Gott, J. R. I., "On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution", ApJ, **176**, 1 (1972).

Guth, A. H., "Inflationary universe : A possible solution to the horizon and flatness problems", Phys. Rev. D, **23**, 347 (1981).

Haiman, Z., Allen, S., Bahcall, N., Bautz, M., Boehringer, H., Borgani, S., Bryan, G., Cabrera, B., Canizares, C., Citterio, O. et al., "An X-ray Galaxy Cluster Survey for Investigations of Dark Energy", Rapport soumis à la Dark Energy Task Force, `ArXiv :astro-ph/0507013` (2005).

Hasinger, G., Miyaji, T. & Schmidt, M., "Luminosity-dependent evolution of soft X-ray selected AGN. New Chandra and XMM-Newton surveys", A&A, **441**, 417 (2005).

Helsdon, S. F. & Ponman, T. J., "The intragroup medium in loose groups of galaxies", MNRAS, **315**, 356 (2000).

Henry, J. P., "A Measurement of the Density Parameter Derived from the Evolution of Cluster X-Ray Temperatures", ApJ, **489**, L1+ (1997).

Henry, J. P., Gioia, I. M., Mullis, C. R., Voges, W., Briel, U. G., Böhringer, H. & Huchra, J. P., "Overview of the ROSAT North Ecliptic Pole Survey", ApJ, **553**, L109 (2001).

Hetterscheidt, M., Simon, P., Schirmer, M., Hildebrandt, H., Schrabback, T., Erben, T. & Schneider, P., "GaBoDS : The Garching-Bonn deep survey. VII. Cosmic shear analysis", A&A, **468**, 859 (2007).

Hoekstra, H., Mellier, Y., van Waerbeke, L., Semboloni, E., Fu, L., Hudson, M. J., Parker, L. C., Tereno, I. & Benabed, K., "First Cosmic Shear Results from the Canada-France-Hawaii Telescope Wide Synoptic Legacy Survey", ApJ, **647**, 116 (2006).

Hogg, D. W. & Turner, E. L., "A Maximum Likelihood Method to Improve Faint-Source Flux and Color Estimates", PASP, **110**, 727 (1998).

Hu, W., "Self-consistency and calibration of cluster number count surveys for dark energy", Phys. Rev. D, **67**, 8, 081304 (2003).

Hubble, E., "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae", Proceedings of the National Academy of Science, **15**, 168 (1929).

Hubble, E. P., "NGC 6822, a remote stellar system.", ApJ, **62**, 409 (1925).

Jansen, F., Lumb, D., Altieri, B., Clavel, J., Ehle, M., Erd, C., Gabriel, C., Guainazzi, M., Gondoin, P., Much, R. et al., "XMM-Newton observatory. I. The spacecraft and operations", A&A, **365**, L1 (2001).

Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P. & Yoshida, N., "The mass function of dark matter haloes", MNRAS, **321**, 372 (2001).

Jones, C. & Forman, W., "The structure of clusters of galaxies observed with Einstein", ApJ, **276**, 38 (1984).

Jones, C. & Forman, W., "Einstein Observatory Images of Clusters of Galaxies", ApJ, **511**, 65 (1999).

Kaiser, N., "Evolution and clustering of rich clusters", MNRAS, **222**, 323 (1986).

King, I., "The structure of star clusters. I. an empirical density law", AJ, **67**, 471 (1962).

Koester, B. P., McKay, T. A., Annis, J., Wechsler, R. H., Evrard, A., Bleem, L., Becker, M., Johnston, D., Sheldon, E., Nichol, R. et al., "A MaxBCG Catalog of 13,823 Galaxy Clusters from the Sloan Digital Sky Survey", ApJ, **660**, 239 (2007).

Kolb, E. W., Matarrese, S. & Riotto, A., "On cosmic acceleration without dark energy", New Journal of Physics, **8**, 322 (2006).

Koopmans, L. V. E. & Fassnacht, C. D., "A Determination of $H_0$ with the CLASS Gravitational Lens B1608+656. II. Mass Models and the Hubble Constant from Lensing", ApJ, **527**, 513 (1999).

Krick, J. E. & Bernstein, R. A., "Diffuse Optical Light in Galaxy Clusters. II. Correlations with Cluster Properties", AJ, **134**, 466 (2007).

Langer, M., Puget, J.-L. & Aghanim, N., "Cosmological magnetogenesis driven by radiation pressure", Phys. Rev. D, **67**, 4, 043505 (2003).

LaRoque, S. J., Bonamente, M., Carlstrom, J. E., Joy, M. K., Nagai, D., Reese, E. D. & Dawson, K. S., "X-Ray and Sunyaev-Zel'dovich Effect Measurements of the Gas Mass Fraction in Galaxy Clusters", ApJ, **652**, 917 (2006).

Lesgourgues, J., Viel, M., Haehnelt, M. G. & Massey, R., "A combined analysis of 3D Weak Lensing, Lyman-alpha forest and WMAP year three data", JCAP, **11**, 8 (2007).

Limber, D. N., "The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field.", ApJ, **117**, 134 (1953).

Lumb, D. H., Bartlett, J. G., Romer, A. K., Blanchard, A., Burke, D. J., Collins, C. A., Nichol, R. C., Giard, M., Marty, P. B., Nevalainen, J. et al., "The XMM-NEWTON $\Omega$ project. I. The X-ray luminosity-temperature relation at z$>$ 0.4", A&A, **420**, 853 (2004).

Mainieri, V., Bergeron, J., Hasinger, G., Lehmann, I., Rosati, P., Schmidt, M., Szokoly, G. & Della Ceca, R., "XMM-Newton observation of the Lockman Hole. II. Spectral analysis", A&A, **393**, 425 (2002).

Majumdar, S. & Mohr, J. J., "Importance of Cluster Structural Evolution in Using X-Ray and Sunyaev-Zeldovich Effect Galaxy Cluster Surveys to Study Dark Energy", ApJ, **585**, 603 (2003).

Mallat, S., "A theory for multiresolution signal decomposition : the wavelet representation", IEEE Transaction on Pattern Analysis and Machine Intelligence, **11**, 674 (1989).

Mandelbaum, R., Seljak, U., Cool, R. J., Blanton, M., Hirata, C. M. & Brinkmann, J., "Density profiles of galaxy groups and clusters from SDSS galaxy-galaxy weak lensing", MNRAS, **372**, 758 (2006).

Markevitch, M., "The $L_X$-T Relation and Temperature Function for Nearby Clusters Revisited", ApJ, **504**, 27 (1998).

Markevitch, M., Gonzalez, A. H., David, L., Vikhlinin, A., Murray, S., Forman, W., Jones, C. & Tucker, W., "A Textbook Example of a Bow Shock in the Merging Galaxy Cluster 1E 0657-56", ApJ, **567**, L27 (2002).

Massey, R. & Refregier, A., "Polar shapelets", MNRAS, **363**, 197 (2005).

Maughan, B., Jones, L., Pierre, M., Andreon, S., Birkinshaw, M., Bremer, M., Pacaud, F., Ponman, T., Valtchanov, I. & Willis, J., "Testing the galaxy cluster mass-observable relations with Chandra and XMM-Newton observations of XLSSC 029", soumis à MNRAS, `ArXiv :0709.2300v1`.

Maughan, B. J., Jones, L. R., Ebeling, H. & Scharf, C., "The evolution of the cluster X-ray scaling relations in the Wide Angle ROSAT Pointed Survey sample at $0.6 < z < 1.0$", MNRAS, **365**, 509 (2006).

Mazure, A., Adami, C., Pierre, M., Le Fèvre, O., Arnouts, S., Duc, P. A., Ilbert, O., Lebrun, V., Meneux, B., Pacaud, F. et al., "Structure detection in the D1 CFHTLS deep field using accurate photometric redshifts : a benchmark", A&A, **467**, 49 (2007).

Mihos, J. C., Harding, P., Feldmeier, J. & Morrison, H., "Diffuse Light in the Virgo Cluster", ApJ, **631**, L41 (2005).

Mitchell, R. J. & Culhane, J. L., "Detection of iron line emission in the Ariel V X-ray spectrum of the Centaurus cluster of galaxies", MNRAS, **178**, 75P (1977).

Mitchell, R. J., Culhane, J. L., Davison, P. J. N. & Ives, J. C., "Ariel 5 observations of the X-ray spectrum of the Perseus Cluster", MNRAS, **175**, 29P (1976).

Mo, H. J. & White, S. D. M., "An analytic model for the spatial clustering of dark matter haloes", MNRAS, **282**, 347 (1996).

Moretti, A., Campana, S., Lazzati, D. & Tagliaferri, G., "The Resolved Fraction of the Cosmic X-Ray Background", ApJ, **588**, 696 (2003).

Morlet, J., Arens, G., Forgeau, I. & Giard, D., "Wave Propagation and Sampling Theory", Geophysics, **47**, 203 (1982).

Navarro, J. F., Frenk, C. S. & White, S. D. M., "A Universal Density Profile from Hierarchical Clustering", ApJ, **490**, 493 (1997).

Neto, A. F., Gao, L., Bett, P., Cole, S., Navarro, J. F., Frenk, C. S., White, S. D. M., Springel, V. & Jenkins, A., "The statistics of $\Lambda$ CDM halo concentrations", MNRAS, **381**, 1450 (2007).

Nord, B., Stanek, R., Rasia, E. & Evrard, A. E., "Effects of selection and covariance on X-ray scaling relations of galaxy clusters", MNRAS, **383**, L10 (2008).

Novicki, M. C., Sornig, M. & Henry, J. P., "The Evolution of the Galaxy Cluster Luminosity-Temperature Relation", AJ, **124**, 2413 (2002).

Nusser, A. & Silk, J., "Hydrostatic equilibrium of a porous intracluster medium : implications for mass fraction and X-ray luminosity", `ArXiv :0711.1281v1`.

Olsen, L. F., Benoist, C., Cappi, A., Maurogordato, S., Mazure, A., Slezak, E., Adami, C., Ferrari, C. & Martel, F., "Galaxy clusters in the CFHTLS. First matched filter candidate catalogue of the Deep fields", A&A, **461**, 81 (2007).

Pacaud, F., Pierre, M., Adami, C., Altieri, B., Andreon, S., Chiappetti, L., Detal, A., Duc, P.-A., Galaz, G., Gueguen, A. et al., "The XMM-LSS survey : the Class 1 cluster sample over the initial 5 deg$^2$ and its cosmological modelling", MNRAS, **382**, 1289 (2007).

Pacaud, F., Pierre, M., Refregier, A., Gueguen, A., Starck, J.-L., Valtchanov, I., Read, A. M., Altieri, B., Chiappetti, L., Gandhi, P. et al., "The XMM Large-Scale Structure survey : the X-ray pipeline and survey selection function", MNRAS, **372**, 578 (2006).

Penzias, A. A. & Wilson, R. W., "A Measurement of Excess Antenna Temperature at 4080 Mc/s.", ApJ, **142**, 419 (1965).

Percival, W. J., Nichol, R. C., Eisenstein, D. J., Weinberg, D. H., Fukugita, M., Pope, A. C., Schneider, D. P., Szalay, A. S., Vogeley, M. S., Zehavi, I. et al., "Measuring the Matter Density Using Baryon Oscillations in the SDSS", ApJ, **657**, 51 (2007).

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E. et al., "Measurements of Omega and Lambda from 42 High-Redshift Supernovae", ApJ, **517**, 565 (1999).

Perola, G. C., Puccetti, S., Fiore, F., Sacchi, N., Brusa, M., Cocchia, F., Baldi, A., Carangelo, N., Ciliegi, P., Comastri, A. et al., "The HELLAS2XMM survey. VI. X-ray absorption in the 1df AGN sample through a spectral analysis", A&A, **421**, 491 (2004).

Pierpaoli, E., Borgani, S., Scott, D. & White, M., "On determining the cluster abundance normalization", MNRAS, **342**, 163 (2003).

Pierre, M., Chiappetti, L., Pacaud, F., Gueguen, A., Altieri, B., Aussel, H., Gandhi, P., Garcet, O., Gosset, E., Paioro, L. et al., "The XMM-LSS catalogue : X-ray sources and associated optical data. Version I", MNRAS, sous presse (2007).

Pierre, M., Pacaud, F., Duc, P.-A., Willis, J. P., Andreon, S., Valtchanov, I., Altieri, B., Galaz, G., Gueguen, A., Fèvre, J.-P. L. et al., "The XMM Large-Scale Structure survey : a well-controlled X-ray cluster sample over the D1 CFHTLS area", MNRAS, **372**, 591 (2006).

Pierre, M., Pacaud, F., Melin, J. B. & the XMM-LSS consortium, "The XMM-LSS cluster sample and its cosmological applications. Prospects for the XMM next decade", accepé par Astronomische Nachrichten, `ArXiv :0712.0262v1` (2008).

Plionis, M., Rovilos, M., Basilakos, S., Georgantopoulos, I. & Bauer, F., "Luminosity Dependent X-ray AGN Clustering ?", accepté par ApJL, `ArXiv :0712.3072v1` (2008).

Pointecouteau, E., Arnaud, M. & Pratt, G. W., "The structural and scaling properties of nearby galaxy clusters. I. The universal mass profile", A&A, **435**, 1 (2005).

Polletta, M., Tajer, M., Maraschi, L., Trinchieri, G., Lonsdale, C. J., Chiappetti, L., Andreon, S., Pierre, M., Le Fèvre, O., Zamorani, G. et al., "Spectral Energy Distributions of Hard X-Ray Selected Active Galactic Nuclei in the XMM-Newton Medium Deep Survey", ApJ, **663**, 81 (2007).

Popesso, P., Biviano, A., Böhringer, H., Romaniello, M. & Voges, W., "RASS-SDSS galaxy cluster survey. III. Scaling relations of galaxy clusters", A&A, **433**, 431 (2005).

Popesso, P., Böhringer, H., Brinkmann, J., Voges, W. & York, D. G., "RASS-SDSS Galaxy clusters survey. I. The catalog and the correlation of X-ray and optical properties", A&A, **423**, 449 (2004).

Pratt, G. W. & Arnaud, M., "The mass profile of A1413 observed with XMM-Newton : Implications for the M-T relation", A&A, **394**, 375 (2002).

Pratt, G. W., Arnaud, M. & Pointecouteau, E., "Structure and scaling of the entropy in nearby galaxy clusters", A&A, **446**, 429 (2006).

Pratt, G. W., Böhringer, H., Croston, J. H., Arnaud, M., Borgani, S., Finoguenov, A. & Temple, R. F., "Temperature profiles of a representative sample of nearby X-ray galaxy clusters", A&A, **461**, 71 (2007).

Predehl, P., Hasinger, G., Böhringer, H., Briel, U., Brunner, H., Churazov, E., Freyberg, M., Friedrich, P., Kendziorra, E., Lutz, D. et al., "eROSITA", dans *Space Telescopes and Instrumentation II : Ultraviolet to Gamma Ray. Edited by Turner, Martin J. L. ; Hasinger, Günther. Proceedings of the SPIE, Volume 6266, pp. 62660P (2006).*, tome 6266 de *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference* (2006).

Press, W. H. & Schechter, P., "Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation", ApJ, **187**, 425 (1974).

Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P., *Numerical recipes in C. The art of scientific computing*, Cambridge : University Press, 1992, 2nd ed. (1992).

Rasia, E., Ettori, S., Moscardini, L., Mazzotta, P., Borgani, S., Dolag, K., Tormen, G., Cheng, L. M. & Diaferio, A., "Systematics in the X-ray cluster mass estimators", MNRAS, **369**, 2013 (2006).

Read, A. M. & Ponman, T. J., "The XMM-Newton EPIC background : Production of background maps and event files", A&A, **409**, 395 (2003).

Refregier, A. & Bacon, D., "Shapelets - II. A method for weak lensing measurements", MNRAS, **338**, 48 (2003).

Reiprich, T. H. & Böhringer, H., "The Mass Function of an X-Ray Flux-limited Sample of Galaxy Clusters", ApJ, **567**, 716 (2002).

Roberts, M. S., "The Rotation Curves of Galaxies", Comments on Astrophysics, **6**, 105 (1976).

Rodriguez, J., Tomsick, J. A., Foschini, L., Walter, R., Goldwurm, A., Corbel, S. & Kaaret, P., "An XMM-Newton observation of IGR J16320-4751 = AX J1631.9-4752", A&A, **407**, L41 (2003).

Rosati, P., della Ceca, R., Burg, R., Norman, C. & Giacconi, R., "A first determination of the surface density of galaxy clusters at very low x-ray fluxes", ApJ, **445**, L11 (1995).

Rosati, P., della Ceca, R., Norman, C. & Giacconi, R., "The ROSAT Deep Cluster Survey : The X-Ray Luminosity Function Out to z=0.8", ApJ, **492**, L21+ (1998).

Rubin, V. C., Thonnard, N. & Ford, W. K., Jr., "Extended rotation curves of high-luminosity spiral galaxies. IV - Systematic dynamical properties, SA through SC", ApJ, **225**, L107 (1978).

Sanders, J. S., Fabian, A. C. & Dunn, R. J. H., "Non-thermal X-rays, a high-abundance ridge and fossil bubbles in the core of the Perseus cluster of galaxies", MNRAS, **360**, 133 (2005).

Scharf, C., "Optimal Chandra and XMM-Newton Bandpasses for Detecting Low-Temperature Groups and Clusters of Galaxies", ApJ, **572**, 157 (2002).

Scharf, C. A., Jones, L. R., Ebeling, H., Perlman, E., Malkan, M. & Wegner, G., "The Wide-Angle ROSAT Pointed X-Ray Survey of Galaxies, Groups, and Clusters. I. Method and First Results", ApJ, **477**, 79 (1997).

Schechter, P., "An analytic expression for the luminosity function for galaxies.", ApJ, **203**, 297 (1976).

Schuecker, P., Böhringer, H., Collins, C. A. & Guzzo, L., "The REFLEX galaxy cluster survey. VII. $\Omega_m$ and $\sigma_8$ from cluster abundance and large-scale clustering", A&A, **398**, 867 (2003).

Serlemitsos, P. J., Smith, B. W., Boldt, E. A., Holt, S. S. & Swank, J. H., "X-radiation from clusters of galaxies - Spectral evidence for a hot evolved gas", ApJ, **211**, L63 (1977).

Setti, G. & Woltjer, L., "Active Galactic Nuclei and the spectrum of the X-ray background", A&A, **224**, L21 (1989).

Sheth, R. K., Mo, H. J. & Tormen, G., "Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes", MNRAS, **323**, 1 (2001).

Sheth, R. K. & Tormen, G., "Large-scale bias and the peak background split", MNRAS, **308**, 119 (1999).

Smith, G. P., Kneib, J.-P., Smail, I., Mazzotta, P., Ebeling, H. & Czoske, O., "A Hubble Space Telescope lensing survey of X-ray luminous galaxy clusters - IV. Mass, structure and thermodynamics of cluster cores at z= 0.2", MNRAS, **359**, 417 (2005).

Spergel, D. N., Bean, R., Doré, O., Nolta, M. R., Bennett, C. L., Dunkley, J., Hinshaw, G., Jarosik, N., Komatsu, E., Page, L. et al., "Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Implications for Cosmology", ApJS, **170**, 377 (2007).

Stanek, R., Evrard, A. E., Böhringer, H., Schuecker, P. & Nord, B., "The X-Ray Luminosity-Mass Relation for Local Clusters of Galaxies", ApJ, **648**, 956 (2006).

Starck, J.-L., Murtagh, F. & Bijaoui, A., *Image processing and data analysis. The multiscale approach*, Image processing and data analysis. The multiscale approach, Publisher : Cambridge, UK : Cambridge University Press, 1998, ISBN : 0521590841 (1998).

Starck, J.-L. & Pierre, M., "Structure detection in low intensity X-ray images", A&AS, **128**, 397 (1998).

Stocke, J. T., Morris, S. L., Gioia, I. M., Maccacaro, T., Schild, R., Wolter, A., Fleming, T. A. & Henry, J. P., "The Einstein Observatory Extended Medium-Sensitivity Survey. II - The optical identifications", ApJS, **76**, 813 (1991).

Strüder, L., Briel, U., Dennerl, K., Hartmann, R., Kendziorra, E., Meidinger, N., Pfeffermann, E., Reppin, C., Aschenbach, B., Bornemann, W. et al., "The European Photon Imaging Camera on XMM-Newton : The pn-CCD camera", A&A, **365**, L18 (2001).

Sugiyama, N., "Cosmic Background Anisotropies in Cold Dark Matter Cosmology", ApJS, **100**, 281 (1995).

Sunyaev, R. A. & Zeldovich, I. B., "Microwave background radiation as a probe of the contemporary structure and history of the universe", ARA&A, **18**, 537 (1980).

Tegmark, M., "The Angular Power Spectrum of the Four-Year COBE Data", ApJ, **464**, L35+ (1996).

Turner, M. J. L., Abbey, A., Arnaud, M., Balasini, M., Barbera, M., Belsole, E., Bennie, P. J., Bernard, J. P., Bignami, G. F., Boer, M. et al., "The European Photon Imaging Camera on XMM-Newton : The MOS cameras : The MOS cameras", A&A, **365**, L27 (2001).

Valtchanov, I., Pierre, M. & Gastaud, R., "Comparison of source detection procedures for XMM-Newton images", A&A, **370**, 689 (2001).

Valtchanov, I., Pierre, M., Willis, J., Dos Santos, S., Jones, L., Andreon, S., Adami, C., Altieri, B., Bolzonella, M., Bremer, M. et al., "The XMM-LSS survey. First high redshift galaxy clusters : Relaxed and collapsing systems", A&A, **423**, 75 (2004).

Vikhlinin, A. & Forman, W., "Detection of the Angular Correlation of Faint X-Ray Sources", ApJ, **455**, L109+ (1995).

Vikhlinin, A., Kravtsov, A., Forman, W., Jones, C., Markevitch, M., Murray, S. S. & Van Speybroeck, L., "Chandra Sample of Nearby Relaxed Galaxy Clusters : Mass, Gas Fraction, and Mass-Temperature Relation", ApJ, **640**, 691 (2006).

Vikhlinin, A., Markevitch, M., Murray, S. S., Jones, C., Forman, W. & Van Speybroeck, L., "Chandra Temperature Profiles for a Sample of Nearby Relaxed Galaxy Clusters", ApJ, **628**, 655 (2005).

Vikhlinin, A., McNamara, B. R., Forman, W., Jones, C., Quintana, H. & Hornstrup, A., "A Catalog of 200 Galaxy Clusters Serendipitously Detected in the ROSAT PSPC Pointed Observations", ApJ, **502**, 558 (1998).

Vikhlinin, A., VanSpeybroeck, L., Markevitch, M., Forman, W. R. & Grego, L., "Evolution of the Cluster X-Ray Scaling Relations since $z > 0.4$", ApJ, **578**, L107 (2002).

Vogt, C. & Enßlin, T. A., "Measuring the cluster magnetic field power spectra from Faraday rotation maps of Abell 400, Abell 2634 and Hydra A", A&A, **412**, 373 (2003).

Voit, G. M., "Tracing cosmic evolution with clusters of galaxies", Reviews of Modern Physics, **77**, 207 (2005).

Warren, M. S., Abazajian, K., Holz, D. E. & Teodoro, L., "Precision Determination of the Mass Function of Dark Matter Halos", ApJ, **646**, 881 (2006).

White, S. D. M., Efstathiou, G. & Frenk, C. S., "The amplitude of mass fluctuations in the universe", MNRAS, **262**, 1023 (1993).

Willingale, R., Aschenbach, B., Griffiths, R. G., Sembay, S., Warwick, R. S., Becker, W., Abbey, A. F. & Bonnet-Bidaud, J.-M., "New light on the X-ray spectrum of the Crab Nebula", A&A, **365**, L212 (2001).

Willis, J. P., Pacaud, F., Valtchanov, I., Pierre, M., Ponman, T., Read, A., Andreon, S., Altieri, B., Quintana, H., Dos Santos, S. et al., "The XMM Large-Scale Structure survey : an initial sample of galaxy groups and clusters to a redshift $z < 0.6$", MNRAS, **363**, 675 (2005).

Worsley, M. A., Fabian, A. C., Barcons, X., Mateos, S., Hasinger, G. & Brunner, H., "The (un)resolved X-ray background in the Lockman Hole", MNRAS, **352**, L28 (2004).

Yang, Y., Mushotzky, R. F., Barger, A. J. & Cowie, L. L., "Spatial Correlation Function of the Chandra-selected Active Galactic Nuclei", ApJ, **645**, 68 (2006).

Yang, Y., Mushotzky, R. F., Barger, A. J., Cowie, L. L., Sanders, D. B. & Steffen, A. T., "Imaging Large-Scale Structure in the X-Ray Sky", ApJ, **585**, 85 (2003).

Yuan, F. & Narayan, R., "On the Nature of X-Ray-Bright, Optically Normal Galaxies", ApJ, **612**, 724 (2004).

Zwicky, F., "Die Rotverschiebung von extragalaktischen Nebeln", Helvetica Physica Acta, **6**, 110 (1933).

# Deuxième partie:

## Publications référées

# Chapitre 7

# Présentation méthodologique des principales publications

## 7.1 Validation du relevé et premiers échantillons :

### "The XMM-LSS survey : an initial sample of galaxy groups and clusters to a redshift z < 0.6"

#### Contexte/motivation

Cet article (conjointement avec Valtchanov et al. 2004) clos la période de validation du relevé. Après une importante controverse sur la faisabilité d'obtenir un large échantillon d'amas dans des pointés de seulement 10 ks, les deux articles présentent des échantillons préliminaires constitués à partir des premières données. Valtchanov et al. (2004) démontre la capacité de détecter des amas distants ($0.6 < z < 1.0$), tandis que le présent article met en évidence, à partir d'un échantillon plus proche, la possibilité d'étudier les propriétés physiques des amas détectés directement à partir du relevé.

#### Résumé

L'article présente des observations X et spectroscopiques de 12 groupes et amas de galaxies identifiés dans le cadre du XMM-LSS. Ces systèmes furent sélectionnés en tant que sources X étendues à partir d'une mosaic XMM de 3.5 $deg^2$ au delà d'une limite en flux de $8 \times 10^{-15}$ $erg.s^{-1}.cm^{-2}$dans la bande d'énergie [0.5-2] keV. Des images BVRI profondes, ainsi que de la spectroscopie multi-objet permettent de valider chaque source comme une concentration de galaxies appartenant à l'intervalle de reshift $0.29 < z < 0.56$. Nous déterminons les dispersions de vitesses sur la ligne de visée ainsi que les propriétés X de chaque structure, calculées à l'aide d'un modèle de brillance de surface à 2 dimensions, et d'un ajustement au données spectral d'XMM d'une unique composante de température. La distribution résultante des luminosités X, températures et dispersions de vitesse indique que le relevé XMM-LSS détecte des amas de faible masse et des groupes de galaxies jusqu'à des redshifts de $z < 0.6$. Les systèmes confirmés montrent des signes d'une évolution positive de la luminosité X à une temperature donnée comparés aux

échantillons de groupes et amas de plus bas redshift. Une appréciation plus complète de ces tendances ne sera cependant possible que sur la base de l'échantillon statistiquement complet de groupes et amas de galaxies attendu avec la poursuite du relevé XMM-LSS.

### Contribution personnelle

Ma tâche principale pour cet article consistait à mesurer la temperature du milieu intra-amas dans les sources de l'échantillon. Compte tenu du bas S/N des sources (comparativement aux données habituellement utilisées pour ce type d'étude), la faisabilité d'une telle tâche à d'abord été considérée par l'intermédiaire de simulations. Cela m'a conduit à identifier un biais important dans la mesure des températures à bas S/N avec le modèle de C-statistique utilisé par le logiciel `Xspec`, ainsi que les conditions nécessaires pour l'éviter. Une fois ce biais corrigé, j'ai pu démontrer que, pour nos sources, la mesure de température était statistiquement possible même pour un signal faible grace aux basses températures considérées dans le relevé. Ces températures furent ensuite utilisées pour estimer la luminosité bolométrique des sources et étudier la relation $L_X - T$ de l'échantillon.

Mon. Not. R. Astron. Soc. **363,** 675–691 (2005) doi:10.1111/j.1365-2966.2005.09473.x

# The *XMM* Large-Scale Structure survey: an initial sample of galaxy groups and clusters to a redshift $z < 0.6^{\star}$

J. P. Willis,[1]† F. Pacaud,[2] I. Valtchanov,[2,3] M. Pierre,[2] T. Ponman,[4] A. Read,[5] S. Andreon,[6] B. Altieri,[7] H. Quintana,[8] S. Dos Santos,[2] M. Birkinshaw,[9] M. Bremer,[9] P.-A. Duc,[2] G. Galaz,[8] E. Gosset,[10] L. Jones[4] and J. Surdej[10]

[1]*Department of Physics and Astronomy, University of Victoria, Elliot Building, 3800 Finnerty Road, Victoria, BC, V8P 1A1 Canada*
[2]*CEA/Saclay, Service d'Astrophysique, F-91191, Gif–sur–Yvette, France*
[3]*Astrophysics Group, Blackett Laboratory, Imperial College of Science Technology and Medicine, London SW7 2BW*
[4]*School of Physics and Astronomy, University of Birmingham, Edgebaston, Birmingham B15 2TT*
[5]*Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH*
[6]*INAF – Osservatorio Astronomico di Brera, Milan, Italy*
[7]*ESA, Villafranca del Castillo, Spain*
[8]*Departamento de Astronomía y Astrofísica, P.U. Católica de Chile, Santiago, Chile*
[9]*Department of Physics, University of Bristol, Tyndall Avenue, Bristol BS8 1TL*
[10]*Université de Liège, Allée du 6 Août, 17, B5C, 4000 Sart Tilman, Belgium*

Accepted 2005 July 28. Received 2005 July 28; in original form 2004 September 14

**ABSTRACT**

We present X-ray and optical spectroscopic observations of 12 galaxy groups and clusters identified within the *X-ray Multi-Mirror* (*XMM*) Large-Scale Structure (LSS) survey. Groups and clusters are selected as extended X-ray sources from a 3.5 deg$^2$ *XMM* image mosaic above a flux limit $8 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ in the [0.5–2] keV energy band. Deep *BVRI* images and multi-object spectroscopy confirm each source as a galaxy concentration located within the redshift interval $0.29 < z < 0.56$. We combine line-of-sight velocity dispersions with the X-ray properties of each structure computed from a two-dimensional surface brightness model and a single temperature fit to the *XMM* spectral data. The resulting distribution of X-ray luminosity, temperature and velocity dispersion indicate that the *XMM*–LSS survey is detecting low-mass clusters and galaxy groups to redshifts $z < 0.6$. Confirmed systems display little or no evidence for X-ray luminosity evolution at a given X-ray temperature compared to lower-redshift X-ray group and cluster samples. A more complete understanding of these trends will be possible with the compilation of a statistically complete sample of galaxy groups and clusters anticipated within the continuing *XMM*–LSS survey.

**Key words:** surveys – large-scale structure of Universe – X-rays: galaxies: clusters.

## 1 INTRODUCTION

Surveys of distant galaxy clusters map the distribution in the Universe of large amplitude density fluctuations, and so constrain key cosmological parameters and permit secondary studies to determine how X-ray gas and galaxy evolution proceeds as a function of environment. Wide-area X-ray surveys are well placed to compile statistically well-defined samples of distant galaxy clusters because (i) both source confusion and the X-ray background are low compared to optical searches, (ii) computation of the selection function and volume sampled is straightforward, and (iii) the selection of extended X-ray emitting sources is sensitive to the signature of hot gas contained within massive, gravitationally bound structures.

A number of systematic X-ray studies have extended both the maximum redshift (i.e. the most luminous galaxy clusters) and the minimum luminosity (i.e. the least massive structures) to which X-ray clusters can be identified. A comprehensive review is provided by Rosati, Borgani & Norman (2002). The principal aim of such surveys for distant, X-ray emitting clusters is to determine their space density evolution as a function of redshift and to constrain the combination of the rms mass density fluctuations on $8\,h^{-1}$ Mpc scales, $\sigma_8$, and the overall matter density of the Universe expressed as a fraction of the closure density, $\Omega_{\rm M}$ (e.g. Borgani et al. 2001;

⋆Based upon observations performed at Paranal (70.A-0283), Las Campanas and CTIO observatories and on observations obtained with *XMM–Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA.
†E-mail: jwillis@uvic.ca

Schuecker et al. 2001; Allen et al. 2003). In addition to the study of global cosmological parameters, galaxy clusters provide examples of dense cosmic environments in which it is possible to study the evolution of the hot, X-ray emitting gas (e.g. Ettori et al. 2004; Lumb et al. 2004) and to determine the nature of the cluster galaxy populations and the physical processes underlying observed trends in galaxy evolution (e.g. Yee, Ellingson & Carlberg 1996; Dressler et al. 1999; Andreon et al. 2004a).

Extending our current knowledge of low-luminosity (i.e. low-mass) X-ray clusters represents an important challenge for the present generation of X-ray surveys performed with the *X-ray Multi-Mirror* (*XMM*) and *Chandra* facilities. The local ($z < 0.25$) X-ray luminosity function (XLF) for galaxy clusters is currently determined to X-ray luminosities $L_X \sim 10^{42}$ erg s$^{-1}$ in the [0.5, 2] keV energy band (Henry et al. 1992; Rosati et al. 1998; Ledlow et al. 1999; Böhringer et al. 2002). However, our understanding of such systems at redshifts $0.25 < z < 0.8$ is largely restricted to luminosities $L_X \gtrsim 10^{43.5}$ erg s$^{-1}$ (Henry et al. 1992; Burke et al. 1997; Rosati et al. 1998; Vikhlinin et al. 1998; Ebeling, Edge & Henry 2001).

Low-luminosity ($10^{42} < L_X/\mathrm{erg\,s^{-1}} < 10^{43}$) X-ray clusters correspond to low-mass clusters and larger galaxy groups that form a link between poorly defined 'field' environments and X-ray luminous/optically rich clusters. If, as anticipated, X-ray clusters occupying this luminosity range display X-ray temperatures $T < 4$ keV, they are more likely to display the effects of non-gravitational energy input into the intracluster medium (ICM) than hotter, more massive clusters (Ponman, Cannon & Navarro 1999). Deviations of X-ray scaling relations from simple, self-similar expectations have been studied for structures displaying a relatively wide range of mass/temperature scales at $z < 0.2$ (e.g. Sanderson et al. 2003). However, the study of X-ray emitting structures – selected over an extended temperature range – at $z > 0.2$ will provide an important insight into the evolution of their X-ray emitting gas. Although detailed X-ray studies of galaxy clusters at $0.2 < z < 0.6$ are in progress (e.g. Lumb et al. 2004), few examples of X-ray emitting structures displaying temperatures $T < 3$ keV are currently known at such redshifts. Clearly, the compilation of a sample of X-ray emitting galaxy groups and clusters to $z < 0.6$ will greatly increase the range of X-ray gas temperatures over which evolutionary effects can be studied.

Low-mass clusters and groups are predicted to be sites of continuing galaxy evolution at $z < 1$ (e.g. Baugh, Cole & Frenk 1996; Kauffmann 1996). When identifying such systems, it is important to note that extended X-ray emission arises from gravitationally bound structures. This is an important difference when X-ray selected cluster samples are compared to optical/near-infrared selected cluster samples – whose dynamical state can only be assessed with additional velocity data. In addition, the X-ray properties of galaxy structures (luminosity and temperature) constrain the gravitational mass of the emitting structure. Extending the currently known sample of galaxy groups and low-mass clusters at $z > 0.2$ via X-ray observations will provide an important group/cluster sample with consistent mass ordering. A mass-ordered cluster sample will permit several detailed studies of galaxy evolution at look-back times >3–4 Gyr to be undertaken; e.g. morphological segregation and merger-related effects (Heldson & Ponman 2003), Butcher–Oemler effects (Andreon & Ettori 1999) and the evolution of colour and luminosity functions (Andreon et al. 2004a). To date, detailed galaxy evolution studies of moderately distant $z > 0.2$, X-ray selected clusters have been performed typically for only the most X-ray bright (i.e. massive) systems, e.g. $L_X$([0.3–3.5] keV) $> 4 \times 10^{44}$ erg s$^{-1}$ (Yee et al. 1996; Dressler et al. 1999). Clearly, an improved sample of systems covering an extended mass interval will permit a detailed investigation of galaxy evolution effects as a function of changing environment.

The *XMM* Large-Scale Structure (LSS) survey (Pierre et al. 2004) is a wide-area X-ray survey with the *XMM* facility with the primary aim to extend detailed studies of the X-ray cluster correlation function, currently determined at $z < 0.2$ as part of the REFLEX survey (Schuecker et al. 2001), to a redshift of unity. However, the *XMM*–LSS survey features a number of secondary aims including determination of the cosmological mass function to faint X-ray luminosities, the evolution of cluster galaxy populations and the evolution of the X-ray emitting gas in clusters selected over a range of mass scales. The nominal point-source flux limit of the *XMM*–LSS survey is $5 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ in the [0.5–2] keV energy band. Refregier, Valtchanov & Pierre (2002) demonstrate that (assuming a reasonable distribution of cluster surface brightness profiles) the approximate flux limit for typical extended sources is $8 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$, which corresponds to an X-ray luminosity $L_X = 1.2 \times 10^{43}$ erg s$^{-1}$ for a cluster located at a redshift $z = 0.6$ and $L_X = 4.2 \times 10^{43}$ erg s$^{-1}$ for a cluster located at a redshift $z = 1$ within the adopted cosmological model (see below).

The above aims are predicated upon the compilation of a large, well-defined cluster catalogue. In this paper we describe the first results of the *XMM*–LSS survey at identifying X-ray emitting clusters at $z < 0.6$. The first clusters identified at $z > 0.6$ are presented in Valtchanov et al. (2004). The current paper is organized as follows. In Section 2 we summarize the X-ray and optical imaging data and the methods employed to select candidate clusters. In Section 3 we describe spectroscopic observations and reductions performed for a subset of candidate clusters. In Section 4 we present the determination of cluster spectroscopic properties (redshift and line-of-sight velocity dispersion). In Section 5 we present the determination of confirmed cluster X-ray properties (surface brightness and temperature fitting). In Section 6 we present the current conclusions for the properties of the initial $z < 0.6$ sample. Throughout this paper, we assume a Friedmann–Robertson–Walker cosmological model, characterized by the present-day parameters $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$. Where used, $h$ is defined as $h = H_0/(100$ km s$^{-1}$ Mpc$^{-1})$.

## 2 IDENTIFYING X-RAY CLUSTER CANDIDATES

### 2.1 X-ray data reduction and source detection

Galaxy cluster targets presented in this paper were selected from a mosaic of overlapping *XMM* pointings covering a total area of approximately 3.5 deg$^2$. This data set represents all *XMM*–LSS pointings received by 2002 August and includes 15 A0-1 10-ks exposures and 15 Guaranteed Time (GT) 20-ks exposures obtained as part of the *XMM* Medium Deep Survey (MDS).

*XMM* observations were reduced employing the *XMM* Science Analysis System (SAS) tasks EMCHAIN and EPCHAIN for the multi-object spectroscopy (MOS) and pn detectors, respectively. High background periods induced by soft-proton flaring were excluded from the event lists and raw photon images as a function of energy band were created. The raw images for each detector were processed employing an iterative wavelet technique and a Poissonian noise model with a threshold of $10^{-3}$ (equivalent to $3\sigma$ for the Gaussian case) applied to select the significant wavelet coefficients (Starck & Pierre 1998). Each wavelet filtered image was exposure corrected

**Table 1.** Detection thresholds as a function of photometric passband for objects detected in the optical images.

| Filter | Detection threshold (50 per cent AB magnitude completeness limit for stellar sources) |
|---|---|
| *B* | 26.5 |
| *V* | 26.0 |
| *R* | 26.0 |
| *I* | 25.4 |

and an image mask (including deviant pixels, detector gaps and non-exposed detector regions) was created.

Source detection was performed on the wavelet filtered X-ray images employing the SEXTRACTOR package (Bertin & Arnouts 1996). The discrimination between extended (cluster) and point-like sources (mostly active galactic nuclei) was achieved employing a two-constraint test based on the half-energy radius and the SEXTRACTOR stellarity index of the sources. The applied procedure is the optimum method given the *XMM* point spread function (PSF) and the Poisson nature of the signal (Valtchanov, Pierre & Gastaud 2001). The measurement of extended source properties was performed on the EPIC/pn images as they provide the greatest sensitivity. The EPIC/MOS images were used to discard possible artefacts resulting from edge effects associated with the pn CCDs.

### 2.2 Optical imaging

The selection of potential galaxy members within each candidate cluster was performed employing moderately deep *BVRI* images from the CFH12k camera on the Canada–France–Hawaii Telescope (CFHT) obtained as part of the VIRMOS deep imaging survey (McCracken et al. 2003). Observations were processed employing the TERAPIX[1] data reduction pipeline to produce an astrometric and photometric image data set. Object catalogues were produced using SEXTRACTOR. Catalogue detection thresholds as a function of photometric band are displayed in Table 1.

### 2.3 Selecting candidate clusters and member galaxies

The identification of galaxy clusters over an extended redshift interval in X-ray images is limited by the ability of the *XMM* facility to identify extended cluster X-ray emission in a 10-ks exposure. The half-energy width (HEW) of an on-axis point source is approximately 15 arcsec at 1.5 keV. However, the HEW displays marked local variations resulting from off-axis angle, vignetting and detector gaps. The on-axis HEW corresponds to a projected transverse distance of 120 kpc at a redshift $z = 1$ within the assumed cosmological model. Although this HEW is sufficient to resolve the extended emission from massive galaxy clusters to $z \gtrsim 1$, the effect of low central surface brightness, leading to a truncated detectable cluster extension, can lead to a cluster being erroneously identified as an unresolved object. Therefore, both distant clusters and intrinsically compact clusters at all redshifts may potentially appear as only marginally resolved or unresolved sources in *XMM*–LSS X-ray mosaics. A quantitative assessment of the cluster X-ray selection function will form the subject of a future paper (Pacaud et al., in preparation).

The *XMM*–LSS survey incorporates a number of different solutions to the problem of cluster identification, e.g. correlation of extended X-ray sources with optical galaxy structures (this paper and Valtchanov et al. 2004), investigation of the X-ray properties of optically identified structures (and vice versa; Andreon et al. 2004b) and the investigation of extended X-ray sources lacking optical counterparts together with unresolved X-ray sources associated with faint optical structures (Andreon et al. 2004b).

Analysis of the first 3.5 deg$^2$ of the *XMM*–LSS survey led to the identification of 55 extended sources with fluxes greater than $8 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$, which are extended according to the criteria detailed in Section 2.1. The optical imaging data corresponding to a $7 \times 7$ arcmin$^2$ field[2] centred upon the location of each extended X-ray source was analysed for the presence of a significant galaxy structure showing a well-defined, red colour sequence (Kodama et al. 1998; Stanford, Eisenhardt & Dickinson 1998). Candidate cluster members were selected by inspection of the available *BVRI* photometry of objects identified within the field of each extended X-ray source. Colour–magnitude thresholds were applied interactively in order to enhance galaxy structures when viewed in a *VRI* pseudo-colour image of each field with X-ray contours superposed from the wavelet-filtered X-ray image. One exception to this procedure was cluster candidate XLSS J022722.3-032141 (see Table 2); the optical data for this candidate cluster consisted of the *I*-band Very Large Telescope (VLT) Focal Reducing Spectrograph (FORS2) pre-image obtained to define slit locations. Candidate cluster members were selected to include all galaxies up to 1.5 mag fainter than the bright ($I = 17.1$), central galaxy associated with the extended X-ray source.

The sample of candidate cluster members generated by the above procedures was used to design spectroscopic masks for each candidate cluster field. Two multi-slit masks were created for each candidate cluster with brighter galaxies given higher priority in the slit assignment procedure. Unused regions of each multi-slit mask were employed to sample the population of unresolved X-ray sources with bright ($R < 23$) optical counterparts. Further discussion of this additional sample will appear elsewhere.

## 3 SPECTROSCOPIC OBSERVATIONS

The candidate cluster sample was observed by the Las Campanas Observatory Baade telescope with the Low Dispersion Survey Spectrograph (LDSS2) during 2002 October 4–5 and the European Southern Observatory VLT with FORS2 during 2002 October 9–12. Each instrument (LDSS2 and FORS2) is a focal reducing spectrograph with both an imaging and a MOS capability. In each case, MOS observations are performed using slit masks mounted in the instrument focal plane. Details of which cluster candidate was observed with which telescope plus instrument configuration are provided in Table 2. The effective wavelength interval, pixel sampling and spectral resolution generated by each instrument combination are indicated in Table 3.

Spectroscopic observations were reduced employing standard data reduction procedures within IRAF:[3] a zero level, flat-field and cosmic ray correction was applied to all MOS observations prior

[1] See http://terapix.iap.fr.

[2] A practical limit determined by the field of view of the multi-object spectroscopic facilities employed to observe cluster candidates (see Section 3). The field size is large compared to the extent of the X-ray emission and obviates any requirement to adjust the field centre to maximize the number of candidate cluster members.

[3] IRAF is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.

**Table 2.** Observing log of candidate groups and clusters.

| ID[a] | Cluster | RA[b] | Dec. | Instrument | Grism + filter | # of masks | Total exposure time per mask (seconds) |
|---|---|---|---|---|---|---|---|
| 006 | XLSSUJ022145.2-034614 | 02:21:45.22 | −03:46:14.1 | FORS2 | 300V + GG435 | 2 | 2700 + 2400 |
| 007 | XLSSUJ022406.0-035511 | 02:24:05.95 | −03:55:11.4 | FORS2 | 600RI + GG435 | 2 | 2400 + 2400 |
| 008 | XLSSUJ022520.7-034800 | 02:25:20.71 | −03:48:00.0 | FORS2 | 600RI + GG435 | 2 | 1200 + 1200 |
| 009 | XLSSUJ022644.2-034042 | 02:26:44.21 | −03:40:41.8 | FORS2 | 300V + GG435 | 2 | 1800 + 1800 |
| 010 | XLSSUJ022722.2-032137 | 02:27:22.16 | −03:21:37.0 | FORS2 | 600RI + GG435 | 1 | 600 |
| 012 | XLSSUJ022827.5-042554 | 02:28:27.47 | −04:25:54.3 | LDSS2 | medium-red | 2 | 1800 + 1800 |
| 013 | XLSSUJ022726.0-043213 | 02:27:25.98 | −04:32:13.1 | LDSS2 | medium-red | 2 | 900 + 900 |
| 014 | XLSSUJ022633.9-040348 | 02:26:33.87 | −04:03:48.0 | LDSS2 | medium-red | 2 | 1800 + 1800 |
| 016 | XLSSUJ022829.0-045932 | 02:28:29.03 | −04:59:32.2 | LDSS2 | medium-red | 2 | 600 + 900 |
| 017 | XLSSUJ022628.2-045948 | 02:26:28.19 | −04:59:48.1 | LDSS2 | medium-red | 2 | 1800 + 1800 |
| 018 | XLSSUJ022401.5-050525 | 02:24:01.46 | −05:05:24.8 | LDSS2 | medium-red | 2 | 2400 + 2400 |
| 020 | XLSSUJ022627.0-050008 | 02:26:27.08 | −05:00:08.4 | LDSS2 | medium-red | 2 | 1800 + 1800 |

[a]All clusters are referred to via the reference XLSSC plus the identification number, e.g. XLSSC 006, etc. Clusters 001-005 correspond to redshift $z > 0.6$ clusters presented in Valtchanov et al. (2004).
[b]Positions are J2000.0.

**Table 3.** Instrumental characteristics for each spectrograph configuration employed during the observations.

| Instrument | Grism + filter | Wavelength interval (Å) | Pixel sampling (Å pix$^{-1}$) | Spectral resolution[a] (Å) |
|---|---|---|---|---|
| FORS2 | 300V + GG435 | 4000–9000 | 3.2 | 14 |
| FORS2 | 600RI + GG435 | 5000–8500 | 1.6 | 7 |
| LDSS2 | medium-red | 4000–9000 | 5.1 | 14 |

[a]Estimated for each spectrograph via the mean FWHM of the HeI5876 arc emission line. All spectral observations were performed with a slit width of approximately 1.4 arcsec.

to the identification, sky subtraction and extraction of individual spectral traces employing the APEXTRACT package. The dispersion solution for each extracted spectrum was determined employing HeNeAr lamp exposures and all data spectra were resampled to a linear wavelength scale. A single spectrophotometric standard star from the atlas of Hamuy, Walker & Suntzeff (1992) and Hamuy, Suntzeff & Heathcote (1994) was observed during each night and was employed to correct for the relative instrumental efficiency as a function of wavelength. Removal of the relative instrumental efficiency as a function of wavelength does not affect the later determination of galaxy redshifts via cross-correlation analysis. However, it does permit spectra to be displayed on a relative spectral flux scale that aids the visual assessment of low-quality spectra.

### 3.1 Spectral classification and redshift determination

In order to confirm the redshift of each candidate cluster, the spectroscopic sample generated for each field was constructed to maximize the number of potential cluster members according to the photometric criteria described in Section 2.3. Although these criteria were constructed in order to identify the characteristic colour signature generated by an overdensity of early-type galaxies at a particular redshift, it is probable that the spectral sample generated for each candidate cluster is contaminated by the presence of galaxies within the target field that are (gravitationally) unassociated with the cluster and with stars misidentified as galaxies. In order to address this issue and to classify each candidate cluster member, all reduced spectra were inspected visually to identify contaminating stars and to provide an initial estimate of galaxy redshifts based upon the identification of prominent features. Individual spectra were then cross-correlated with a representative early-type galaxy template (Kinney et al. 1996) employing the IRAF routine XCSAO (Tonry & Davis 1979). The cross-correlation procedure was performed interactively in order to improve the identification of a reliable cross-correlation peak. Spectral regions corresponding to the observed locations of prominent night sky emission features and regions of strong atmospheric absorption were masked within the cross-correlation analysis. Computed redshift values have not been corrected to a heliocentric velocity scale.

Errors in the cross-correlation velocity returned by XCSAO are computed based upon the fitted peak height and the antisymmetric noise component associated with the identified cross-correlation peak (Tonry & Davis 1979; Heavens 1993). The typical median velocity error computed for spectra observed with each spectrograph combination described in Table 3 is 75 km s$^{-1}$ for FORS2+600RI and 150 km s$^{-1}$ for FORS2+300V and LDSS2/medium-red. The random error associated with uncertainties in the dispersion solution applied to each spectrum was characterized via determination of the error in the wavelength location of prominent emission features in the night sky spectrum associated with each data spectrum when compared to their reference values. The distributions of wavelength residuals are considered in velocity space and the ROSTAT statistics package (Beers, Flynn & Gebhardt 1990; see Section 4) is employed to calculate the bi-weight mean and scale for each cluster field. Typical values (for all spectrograph settings) of the mean wavelength shift and dispersion computed via this method are $<\pm 100$ and $<50$ km s$^{-1}$, respectively. Radial velocities (from which the redshift of each galaxy is determined) are corrected for the mean velocity residual in each field and we further assume that the distribution of errors in the cross-correlation velocities and in the dispersion solutions are Gaussian.

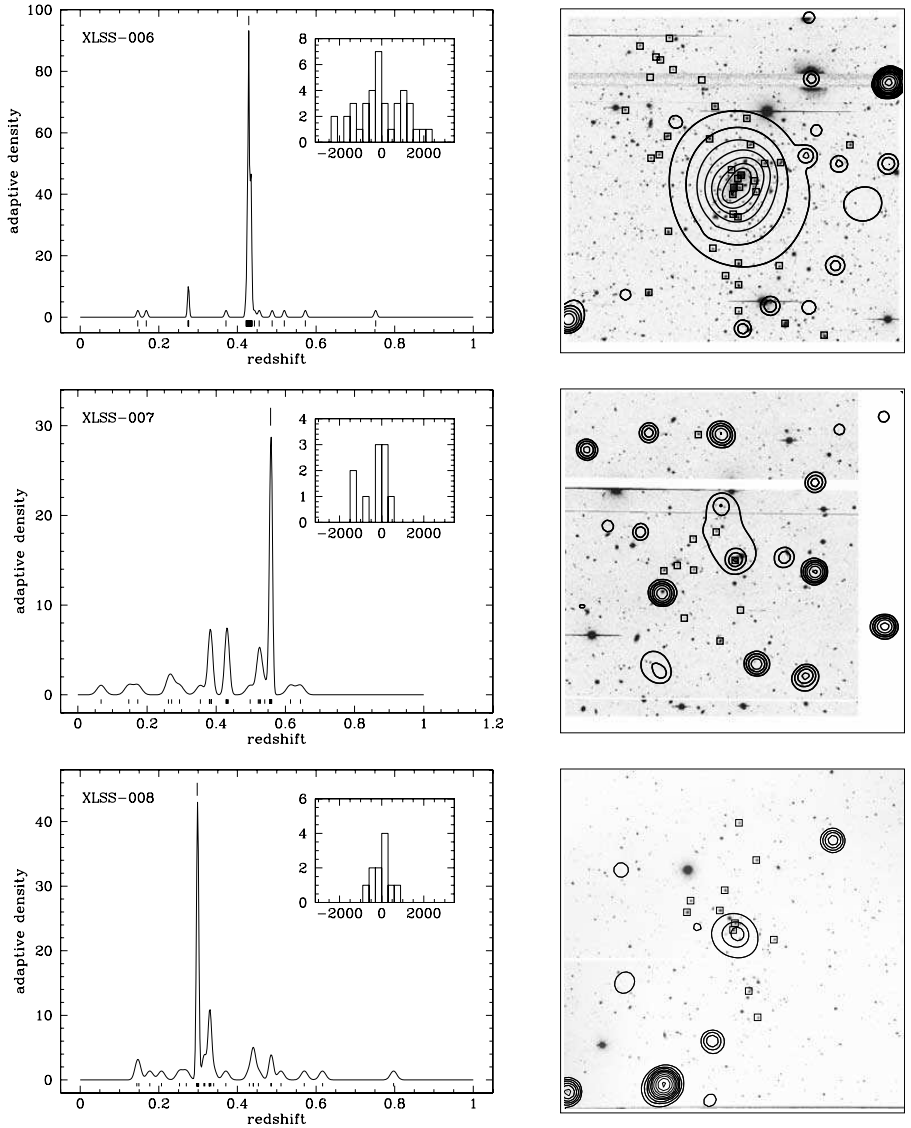


**Figure 1.** Optical images, X-ray contours and spectroscopic information for each cluster presented in Table 4. Left-hand panel: adaptively smoothed density profile versus redshift. Individual redshifts are marked by short vertical lines below the density curve. A short vertical line above the density curve marks the redshift location of the cluster. The inset histogram displays the rest-frame velocity distribution of confirmed cluster members. Right-hand panel: a 7 × 7 arcmin$^2$ *R*-band image centred on the extended X-ray source. Wavelet filtered X-ray contours are overplotted and squares indicate the position of confirmed cluster members. The X-ray contours typically run from two times the background level in each frame to 5 photons pixel$^{-1}$ with 10 logarithmic levels. Note that the optical images form a heterogeneous data set and are presented to indicate the visual appearance of each cluster. All optical images are orientated with north up and east left. Where more than one cluster is detected in the same field, a black arrow indicates the location of cluster being considered.

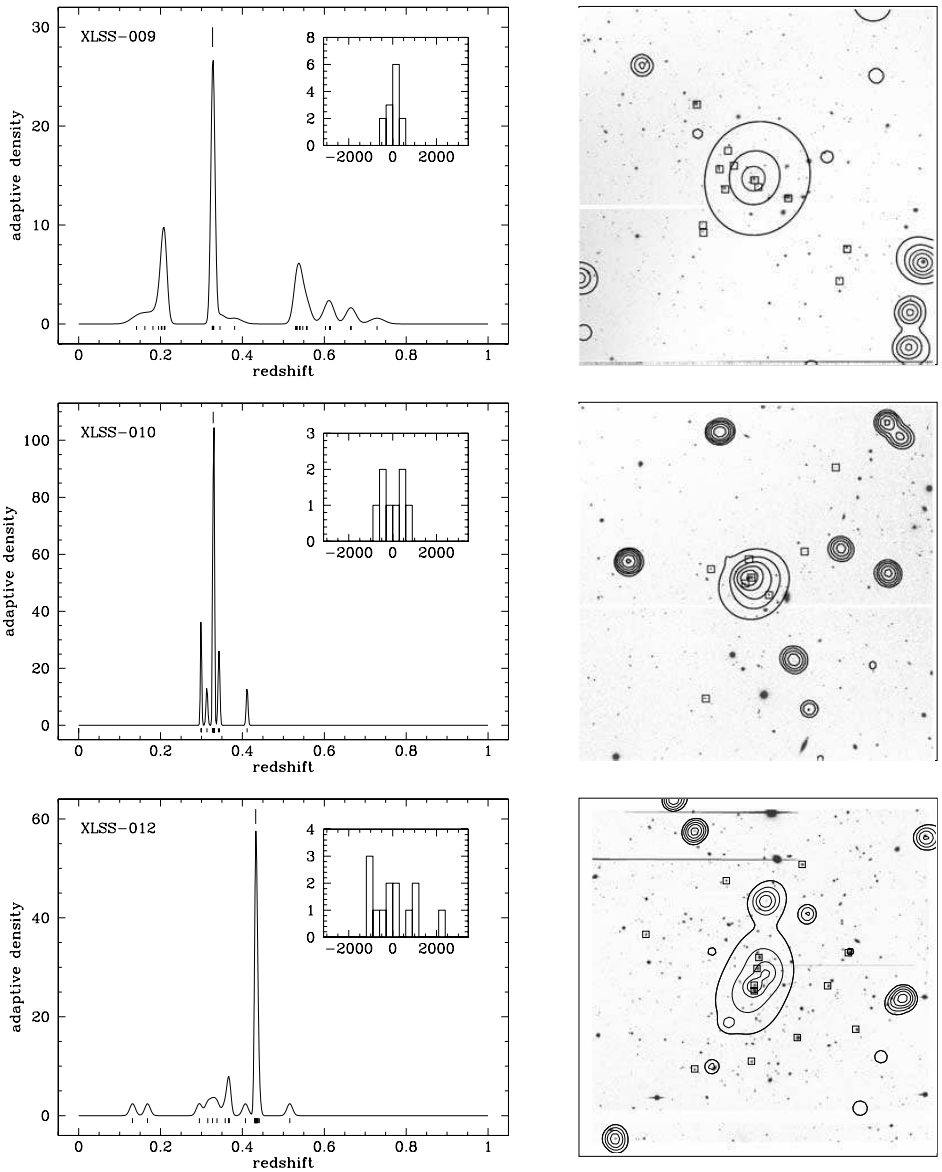


**Figure 1 –** *continued*

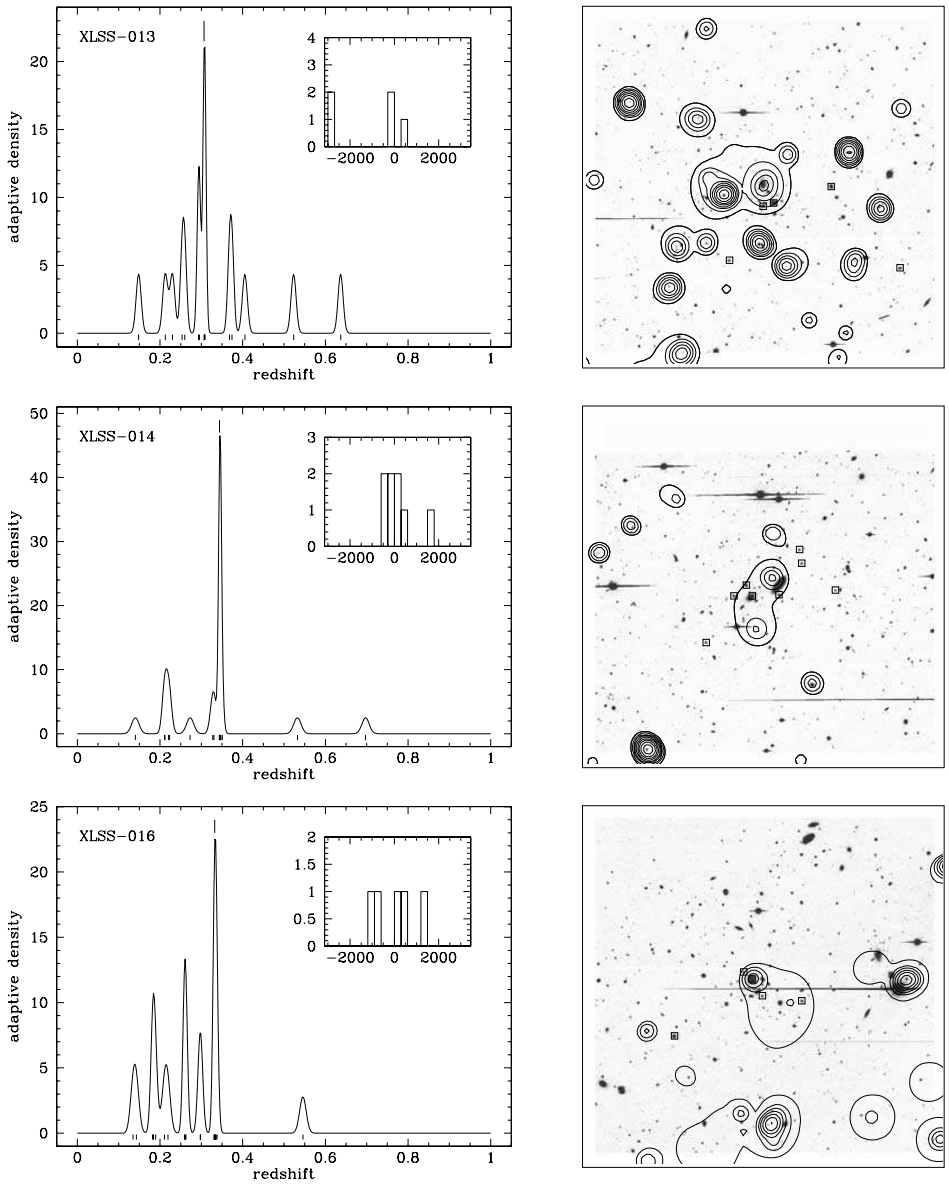


**Figure 1 –** *continued*

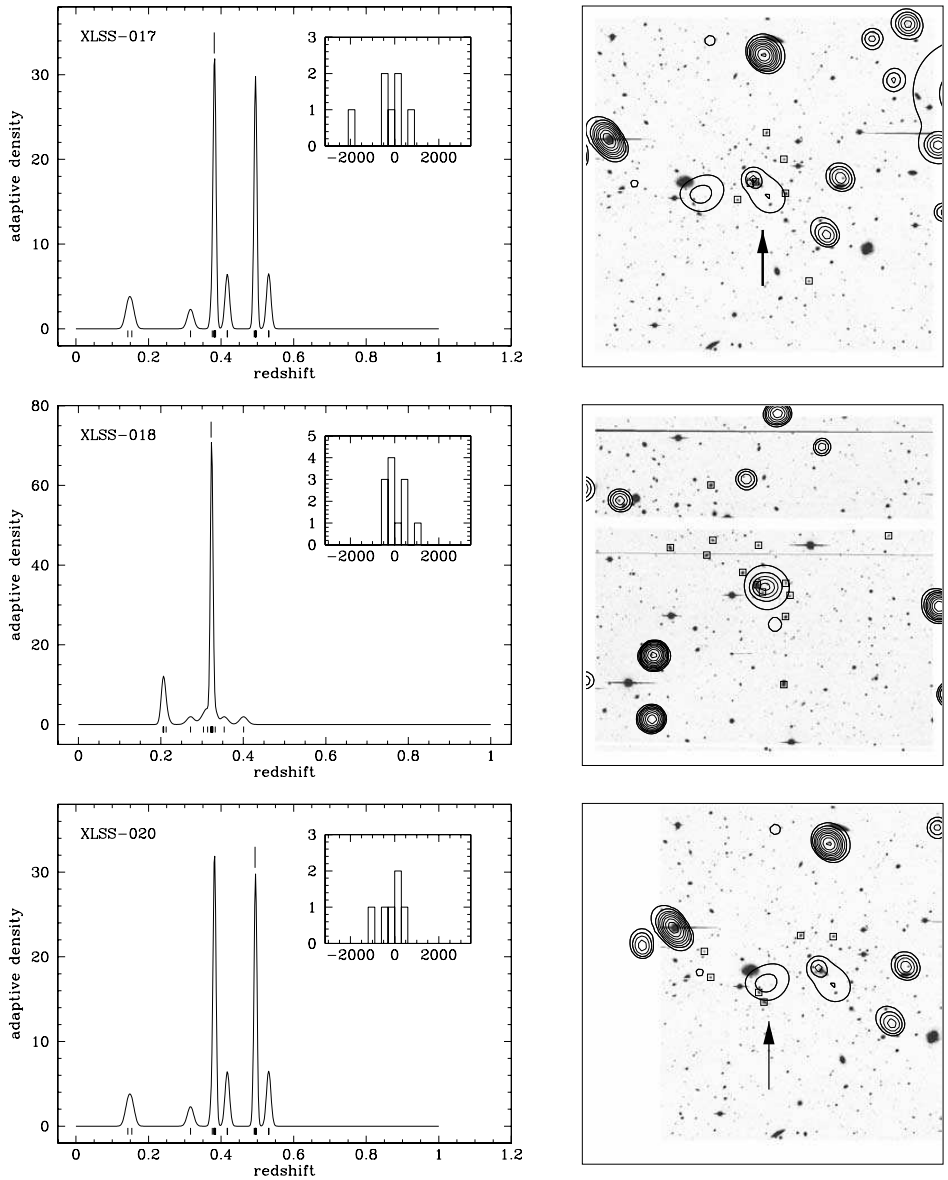


**Figure 1 –** *continued*

Therefore, we calculate a total uncertainty in the corrected radial velocity by combining these two sources of error in quadrature.

## 4 DETERMINATION OF CLUSTER SPECTROSCOPIC PROPERTIES

The nature of each candidate cluster contained within the spectroscopic sample is assessed employing the available X-ray and optical images and the spectroscopic information accumulated for each field. The field of each candidate cluster is inspected visually employing a composite image containing the CFH12k *R*-band grey-scale image, X-ray contours derived from the wavelet-filtered *XMM* mosaic and the available redshifts of all galaxies contained within the field (see Fig. 1). The redshift distribution generated by all spectroscopic redshifts obtained in each candidate cluster field is also displayed. To illustrate redshift space overdensities, the redshift density computed by applying an adaptive kernel (Silverman 1986) to the redshift data is also shown. This process provides an initial estimate of the cluster redshift via the identification of three-dimensional (position and redshift) structures associated with the extended X-ray emission. This redshift estimate is then employed to select the corresponding peak in the redshift histogram of the field. In the case of clusters XLSSC 017 and 020, the redshift of each extended X-ray source was determined by determining the spatial barycentre of each redshift peak displayed in Fig. 1 and assigning the redshift grouping closest to each X-ray source as the cluster redshift.

The typically small number (<20) of objects observed spectroscopically in each cluster field limits the usefulness of any assessment of spectroscopic completeness and redshift confirmation frequency. However, the spectroscopic redshift reported for each cluster is in excellent agreement ($\Delta z < 0.02$) with the redshift determined independently from the location of the red envelope of the corresponding cluster colour–magnitude relation in *R*–*z* colour space (Andreon et al. 2004a).

The sample of cluster members is selected in radial velocity space employing an iterative method similar to that of Lubin, Oke & Postman (2002): the initial cluster sample is selected to lie within the redshift interval $\Delta z = \pm 0.06$ of the estimated cluster redshift. Radial velocities relative to the cluster centre are calculated within the cluster rest frame, i.e. $\Delta v = c\,(z - \bar{z})/(1 + \bar{z})$ where $\bar{z}$ is the median redshift within the specified interval. The bi-weight mean and scale of the radial velocity distribution within this interval is computed using ROSTAT and galaxies that display a velocity difference relative to the central location of greater than 3500 km s$^{-1}$, or three scale measures, are rejected and the statistical measures recalculated. This procedure is repeated until no further galaxies are rejected. Errors in the bi-weight mean and scale are estimated employing a bootstrap or jackknife calculation with 10 000 resamplings for clusters with greater than or less than 10 confirmed members, respectively. Estimates of the bi-weighted mean radial velocity and line-of-sight velocity dispersion of each cluster, and the associated uncertainties, are corrected for biases arising from measurement errors employing the prescription of Danese, de Zotti & di Tullio (1980).

Although galaxy redshift and positional information is compared to the location of the X-ray source to define the initial cluster redshift input to the velocity search algorithm, no additional spatial filtering of potential galaxy cluster members is performed. However, the projected transverse distances sampled by the detector fields of view employed to perform the spectroscopy vary between 1.5 and $2\,h^{-1}$ Mpc for clusters located over the redshift interval $0.3 < z < 0.6$. Studies of velocity dispersion gradients in both local (Girardi et al. 1996) and distant (Borgani et al. 1999) X-ray clusters – albeit hotter/more luminous systems than those presented in the current paper – indicate that integrated velocity dispersion profiles typically converge within radii $r < 1$–$2\,h^{-1}$ Mpc of the X-ray cluster centre. Upon initial inspection, the cluster regions sampled by the projected field of view sampled by each telescope plus spectrograph combination would appear to be well matched to the convergent velocity profile of typical hot/luminous clusters. However, inspection of Fig. 1 indicates that cluster galaxies are typically confirmed within the central regions in each field – a strategy required by the necessity to confirm the redshift of galaxy structures near the X-ray source. Although the *XMM*–LSS clusters presented in this paper are typically cooler (i.e. less massive) than those presented by Girardi et al. (1996) and Borgani et al. (1999), and may reasonably be assumed to be intrinsically less extensive, an unknown and potentially significant uncertainty is associated with the assumption that the computed velocity dispersion figures represent the convergent velocity dispersion for each cluster. The resulting spectroscopic properties of all cluster candidates listed in Table 2 are given in Table 4.

**Table 4.** Spectroscopic properties of all redshift $z < 0.6$ groups and clusters.

| Cluster | Redshift[a] | # of members | $\sigma_v$ [b] (km s$^{-1}$) |
|---|---|---|---|
| XLSSC 006 | 0.429 | 39 | $821^{+92}_{-74}$ |
| XLSSC 007 | 0.558 | 10 | $323^{+178}_{-191}$ |
| XLSSC 008 | 0.298 | 11 | $351^{+98}_{-35}$ |
| XLSSC 009 | 0.327 | 13 | $232^{+60}_{-31}$ |
| XLSSC 010 | 0.329 | 8 | $420 \pm 72$ |
| XLSSC 012 | 0.433 | 13 | $694^{+204}_{-91}$ |
| XLSSC 013[c] | 0.307 | 5 | N/A |
| XLSSC 014 | 0.344 | 8 | $416 \pm 246$ |
| XLSSC 016 | 0.332 | 5 | $703 \pm 266$ |
| XLSSC 017 | 0.382 | 7 | $571 \pm 282$ |
| XLSSC 018 | 0.322 | 12 | $342^{+104}_{-35}$ |
| XLSSC 020 | 0.494 | 6 | $265^{+240}_{-146}$ |

Notes. [a]The uncertainty associated with cluster redshifts is less than $\Delta z = 0.001$ in all cases. [b]Velocity dispersion uncertainties are quoted at the 68 per cent confidence level. [c]The available data for cluster 013 do not generate a well-defined velocity dispersion.

## 5 DETERMINATION OF GROUP AND CLUSTER X-RAY PROPERTIES

In order to determine the nature of the spectroscopically confirmed clusters presented in this paper, additional analyses of the available *XMM* data were performed to characterize the spatial and spectral properties of the X-ray emitting gas. When combined with the optical redshift and line-of-sight velocity dispersion (where available) information, these X-ray measures permit a comparison of the cluster sample with lower-redshift samples and in particular permit the approximate mass interval occupied by $z < 0.6$ *XMM*–LSS clusters to be understood.

### 5.1 Morphological properties

The X-ray surface brightness distribution of each cluster was modelled employing a circular $\beta$-model, of the form

$$f(r) = \frac{A}{[1 + (r/r_0)^2]^{\alpha}}, \quad (1)$$

where the coordinate $r$ is measured in arcsec with respect to the centre of the X-ray photon distribution, $r_0$ is the core radius, $A$ is the amplitude at $r = 0$, and $\alpha = 3\beta - (1/2)$.

Images and exposure maps for each cluster field were created for the three EPIC instruments (MOS1, MOS2 and pn) separately in the [0.5–2] keV energy band. Images of the appropriate PSF were created (using SAS-CALVIEW), with the appropriate energy weighting together with the off-axis and azimuthal angles of the source. Square regions (of sizes ranging between 175 and 500 arcsec on a side) were selected around each source and around a nearby, source-free, background region. Mask images were also created and employed to remove from further analysis regions associated with chip gaps and serendipitous point sources lying close to each cluster source.

We used the SHERPA package from the CIAO analysis system to fit a model of the form given by equation (1) to the X-ray data for each EPIC instrument. The quality of fit to the three instruments was optimized using the Cash statistic, providing a maximum likelihood fit, which accounts properly for the Poissonian nature of the data. Each model incorporates a flat background model (where the background level is determined employing the associated background region) and a $\beta$-model convolved with the appropriate PSF. For each of the three instrument models determined for each cluster, the values of the core radius $r_0$, position $(x_0, y_0)$ and slope $\alpha$ were constrained to be identical. Only the normalization for each model was permitted to vary.

Table 5 lists the best-fitting structural parameters for each of the confirmed cluster sources. Fig. 2 shows flux contours corresponding to the best-fitting surface brightness model overplotted on the X-ray emission for each candidate. The most interesting outcome from these fits is the low value of the fitted $\beta$ parameter for many of these systems, compared with the typical value of $\beta = 0.66$ determined for clusters (Arnaud & Evrard 1999). The median value of $\beta = 0.45$ found here agrees with the value of 0.46 derived for a sample of low-redshift groups by Heldson & Ponman (2000). Table 5 also lists the value of the normalized $\chi^2$ statistic computed from a comparison of the best-fitting model to the radially averaged surface brightness profile for each source. The value of the $\chi^2$ statistic for each cluster generally indicates that the computed model provides a statistically acceptable fit. The two systems with computed $\chi^2$ values significantly less than one (14 and 16) display some of the lowest count levels in the sample. In these cases, the radially averaged counts may be better described by a Poissonian rather than Gaussian noise distribution – partially invalidating the application of a $\chi^2$ merit function.

**Table 5.** Morphological X-ray parameters determined for confirmed groups and clusters. Displayed errors are $1\sigma$ and, in the case of cluster 19, no reliable error information could be determined. Values of $\chi^2$ were computed over a sum of radial bins for each cluster. These values are provided to indicate the overall merit of each fit.

| Cluster | boxside arcsec | $r_0$ arcsec | $\beta$ | $\chi^2$ (per d.o.f.) |
|---|---|---|---|---|
| XLSSC 006 | 250 | $24.0^{+3.1}_{-2.4}$ | $0.58^{+.04}_{-.02}$ | 1.34 |
| XLSSC 007 | 400 | $24.0^{+14.0}_{-16.0}$ | $0.40^{+.07}_{-.06}$ | 1.15 |
| XLSSC 008 | 250 | $9.6^{+8.7}_{-6.0}$ | $0.44^{+.09}_{-.06}$ | 1.26 |
| XLSSC 009 | 250 | $44.0^{+52.0}_{-19.0}$ | $0.67^{+.73}_{-.17}$ | 0.96 |
| XLSSC 010 | 400 | $8.5^{+3.3}_{-2.6}$ | $0.44^{+.03}_{-.02}$ | 1.68 |
| XLSSC 012 | 500 | $29.2^{+10.0}_{-7.2}$ | $0.54^{+.08}_{-.48}$ | 1.39 |
| XLSSC 013 | 175 | $11.8^{+14.9}_{-6.2}$ | $0.67^{+.67}_{-.16}$ | 1.34 |
| XLSSC 014 | 175 | $4.7^{+39}_{-4.7}$ | $0.40^{+.20}_{-.16}$ | 0.57 |
| XLSSC 016 | 200 | $2.5^{+3.5}_{-2.4}$ | $0.45^{+.10}_{-.07}$ | 0.79 |
| XLSSC 017 | 200 | $12.1^{+7.6}_{-12.0}$ | $0.55^{+.04}_{-.26}$ | 1.11 |
| XLSSC 018 | 250 | $5.2^{+3.4}_{-2.8}$ | $0.42^{+.03}_{-.03}$ | 1.14 |
| XLSSC 020 | 200 | $24.8^{+8.1}_{-22.2}$ | $0.40^{+.04}_{-.16}$ | 0.99 |

### 5.2 Spectral properties

X-ray spectra for each cluster were extracted within a source circle of radius ranging between 30 and 90 arcsec. The corresponding background spectrum was extracted from a surrounding annulus. Sources adjacent to the cluster were flagged and removed from the spectral analysis employing the source region file generated by the original source extraction software. Source mapping used the stacked pn + MOS1 + MOS2 image of each cluster field in the [0.5–2] keV band only. The source extraction and background regions applied to XLSSC 013 are shown in Fig. 3 as an example of the procedure.

Extracted spectral data corresponding to pn + MOS1 + MOS2 detectors were fitted simultaneously. The fitting model consists of an absorbed APEC hot plasma model (Smith et al. 2001) with a metal abundance ratio set to Grevesse & Sauval (1999) values. The hydrogen absorption is modelled using a WABS model with $N_{\rm H}$ fixed at the Galactic value, i.e. $N_{\rm H} \sim 2.6 \times 10^{20}$ cm$^{-2}$. Spectral data are resampled such that the associated background spectrum displays five counts per bin and model values are compared to the data by computing the corresponding value of the C-statistic (see Appendix A for a justification for this approach). Model fitting is performed in two stages. First, the temperature and abundance are fixed ($T = 0.5$ keV, $Z/Z_{\odot} = 0.3$) and only the count normalization is fitted. Once a best-fitting spectrum normalization has been computed, the best-fitting temperature is computed assuming a fixed metal abundance. The fitting results are displayed in Table 6. Table 6 also shows the value of the normalized C-statistic computed from a comparison of the spectral data for each cluster to the quoted model. In each case a statistically reasonable agreement is obtained. Spectral data for all clusters possessing a fitted temperature in Table 6 are displayed in Fig. 4. In the cases of clusters 7, 14, 16, 17 and 20, no spectral fit was possible and a temperature of 1.5 keV (typical of the sample as a whole) was assumed for the purpose of calculating the flux from the system.

X-ray luminosities for each cluster source have been determined within a uniform physical scale derived from the cluster overdensity radius. In the present study we employ the radius $r_{500}$, within which the total mean density of the system is 500 times the critical density of the Universe at the redshift of the system. The value of $r_{500}$ is computed using an isothermal $\beta$-model (Ettori 2000) and employing the fitted gas temperature and $\beta$ value for each system.[4]

[4] We modify equation (A2) of Ettori (2000) to reflect the varying redshift dependence of the Hubble parameter in a matter plus lambda cosmological model.

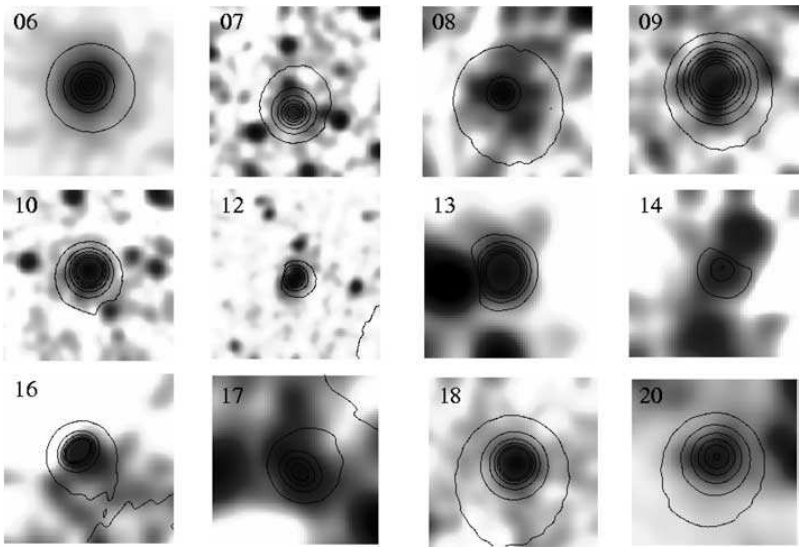


**Figure 2.** Model surface brightness distributions for the X-ray group and cluster sample. Contours represent the best-fitting model convolved with the detector response and are overplotted on the full (pn+MOS1+MOS2) Gaussian smoothed X-ray emission corresponding to the [0.5–2] keV energy band. The Gaussian smoothing scale (i.e. the standard sigma) is 5 pixel or 12.5 arcsec. The linearly spaced contours represent the MOS1 model (only the normalization varies between the three instrument models), except for XLSSC 007 and 013, where the pn model is displayed. The images are of size 'boxside' (see Table 5).

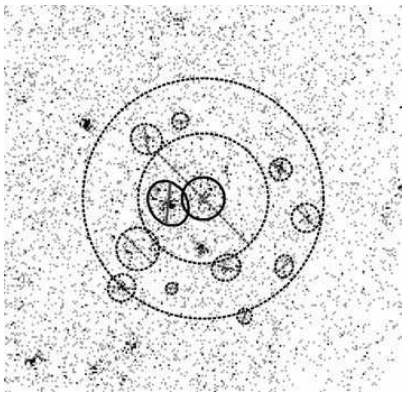

**Figure 3.** Source and background extraction regions applied to XLSSC 013. The grey-scale shows the photon map generated by the stacked pn + MOS1 + MOS2 image. The central solid circle represents the source extraction region. Solid circles crossed with a diagonal line indicate regions excluded from the source aperture. The two large concentric dashed circles define the background annulus and dashed circles crossed with diagonal lines indicate regions excluded from the background aperture.

We have compared the value of $r_{500}$ derived from the above isothermal model for each cluster to the corresponding values obtained using the observed relationship between cluster mass and temperature. Employing the data presented by Finoguenov, Reiprich & Böhringer (2001), based on systems ranging from $T_{\rm X} = 0.75$ to 14 keV, we obtain the following relation

$$r_{500} = 0.391 T_{\rm X}^{0.63}\, h_{70}(z)^{-1}\ {\rm Mpc}, \qquad (2)$$

where $h_{70}\ (z)$ describes the redshift evolution of the Hubble parameter (scaled to 70 km s$^{-1}$ Mpc$^{-1}$) in the assumed cosmological model. The value for $r_{500}$ for each cluster derived using the above methods typically agree to within $\pm 10$ per cent with the maximal difference being $\pm 20$ per cent. Based upon this comparison, we employ $r_{500}$ values based upon the isothermal model in the rest of this paper.

In those cases where a successful spectral fit was obtained, the derived bolometric source flux was extrapolated to $r_{500}$ using the $\beta$-model determined in Section 5.1. For the four systems with detected flux but with no successful spectral fit, the corresponding best-fitting spatial model was employed to compute the bolometric flux within $r_{500}$, assuming a $T = 1.5$ keV emission spectrum. Uncertainties on the resulting value of bolometric luminosity are available only for those systems with spectral fits, in which cases the uncertainty is simply derived from the error on the fitted spectral model normalization. Computation of luminosity uncertainties employing this method does not include the effects of the uncertain extrapolation of the surface brightness model to $r_{500}$. Table 6 shows the aperture correction factor for each cluster, $A$, defined as the

**Table 6.** Spectral X-ray parameters determined for confirmed groups and clusters. Values for exposure time, $t_{\rm exp}$, and total counts are summed over all three detectors. Where the letter 'F' follows a tabulated temperature ($T$) value, this indicates that the value was fixed in the fitting procedure. The definition of the aperture correction factor, $A$, is provided in the text. Displayed errors are $1\sigma$.

| Cluster | $t_{\rm exp}$ s | Total counts | $r_{\rm spec}$ arcsec | $T$ keV | C-stat (per d.o.f.) | $r_{500}$ Mpc | $A$ | $L_{\rm bol}\,(r_{500})$ $\times\,10^{43}$ erg s$^{-1}$ |
|---|---|---|---|---|---|---|---|---|
| XLSSC 006 | 17789 | 1943 | 82.5 | $4.80^{+1.12}_{-0.84}$ | 0.85 | 0.809 | 1.29 | $36.2 \pm 2.3$ |
| XLSSC 007 | 28094 | 138 | 90 | 1.5F | 1.10 | 0.284 | 0.65 | 1.1 |
| XLSSC 008 | 32358 | 94 | 60 | $1.25^{+1.44}_{-0.38}$ | 1.04 | 0.393 | 1.62 | $0.5 \pm 0.2$ |
| XLSSC 009 | 10709 | 112 | 90 | $0.91^{+0.20}_{-0.17}$ | 1.12 | 0.292 | 0.93 | $1.1 \pm 0.3$ |
| XLSSC 010 | 22635 | 505 | 67.5 | $2.40^{+0.82}_{-0.53}$ | 1.00 | 0.539 | 1.50 | $4.6 \pm 0.5$ |
| XLSSC 012 | 37726 | 635 | 60 | $2.00^{+1.28}_{-0.51}$ | 1.20 | 0.462 | 1.52 | $3.0 \pm 0.4$ |
| XLSSC 013 | 34383 | 133 | 35 | $1.03^{+0.18}_{-0.25}$ | 0.92 | 0.437 | 1.38 | $0.5 \pm 0.1$ |
| XLSSC 014 | 14801 | 286 | 50 | 1.5F | 1.26 | 0.404 | 1.59 | 0.4 |
| XLSSC 016 | 27202 | 25 | 30 | 1.5F | 0.99 | 0.432 | 1.76 | 0.4 |
| XLSSC 017 | 25506 | 79 | 30 | 1.5F | 1.14 | 0.456 | 1.50 | 0.6 |
| XLSSC 018 | 62573 | 295 | 45 | $2.66^{+2.47}_{-0.91}$ | 1.40 | 0.558 | 2.32 | $1.3 \pm 0.2$ |
| XLSSC 020 | 16770 | 61 | 37.5 | 1.5F | 1.09 | 0.305 | 1.01 | 2.0 |

relative change in the integrated $\beta$-model profile obtained by varying the integration limit from $r_{\rm spec}$ to $r_{500}$.

## 6 NATURE OF *XMM*–LSS SELECTED GROUPS AND CLUSTERS AT $Z < 0.6$

In this section we compare the properties of the *XMM*–LSS groups and clusters at $z < 0.6$ with X-ray group and cluster samples in the literature. Fig. 5 compares the $L_{\rm X}$ versus $T_{\rm X}$ distribution of *XMM*–LSS clusters confirmed at $z < 0.6$ to both the distribution formed by the Group Evolution Multiwavelength Study (GEMS) sample of local ($z < 0.03$) X-ray emitting galaxy groups (Osmond & Ponman 2004, hereafter OP04) and the sample of Markevitch (1998, hereafter M98) containing clusters at $z < 0.09$ with *ASCA* temperatures and *ROSAT* luminosities.[5] Fig. 5 indicates that *XMM*–LSS clusters occupy a region of the $L_{\rm X}$ versus $T_{\rm X}$ plane ranging from cool ($T_{\rm X} > 0.9$ keV), low-luminosity ($L_{\rm X}\,(r_{500}) > 4 \times 10^{42}$ erg s$^{-1}$) X-ray groups to moderate temperature ($T_{\rm X} = 5$ keV), moderate-luminosity ($L_{\rm X}\,(r_{500}) = 4 \times 10^{44}$ erg s$^{-1}$) clusters. Although the sample of X-ray systems presented in this paper is not statistically complete, it is representative of the properties of the complete flux-limited sample currently under construction. Due to the steeply rising nature of the XLF toward faint X-ray systems, it is anticipated that the larger, statistically complete sample of X-ray structures identified by the *XMM*–LSS survey at $z < 0.6$ will be dominated by such galaxy group and low-mass cluster systems.

It can be seen from Fig. 5, that our *XMM*–LSS groups and clusters appear to be in good agreement with a linear fit to the $L_{\rm X}$ versus $T_{\rm X}$ distribution of lower-redshift group and cluster samples. The local fit to the $L_{\rm X}$ versus $T_{\rm X}$ distribution takes the form $\log L_{\rm X} = 2.91 \log T_{\rm X} + 42.54$ and is derived from an orthogonal regression fit to the combined OP04 plus M98 samples incorporating a treatment of the selection effects present in each sample (Heldson and Ponman, in preparation). The comparison of *XMM*–LSS groups and clusters to this local relationship may be quantified (Fig. 6) by calculating a luminosity enhancement factor, $F = L_{\rm obs}/L_{\rm pred}$, where $L_{\rm obs}$ is the observed cluster X-ray luminosity within a radius, $r_{500}$, and $L_{\rm pred}$ is the luminosity expected applying the fitted $L_{\rm X}$ versus $T_{\rm X}$ relation computed for the local fit and the *XMM*–LSS measured temperature. Neglecting the five systems assigned a fixed temperature (XLSSC 007, 014, 016, 017 and 020 – for which the temperature uncertainty is unknown), the median enhancement factor of the remaining six systems is $F = 1.09$. For comparison, the expected enhancement in $L_{\rm X}$ due to self-similar evolution, scaling to $r_{500}$, is a factor of 1.23 at the typical redshift ($z = 0.4$) of our sample. Therefore, given the observed spread in the enhancement values, the observed negative deviation from the self-similar expectation is not large.

Given the modest size and the statistically incomplete nature of our current sample, we regard these results as in need of confirmation. Ettori et al. (2004) also report evolution weaker than the self-similar expectation from a sample of 28 clusters at $z > 0.4$ with gas temperatures $3 < T < 11$ keV. The combined effect of self-similar scaling with the negative evolution reported by Ettori et al. (2004) would result in an enhancement factor $F = 0.86$[6] at a $z = 0.4$ (see Fig. 6). Although the overlap of the Ettori et al. (2004) sample and the systems contained in the present work is limited, the X-ray luminosities appear to describe similar trends.

The relationship between the specific energy contained within the cluster galaxy motions, compared to the X-ray emitting gas, is described via the $\beta_{\rm spec}$ parameter

$$\beta_{\rm spec} \equiv \frac{\sigma_v^2}{kT_{\rm X}/\mu m_{\rm p}}, \qquad (3)$$

[5] Note that, although luminosities in M98 are quoted within $1\,h_{100}^{-1}$ Mpc apertures and not $r_{500}$ as used in this paper, the appropriate correction factors are typically considerably smaller than the 5 per cent calibration uncertainty associated with the luminosities themselves.

[6] This enhancement factor assumes self-similar evolution and an additional factor $(1 + z)^{B_z}$ where $B_z = -1.04$, following the nomenclature of Ettori et al. (2004).

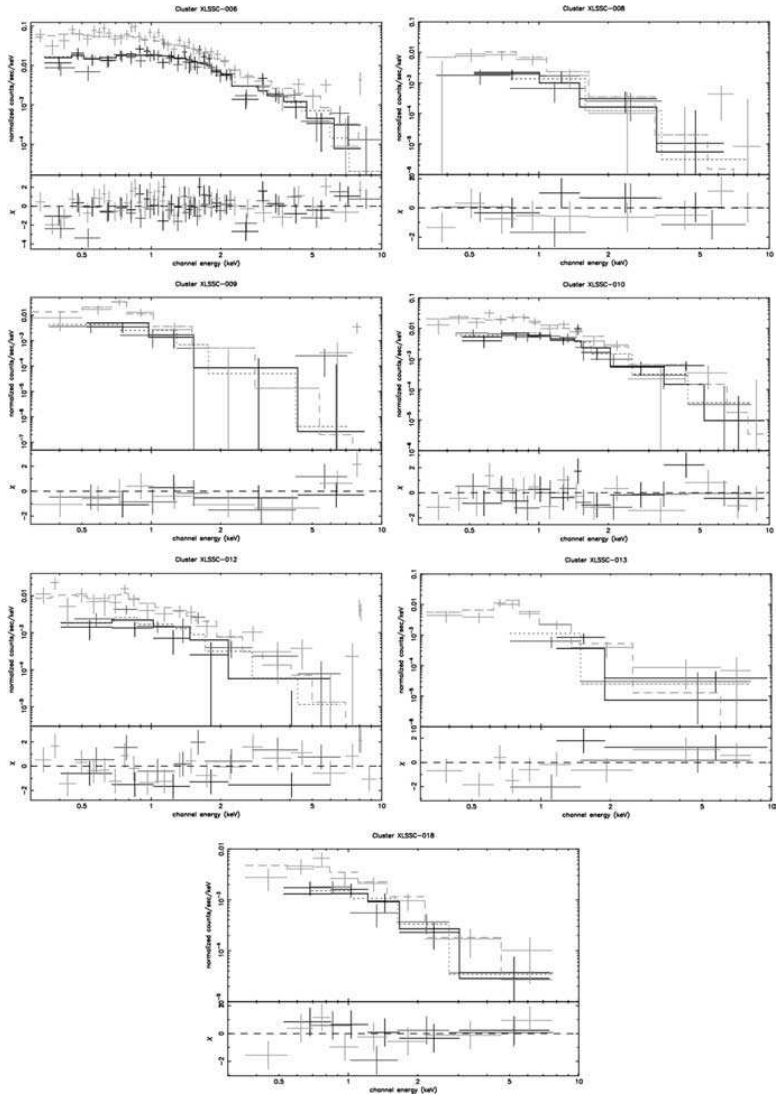


**Figure 4.** Spectral data for groups and clusters described in Table 6 for which a temperature was fitted. For each system the upper panel depicts the spectral photon flux for the MOS1 (thick black crosses), MOS2 (thin black crosses) and pn (grey crosses). The data have been resampled to 20 photons per spectral bin for display purposes only. The spectral model applicable to each detector is also shown by the appropriately coloured solid line, i.e. MOS1 (solid black line), MOS2 (dotted black line) and pn (grey dashed line). The lower panel shows the deviation of the data from the model for each detector expressed in units of normalized $\chi^2$ per spectral bin.

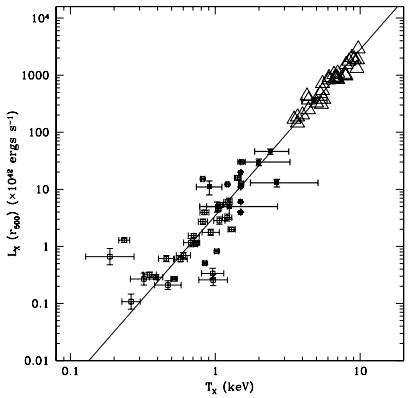


**Figure 5.** Distribution of X-ray luminosity computed within a scale radius $r_{500}$ and temperature for all *XMM*–LSS groups and clusters currently identified at $z \leqslant 0.6$ (solid squares). Also indicated are values of X-ray luminosity and temperature determined for the low-redshift group sample of OP04 (open squares) and for the cluster sample of M98 (open triangles). The solid line indicates an orthogonal regression fit to the $L_X$ versus $T_X$ relation for both the group and cluster sample incorporating a treatment of the selection effects present in each sample (Heldson & Ponman, in preparation) – see text for details.

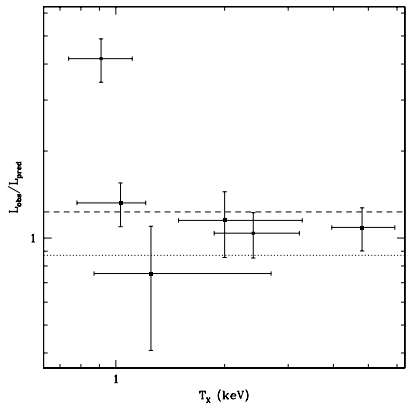


**Figure 6.** Enhancement factor, $F = L_{\rm obs}/L_{\rm pred}$, computed for six *XMM*–LSS groups and clusters located at $z \leqslant 0.6$ plotted versus the X-ray temperature of each system (see text for additional details). Horizontal lines indicate expected values of $F$; the short dashed line indicates the value $F = 1.23$ expected from self-similar considerations. The dotted line indicates the value of $F$ expected at $z = 0.4$ based upon Ettori et al. (2004).

where $\sigma_v$ is the line-of-sight galaxy velocity distribution, $k$ is the Boltzmann constant, $T_X$ is the X-ray gas temperature and $\mu m_p$ is the mean particle mass within the gas (Bahcall & Lubin 1994). Fig. 7 displays the value of $\beta_{\rm spec}$ computed for the *XMM*–LSS $z < 0.6$ cluster sample as a function of computed X-ray luminosity extrapolated to a radius $r_{500}$ (i.e. the consistent measure adopted in this paper).

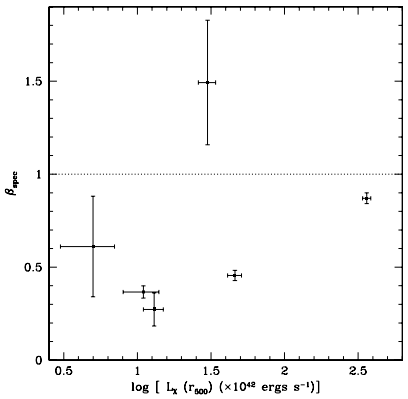


**Figure 7.** Values of $\beta_{\rm spec}$ computed for eight *XMM*–LSS groups and clusters at $z \leqslant 0.6$ (see text for details) plotted versus the X-ray luminosity for each system. The horizontal dashed line indicates the value $\beta_{\rm spec} = 1$.

A number of systems have been excluded from Fig. 7: in addition to the systems excluded from the computation of the luminosity enhancement factor above, XLSSC 013 does not possess well-defined galaxy velocity dispersion, thus preventing computation of $\beta_{\rm spec}$. Fig. 7 indicates typical values $\beta_{\rm spec} < 1$ for the *XMM*–LSS $z < 0.6$ sample (with the exclusion of the above-mentioned systems), the median value of $\beta_{\rm spec}$ for this restricted sampled is $\langle \beta_{\rm spec} = 0.61 \rangle$. This value may be compared to values of $\beta_{\rm spec} \approx 1$ reported by OP04 for luminous X-ray groups, i.e. $L_X(r_{500}) > 10^{42}$ erg s$^{-1}$.

A clear concern when interpreting the trend of low $\beta_{\rm spec}$ values is the extent to which the cluster galaxy velocity dispersion estimates may be biased toward lower values. Potential uncertainties associated with the velocity dispersion values presented in this paper are discussed in Section 4. However, in the overwhelming majority of clusters observed in detail, the integrated velocity dispersion profile of galaxy clusters is a decreasing function of projected radius from the cluster centre (Girardi et al. 1996; Borgani et al. 1999). If the integrated velocity dispersion profiles of the *XMM*–LSS clusters presented in this paper display similar behaviour to hotter clusters, then the expectation arising from computation of the cluster velocity at some fraction of the convergent radius is that the velocity dispersion will be overestimated and will result in values of $\beta_{\rm spec}$ biased to higher values. The extent of any such bias is difficult to quantify in the current data set. However, the implication is that the value of $\beta_{\rm spec}$ displayed in Fig. 7 would not increase with the addition of velocity dispersion measurements extending to larger radii.

The low values of $\beta_{\rm spec}$ apparent in Fig. 7 are similar to those seen by OP04 in lower-luminosity groups ($L_X < 10^{42}$ erg s$^{-1}$) at low redshift. The origin of these low values of $\beta_{\rm spec}$ is far from clear, but OP04 argue that it appears to result primarily from a reduction in $\sigma_v$, rather than an enhancement in $T_X$. Whatever the cause, our results provide tentative evidence that these effects are operating in hotter and more X-ray luminous systems at higher redshift. We are currently in the process of conducting magnitude-limited spectroscopy of a sample of $T_X \sim 1$ keV systems at $z = 0.3$ in order to provide a more robust picture of the dynamics of low temperature X-ray systems.

## 7 CONCLUSIONS

We have presented 12 newly identified X-ray selected groups and clusters as part of the *XMM*–LSS survey. The procedures employed to detect and classify sources in X-rays, and to subsequently confirm each source via optical imaging and spectroscopic observations have been described in detail.

We have emphasized throughout this paper that the current sample of clusters is not complete in any statistical sense. The presentation of a larger, complete sample of X-ray clusters located at $z < 0.6$ will form part of a future publication. However, the current sample of X-ray clusters at $z < 0.6$ presents a number of interesting features: most importantly, the sample is dominated by low X-ray temperature systems located at redshifts much greater than that presented by previous X-ray studies. Such systems are predicted to display the effects of pre-heating or additional energy input into the ICM to a greater extent than hotter, more massive systems. The identification of such low-temperature systems at look-back times up to 5.7 Gyr provides an important baseline over which to study the extent to which such systems evolve.

We find tentative evidence that these high-redshift groups are more luminous than local systems, at a given temperature, in agreement with recent work on richer clusters. However, our results suggest that group luminosities may be evolving less rapidly than higher temperature clusters when compared to self-similar models. If this is confirmed to be the case, then the steepening of the $L_X$–$T_X$ relation at low temperatures reported in local samples, may continue at higher redshift. We also find preliminary indications that the poorly understood tendency for the specific energy in the gas to exceed that in the galaxies in poor groups, extends to systems with higher values of $L_X$ and $T_X$ at $z \sim 0.4$. The completion of a larger and statistically complete sample of intermediate redshift groups from the *XMM*–LSS survey, should allow these results to be placed on a firm statistical footing in the near future.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge Steve Heldson for his assistance with comparisons with the low-redshift X-ray group and cluster samples. The authors additionally thank Jean Ballet and Keith Arnaud for useful discussions on the statistical treatment inside XSPEC and Jean-Luc Sauvageot for technical discussions regarding *XMM* calibration for fitting purposes.

## REFERENCES

Allen S., Schmidt R., Fabian A., Ebeling H., 2003, MNRAS, 342, 287
Andreon S., Ettori S., 1999, ApJ, 516, 647
Andreon S., Willis J., Quintana H., Valtchanov I., Pierre M., Pacaud F., 2004a, MNRAS, 353, 353
Andreon S., Valtchanov I., Jones L. R., Altieri B., Bremer M., Willis J., Pierre M., Quintana H., 2004b, MNRAS, 359, 1250
Arnaud M., Evrard A., 1999, MNRAS, 305, 631
Bahcall N., Lubin L., 1994, ApJ, 426, 513
Baugh C., Cole S., Frenk C., 1996, MNRAS, 283, 1361
Beers T., Flynn K., Gebhardt K., 1990, AJ, 100, 32
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Böhringer H. et al., 2002, ApJ, 566, 93
Borgani S., Girardi M., Carlberg R., Yee H., Ellingson E., 1999, ApJ, 527, 561
Borgani S. et al., 2001, ApJ, 561, 13
Burke D., Collins C., Sharples R., Romer A. K., Holden B. P., Nichol R. C., 1997, ApJ, 488, 83
Danese L., de Zotti G., di Tullio G., 1980, A&A, 82, 322
Dickey J. M., Lockman F. J., 1990, ARA&A, 28, 215
Dressler A., Smail I., Poggianti B., Butcher H., Couch W. J., Ellis R. S., Oemler A. Jr, 1999, ApJS, 122, 51
Ebeling H., Edge A., Henry J., 2001, ApJ, 553, 668
Ettori S., 2000, MNRAS, 311, 313
Ettori S., Tozzi P., Borgani S., Rosati P., 2004, A&A, 417, 13
Finoguenov A., Reiprich T. H., Böhringer H., 2001, A&A, 368, 749
Girardi M., Fadda D., Giuricin G., Mardirossian F., Mezzetti M., 1996, ApJ, 457, 61
Grevesse N., Sauval A. J., 1999, A&A, 347, 348
Hamuy M., Walker A., Suntzeff N., 1992, PASP, 104, 533
Hamuy M., Suntzeff N., Heathcote S., 1994, PASP, 106, 566
Heavens A., 1993, MNRAS, 263, 735
Heldson S., Ponman T., 2000, MNRAS, 319, 933
Heldson S., Ponman T., 2003, MNRAS, 339, 29
Henry J., Gioia I., Maccacaro T., Morris S. L., Stocke J. T., Wolter A., 1992, ApJ, 386, 408
Kauffmann G., 1996, MNRAS, 281, 487
Kinney A., Calzetti D., Bohlin R., McQuade K., Storchi-Bergmann T., Schmitt H. R., 1996, ApJ, 467, 38
Kodama T., Arimoto N., Barger A., Aragon-Salamanca A., 1998, A&A, 334, 99
Ledlow M., Loken C., Burns J., Owen F., Voges W., 1999, ApJ, 516, L53
Lubin L., Oke B., Postman M., 2002, AJ, 124, 1905
Lumb D. et al., 2004, A&A, 420, 853
McCracken H. et al., 2003, A&A, 410, 17
Markevitch M., 1998, ApJ, 504, 27 (M98)
Morrison R., McCammon D., 1983, ApJ, 270, 119
Osmond J. P. F., Ponman T. J., 2004, MNRAS, 350, 1511 (OP04)
Pierre M. et al., 2004, JCAP, 9, 11
Ponman T., Cannon D., Navarro J., 1999, Nat, 397, 135
Refregier A., Valtchanov I., Pierre M., 2002, A&A, 390, 1
Rosati P., della Ceca R., Norman C., Giacconi R., 1998, ApJ, 492, 21
Rosati P., Borgani S., Norman C., 2002, ARA&A, 40, 539
Sanderson A., Ponman T., Finoguenov A., Lloyd-Davies E. J., Markevitch M., 2003, MNRAS, 340, 989
Schuecker P. et al., 2001, A&A, 368, 86
Silverman B., 1986, Density Estimation for Data Analysis and Statistics. Chapman and Hall, London
Smith R. K., Brickhouse N. S., Liedahl D. A., Raymond J. C., 2001, ApJ, 556, L91
Stanford S., Eisenhardt P., Dickinson M., 1998, ApJ, 492, 461
Starck J., Pierre M., 1998, A&AS, 128, 397
Tonry J., Davis M., 1979, AJ, 84, 1511
Valtchanov I., Pierre M., Gastaud R., 2001, A&A, 370, 689
Valtchanov I. et al., 2004, A&A, 423, 75
Vikhlinin A., McNamara B., Forman W., Jones C., Quintana H., Hornstrup A., 1998, ApJ, 498, L21
Yee H., Ellingson E., Carlberg R., 1996, ApJS, 102, 269

## APPENDIX A: FITTING SIMULATED CLUSTER SPECTRA

The current study extends X-ray spectral observations of distant galaxy groups and clusters to low integrated signal levels (∼100 photons above the background). It is therefore prudent to assess the reliability of temperature measures and associated uncertainties computed via model fits to such faint spectra by repeating the fitting procedure for a grid of simulated spectra created to reproduce the properties of the observed sample.

The source model used to simulate group and cluster spectra employs an APEC of an optically thin plasma (Smith et al. 2001). This model depends upon four parameters: temperature, metal abundance, redshift and a normalization representative of the emission integral. To simulate galaxy groups and clusters the abundance is

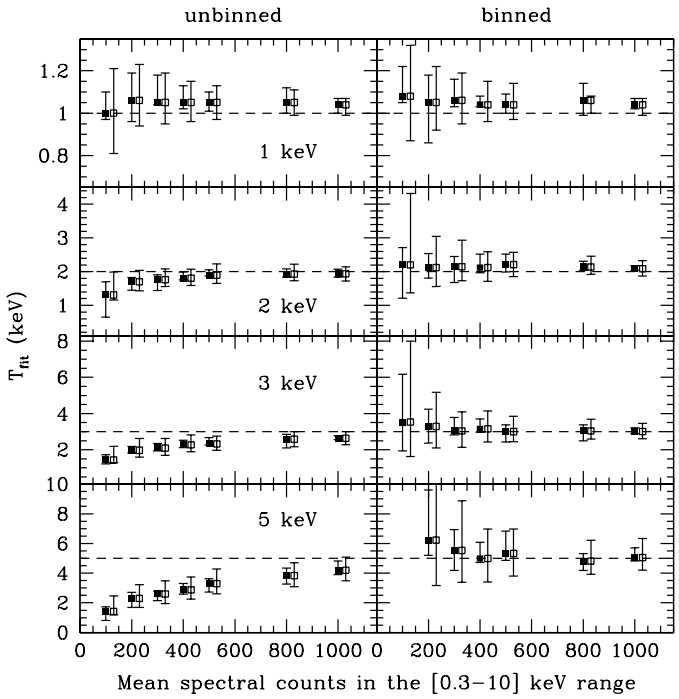


**Figure A1.** A comparison of XSPEC computed temperatures for simulated group and cluster spectra employing the C-statistic with two different resampling approaches. Left-hand panels indicate the results for unbinned spectra. Right-hand panels indicate the results for binned spectra such that the background spectrum displays a minimum of five counts per spectral bin. Panels in each row correspond to spectral models with the indicated input temperature (also shown by the horizontal dashed line). In each panel, data points represent the mean XSPEC computed temperature returned from the set of simulated spectra as a function of total input counts. Filled squares plus error bars indicate the mean computed temperature and the distribution of temperatures accounting for 68 per cent of the sample. Open squares plus error bars indicate the mean computed temperature and the median $1\sigma$ uncertainty returned by XSPEC (open squares are shifted to the right by 30 counts with respect to the filled squares for clarity).

set to $Z/Z_{\odot} = 0.3$ with solar abundance ratios set to Grevesse & Sauval (1999) values. The redshift of the simulated source is set to $z = 0.3$, typical of the sources presented in this paper. Photoelectric absorption described by a WABS model within XSPEC using Morrison & McCammon (1983) cross-sections was applied to the source model with the neutral hydrogen column density fixed to $N_{\rm H} = 2.6 \times 10^{20}$ cm$^{-2}$ – the mean value for our sample according to the H I distribution map of Dickey & Lockman (1990). The instrumental response was modelled using the redistribution matrix and ancillary response files from observations of XLSSC 006.

A model pn + MOS1 + MOS2 spectra was created and a conversion factor applied to generate spectra of the required integrated count level over the spectral interval [0.3–10] keV in a 10 000-s exposure. Each simulated spectrum is generated from this model using Poissonian considerations. Each source spectrum is accompanied by a background spectrum created from a Poisson realization of a background model of normalization and shape consistent with observed cluster backgrounds. Spectra were simulated according to this procedure for temperatures equal to 1, 2, 3 and 5 keV. At each temperature, 50 spectra were simulated at each point of a grid of integrated count levels 100, 200, . . . , 1000.

Spectral fitting follows the same approach as applied to observed data, i.e. temperature and spectrum normalization are permitted to vary while the abundance is fixed. Data from pn + MOS1 + MOS2 are are combined within XSPEC with the response files from XLSSC 006 and the energy range [0.3–10] keV is conserved. The spectral energy range corresponding to [7.5–8.5] keV measured by the pn detector is ignored as it contains strong instrumental line emission that is not well corrected by our data modelling process. The best-fitting model is then determined by minimizing a modified C-statistic and the $1\sigma$ uncertainty about the best-fitting model is computed. While the C-statistic is intended to work efficiently on unbinned data, a comparison of the fitted temperature to the input value indicates a tendency to underestimate the temperature of spectra of input temperatures <5 keV displaying count values <1000 counts using this procedure. This bias is indicated in Fig. A1 and appears to arise from the fact that the statistic used in XSPEC represents a modified C-statistic that accounts for statistical fluctuations in the

background estimation. Our understanding of the problem is that this modified statistic fails at estimating model parameters when there is a significant number of background bins containing zero photons.

Resampling the data to prevent the occurrence of spectral bins containing zero counts minimizes this negative temperature bias. The resampling factor is determined by requiring that the background spectrum associated with each source display a specified minimum count level per spectral bin. Determining the resampling factor from the background spectrum represents a sensible approach as the background counts are more numerous and therefore minimize the loss of spectral information in the source spectrum. Although a small positive bias is introduced to the fitted temperatures of very low count level spectra (<300 counts) when the data are resampled, the amplitude of this bias is less than 10 per cent when the data are resampled to contain five counts per spectral bin (Fig. A1). Applying a larger resampling factor increases the positive temperature bias, which can be understood in terms of the spectral smoothing that the resampling procedure represents.

We therefore resample the observed data to generate a background spectrum containing five counts per spectral bin. This approach generates fitted temperatures that agree with the input temperature to <10 per cent. In addition, comparison of the distribution of fitted temperatures at any given combination of temperature and count level to the temperature uncertainty returned by XSPEC indicates that the XSPEC quoted errors on fitted temperatures overestimate the distribution of fitted temperatures by a factor typically less than 2. Due to the various assumptions that enter the simulation procedure it therefore seems reasonable to provide XSPEC quoted temperature uncertainties as a conservative error estimate.

This paper has been typeset from a TEX/LATEX file prepared by the author.

Mon. Not. R. Astron. Soc. **364,** 751–752 (2005) doi:10.1111/j.1365-2966.2005.09651.x

# Erratum: The *XMM* Large-Scale Structure survey: an initial sample of galaxy groups and clusters to a redshift $z < 0.6$

by J. P. Willis,⋆ F. Pacaud, I. Valtchanov, M. Pierre, T. Ponman, A. Read, S. Andreon, B. Altieri, H. Quintana, S. Dos Santos, M. Birkinshaw, M. Bremer, P.-A. Duc, G. Galaz, E. Gosset, L. Jones and J. Surdej

**Key words:** errata, addenda – surveys – large-scale structure of the Universe – X-rays: galaxies: clusters.

The paper 'The *XMM* Large-Scale Structure Survey: An initial sample of galaxy groups and clusters to a redshift $z < 0.6$' was published in Mon. Not. R. Astron. Soc., **363**, 675-691. The values of bolometric luminosity computed within an aperture of radius $r_{500}$ for each cluster, and the associated uncertainty, were obtained using fluxes that were not corrected for Galactic neutral hydrogen absorption. This error affects the results presented in Table 6 and Figs 5, 6 and 7, revised versions of which are presented here. The typical absorption correction per cluster results in a luminosity change of factor 1.5 and this results in a minor modification of our original conclusions.

In the original paper we compared the luminosity of the *XMM*–LSS group and cluster sample to local X-ray groups and clusters by means of a luminosity enhancement factor. The median luminosity enhancement factor of the clusters with reliable temperature information is now $F = 1.46$ (the previous value based upon unabsorbed luminosities was 1.09) – indicating that the systems we are observing are slightly brighter than the expectation for self–similar evolution from systems at $z = 0$ displaying the same temperature, i.e. $F_{\rm ss} = 1.23$. However, we repeat the caveat that these initial trends require further confirmation within a larger, more complete sample, the compilation of which is currently underway.

**Table 6.** Spectral X-ray parameters determined for confirmed groups and clusters. Values for exposure time, $t_{\rm exp}$, and total counts are summed over all three detectors. Where the letter 'F' follows a tabulated temperature (T) value, this indicates that the value was fixed in the fitting procedure. The definition of the aperture correction factor, $A$, is provided in the original paper text. Displayed errors are $1\sigma$.

| Cluster | $t_{\rm exp}$ seconds | total counts | $r_{\rm spec}$ arcsec | T keV | C–stat (per d.o.f.) | $r_{500}$ Mpc | $A$ | $L_{\rm bol}(r_{500})$ $\times 10^{43}$ erg s$^{-1}$ |
|---|---|---|---|---|---|---|---|---|
| XLSSC 006 | 17789 | 1943 | 82.5 | $4.80^{+1.12}_{-0.84}$ | 0.85 | 0.809 | 1.29 | $44.8 \pm 2.8$ |
| XLSSC 007 | 28094 | 138 | 90 | 1.5F | 1.10 | 0.284 | 0.65 | 1.9 |
| XLSSC 008 | 32358 | 94 | 60 | $1.25^{+1.44}_{-0.38}$ | 1.04 | 0.393 | 1.62 | $0.8 \pm 0.3$ |
| XLSSC 009 | 10709 | 112 | 90 | $0.91^{+0.20}_{-0.17}$ | 1.12 | 0.292 | 0.93 | $2.0 \pm 0.5$ |
| XLSSC 010 | 22635 | 505 | 67.5 | $2.40^{+0.82}_{-0.53}$ | 1.00 | 0.539 | 1.50 | $6.5 \pm 0.7$ |
| XLSSC 012 | 37726 | 635 | 60 | $2.00^{+1.28}_{-0.51}$ | 1.20 | 0.462 | 1.52 | $4.5 \pm 0.6$ |
| XLSSC 013 | 34383 | 133 | 35 | $1.03^{+0.18}_{-0.25}$ | 0.92 | 0.437 | 1.38 | $0.9 \pm 0.2$ |
| XLSSC 014 | 14801 | 286 | 50 | 1.5F | 1.26 | 0.404 | 1.59 | 0.6 |
| XLSSC 016 | 27202 | 25 | 30 | 1.5F | 0.99 | 0.432 | 1.76 | 0.6 |
| XLSSC 017 | 25506 | 79 | 30 | 1.5F | 1.14 | 0.456 | 1.50 | 1.0 |
| XLSSC 018 | 62573 | 295 | 45 | $2.66^{+2.47}_{-0.91}$ | 1.40 | 0.558 | 2.32 | $1.8 \pm 0.3$ |
| XLSSC 020 | 16770 | 61 | 37.5 | 1.5F | 1.09 | 0.305 | 1.01 | 3.4 |

⋆E-mail: jwillis@uvic.ca

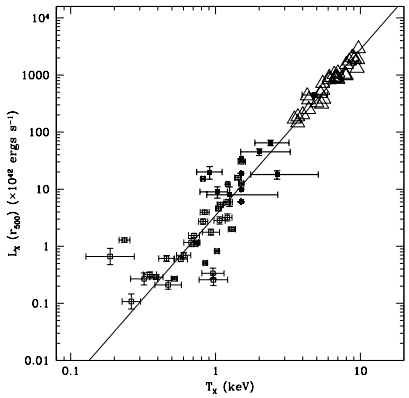


**Figure 5.** Distribution of X-ray luminosity computed within a scale radius $r_{500}$ and temperature for all *XMM*–LSS groups and clusters currently identified at $z \leqslant 0.6$ (solid squares). Also indicated are values of X-ray luminosity and temperature determined for the low-redshift group sample of Osmond & Ponman (2004) (open squares) and for the cluster sample of Markevitch (1998) (open triangles). The solid line indicates an orthogonal regression fit to the $L_X$ versus $T_X$ relation for both the group and cluster sample incorporating a treatment of the selection effects present in each sample – see original paper text for details.

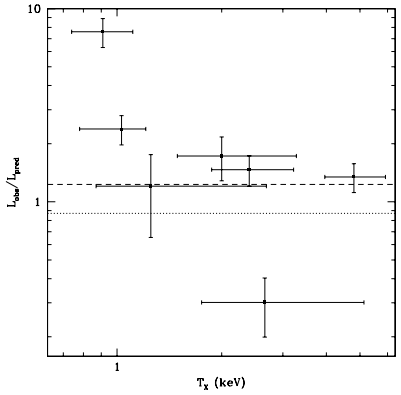


**Figure 6.** Enhancement factor, $F = L_{\rm obs}/L_{\rm pred}$, computed for six *XMM*–LSS groups and clusters located at $z \leqslant 0.6$ plotted versus the X-ray temperature of each system (see original paper text for additional details). Horizontal lines indicate expected values of $F$: the short dashed line indicates the value $F = 1.23$ expected from self–similar considerations. The dotted line indicates the value of $F$ expected at $z = 0.4$ based upon Ettori et al. (2004).

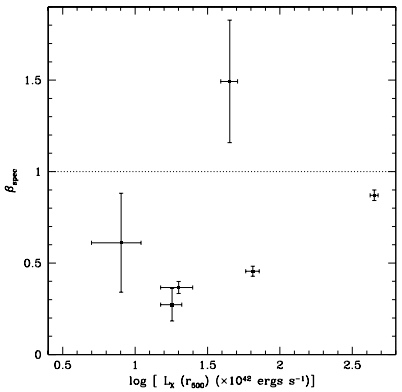


**Figure 7.** Values of $\beta_{\rm spec}$ computed for six *XMM*–LSS groups and clusters at $z \leqslant 0.6$ (see original paper text for details) plotted versus the X-ray luminosity for each system. The horizontal dashed line indicates the value $\beta_{\rm spec} = 1$.

## REFERENCES

Ettori S., Tozzi P., Borgani S., Rosati P., 2004, A&A, 417, 13
Markevitch M., 1998, ApJ, 504, 27
Osmond J. P. F., Ponman T. J., 2004, MNRAS, 350, 1511

This paper has been typeset from a TEX/LATEX file prepared by the author.

## 7.2 La définition d'échantillons controllés :

## "The XMM LSS survey : the X-ray pipeline and survey selection function"

### Contexte/motivation

En vu des applications cosmologiques du relevé, il s'avère indispensable de pouvoir construire des échantillons controllés, dont la fonction de sélection peut être précisément estimée. C'est pourquoi le développement du pipeline X et des outils de simulations à pris une place importante dans mon travail de thèse. Une fois les premières détections publiées, montrer que de tels outils étaient disponibles pour le relevé constituait donc une priorité.

### Résumé

L'article présente le pipeline X developpé dans le but de rechercher des amas dans le relevé XMM-LSS. Celui-ci est basé sur une procédure en deux étapes visant à prendre en compte la nature poissonnienne du signal : (i) detection de source sur des images filtrées en ondelette dans le cadre d'une analyse multirésolution et (ii) analyse des sources au moyen d'un ajustement des photons de l'image par la méthode du maximum de vraisemblance. Les performances pour détecter et caractériser les sources sont testées via des simulations Monte Carlo exhaustives. Cela nous a conduit à définir deux écantillons de sources étendues : la classe C1 qui est dépourvu de contamination, et la classe C2 moins restrictive qui autorise une contamination de 50%. La fonction de sélection correspondante est présentée et la comparaison aux amas XMM-LSS d'hors et déjà confirmés montre un bon accord. Nous arrivons à une densité moyenne de sources prédite d'environ 7 source C1 et 12 C2 par deg$^2$, ce qui est plus élevé qu'aucun relevé X à grand champ disponible à ce jour. Nous remarquons finalement une différence substentielle entre la distribution en redshift attendue pour notre échantillon et celle obtenue en utilisant l'approximation d'un échantillon limité en flux.

### Contribution personnelle

L'ensemble de l'article correspond essentiellement à des travaux personnels (plus de détails sont disponibles dans le chapitre 4). La modélisation cosmologique utilise certaines procédures originellement créées par A. Refregier, adaptées et étendues pour nos besoins.

Mon. Not. R. Astron. Soc. **372,** 578–590 (2006) doi: 10.1111/j.1365-2966.2006.10881.x

# The *XMM* Large-Scale Structure survey: the X-ray pipeline and survey selection function

F. Pacaud,[1⋆] M. Pierre,[1] A. Refregier,[1] A. Gueguen,[1] J.-L. Starck,[1] I. Valtchanov,[2†] A. M. Read,[3] B. Altieri,[4] L. Chiappetti,[5] P. Gandhi,[6] O. Garcet,[7] E. Gosset,[7] T. J. Ponman[8] and J. Surdej[7]

[1] *CEA/DSM/DAPNIA, Service d'astrophysique, F-91191 Gif-sur-Yvette, France*
[2] *Astrophysics Group, Blackett Laboratory, Imperial College of Science Technology and Medicine, London SW7 2BW*
[3] *Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH*
[4] *ESA, Villafranca del Castillo, Spain*
[5] *INAF IASF, Sezione di Milano 'G.Occhialini', via Bassini 15, 20133 Milano, Italy*
[6] *European Southern Observatory, Casilla 19001, Santiago, Chile*
[7] *Institut d'Astrophysique et de Géophysique, Université de Liège, Allée du 6 Août, 17, B5C, 4000 Sart Tilman, Belgium*
[8] *School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT*

Accepted 2006 July 12. Received 2006 July 10; in original form 2005 December 20

**ABSTRACT**

We present the X-ray pipeline developed for the purpose of the cluster search in the *XMM* Large-Scale Structure (*XMM*-LSS) survey. It is based on a two-stage procedure via a dedicated handling of the Poisson nature of the signal: (i) source detection on multiresolution wavelet filtered images and (ii) source analysis by means of a maximum likelihood fit to the photon images. The source detection efficiency and characterization are studied through extensive Monte Carlo simulations. This led us to define two samples of extended sources: the C1 class that is uncontaminated, and the less restrictive C2 class that allows for 50 per cent contamination. The resulting predicted selection function is presented and the comparison to the current *XMM*-LSS confirmed cluster sample shows very good agreement. We arrive at average predicted source densities of about seven C1 and 12 C2 per deg$^2$, which is higher than any available wide field X-ray survey. We finally note a substantial deviation of the predicted redshift distribution for our samples from the one obtained using the usual assumption of a flux-limited sample.

**Key words:** methods: data analysis – surveys – large-scale structure of Universe – X-rays: galaxies: clusters – X-rays: general.

## 1 INTRODUCTION

X-ray imaging is recognized to be one of the most sensitive and reliable methods to detect galaxy clusters. The main reason for this comes from the extended nature of the cluster emission, whose intensity is closely related to the depth of the associated potential well. Moreover, at high galactic latitude and medium-deep X-ray sensitivity ($10^{-14}$–$10^{-15}$ erg s$^{-1}$ cm$^{-2}$ in the [0.5–2] keV band), the mean source density[1] is much lower than in the optical or near-infrared wavelengths. Both aspects concur to significantly lower projection effects that become critical in the optical bands above $z > 0.5$. The task of discovering and characterizing X-ray clusters is, however, complicated by the Poisson nature of X-ray data combined with several instrumental effects (PSF, vignetting, CCD patterns, X-ray and particle background) that have to be disentangled from the intrinsic emission profile of the sources.

With its mosaic of overlapping $10^4$ s *XMM* pointings, the *XMM* Large-Scale Structure (*XMM*-LSS) survey has been designed to detect a significant fraction of the cluster population out to $z = 1$, over an area of several tens of deg$^2$, so as to constitute a sample suitable for cosmological studies (Pierre et al. 2004). Trade-off in the survey design was depth versus coverage, keeping within reasonable limits the total observing time. The two major requirements of the X-ray processing were thus to reach the sensitivity limit of the data in a statistically tractable manner in terms of cluster detection efficiency, and to subsequently provide the selection function of the detected objects.

To achieve these goals, it was necessary to design a new two-step X-ray pipeline, combining wavelet multiresolution analysis

⋆E-mail: pacaud@discovery.saclay.cea.fr
†Present address: ESA, Villafranca del Castillo, Spain.
[1] 100–800 sources per deg$^2$ of which about one-tenth is extended.

and maximum likelihood fits, both using Poisson statistics. This substantial development was required, as our controlled tests on simulated cluster fields revealed unsatisfactory performances for extended sources using the early versions of the official detection software provided by the *XMM*-SAS[2] (see Valtchanov, Pierre & Gastaud 2001). Our approach follows the principles pioneered by Vikhlinin et al. (1998) which were originally established for the *ROSAT* Position Sensitive Proportional Counter (PSPC) data, and that we have totally revised to optimally handle the complex *XMM* instrumental characteristics.

The present paper provides a detailed description of our pipeline – a two-year effort – of its performances and of the resulting computation of the selection function both for point-like and extended sources. Section 2 describes the various steps and parameters of the procedure. Section 3 presents a global evaluation of the pipeline using Monte Carlo image simulations. These are in turn used to define a system of classes for cluster candidate sources, allowing for various degrees of completeness or contamination. Finally, in Section 4, we present a case study for the computation of the survey selection function, relying on the pipeline source classification, in a standard cosmological context.

## 2 PIPELINE DESCRIPTION

The pipeline proceeds in three steps.

(i) Starting from raw observation data files (ODFs), calibrated event lists are created using the *XMM*-SAS tasks EMCHAIN and EPCHAIN. These are then filtered for solar soft proton flares and used to produce images.

(ii) The images are filtered in wavelet space, then scanned by a source detection algorithm set to a very low threshold to obtain a primary source list.

(iii) Detailed properties of each detected source is assessed from the photon images using XAMIN, a maximum likelihood profile fitting procedure. This package was designed for the purpose of the *XMM*-LSS survey, with the specific goal of monitoring in a clean and systematic way the characterization of extended X-ray sources and associated selection effects.

### 2.1 Image extraction

Once event lists have been created, proton flare periods are filtered following the method proposed by Pratt & Arnaud (2002), i.e. using the light curves of high-energy events (10–12 keV for the MOS detectors, 12–14 keV for PN). Histograms of each light curve, binned by 104 s, are produced and fitted by a Poisson law to determine the mean of the distribution, $\lambda$. We then apply a $3\sigma$ threshold, so that time intervals where the emission exceeds $\lambda + 3 * \sqrt{\lambda}$ are thrown out as contaminated.

Images of 2.5 arcsec pixel$^{-1}$ containing single and double events are then produced using EVSELECT in each of the five energy bands: [0.3–0.5], [0.5–2], [2–4.5], [4.5–10] and [2–10] keV.

### 2.2 Source detection

In order to maximize detection rates, and provide good input to the maximum likelihood fit for both point-like and extended sources within an acceptable computation time, we follow the prescription of Valtchanov et al. (2001), extensively tested over numerical simulations, to use a mixed approach combining wavelet filtering of the images and detection with a procedure initially developed for optical images (SEXTRACTOR).

In each band, the three European Photon Imaging Camera (EPIC) detector images are co-added and the resulting image is filtered using the wavelet task MR_FILTER from the multiresolution package MR/1 (Starck, Murtagh & Bijaoui 1998). This task incorporates a statistically rigorous treatment of the Poisson noise which enables the removal of insignificant signal directly in the wavelet space using a thresholding algorithm. A subsequent iterative image reconstruction process accurately recovers the flux and shape of the relevant structures contained in the data. The details of the procedure and an evaluation of its ability to properly reconstruct faint sources in the Poisson regime are given in Starck & Pierre (1998) and Valtchanov et al. (2001).

The primary source catalogues are then derived running SEXTRACTOR (Bertin & Arnouts 1996) on the filtered image. The use of this software is now possible because the multiresolution filtering has removed most of the noise and produced a smoothed background. To avoid border effects, we restrict our analysis to the inner 13 arcmin of the field.[3] With our current settings, the software essentially proceeds in four steps. First, the background level is iteratively estimated in image cells by $3\sigma$ clipping and a full-resolution background map is constructed by bicubic-spline interpolation. Sources are then identified as groups of adjacent pixels matching an intensity level. The software subsequently tries to split blended sources by rethresholding at some sublevels between the original threshold and the peak value of each source and looking for features containing a significant amount of the flux in the emission profile. Finally, a detailed analysis of the source is performed: isophotal analysis to determine source position and shape, and photometry in a flexible elliptical aperture as defined in Kron (1980) and Infante (1987).

Parameters of the source detection steps are summarized in Table 1.

### 2.3 Source validation and characterization: XAMIN

At the end of the pipeline processing, all the sources detected by SEXTRACTOR are analysed by XAMIN using the binned photon images.

For each source, XAMIN determines a model that maximizes the probability of generating the observed spatial photon distribution. First, a point-source model is tested, then an extended source profile parametrized by a spherically symmetric $\beta$-model (Cavaliere & Fusco-Femiano 1976),

$$S_{\rm X}(r) \propto \left[1 + \left(\frac{r}{r_{\rm c}}\right)^2\right]^{-3\beta+1/2}, \quad (1)$$

convolved with the *XMM* point spread function. As we generally do not have enough signal-to-noise ratio with our data to estimate simultaneously both $r_{\rm c}$ and $\beta$, especially with a 2D fit, we decided to fix $\beta$ to the canonical value of $2/3$ that is widely used to model

[2] *XMM* Science Analysis System, http://xmm.vilspa.esa.es/sas/; for subsequent data analysis we used v6.1 of this package.

[3] The centre of the pointing is computed as a sensitivity-weighted average of the optical axis positions of the three telescopes, taken from the exposure map header keywords XCEN and YCEN.

**Table 1.** Relevant parameters of the *XMM*-LSS pipeline detection stage. Note that the high SExtractor detection threshold does not imply that we are being restrictive, but rather reflects the fact that the software is run on already adaptively smoothed images.

| Parameter | Value |
|---|---|
| Event selection: | |
| MOS event flag selection | #XMMEA_EM |
| PN event flag selection | (FLAG & 0x2fb002c)==0 |
| MOS patterns | [0:12] |
| PN patterns | [0:4] |
| Image: | |
| Type | Sky |
| Configuration | Co-addition of EPIC detectors |
| Pixel size | 2.5 arcsec |
| MR/1: | |
| Wavelet type | B-spline |
| Transform algorithm | "à trou" |
| Poisson noise threshold | $10^{-3}$ ($\sim$ Gaussian 3.1$\sigma$) |
| Lowest significant scale | 2 pixel |
| Highest significant scale | 256 pixel |
| SExtractor: | |
| Background cell side | 64 pixel |
| Background median filtering | 4 cells |
| Detection threshold | 6$\sigma$ |
| Detection minimum area | 12 pixel |
| Deblending subthresholds | 64 |
| Deblend min. contrast | 0.003 |

the X-ray emission profile of massive galaxy clusters. Similarly, fitting more sophisticated models (e.g. elliptical) is not justified. The best-fitting parameters for both models are listed in output along with relevant parameters characterizing the source (see list of Table 2).

**Table 2.** XAMIN output parameters.

| Parameter | Content |
|---|---|
| CUTRAD | Size of the fitting box |
| EXP[b] | Mean exposure time in the box |
| GAPFLAG[b] | Distance to nearest CCD gap |
| GAP_NEIGHBOUR | Distance to nearest detected neighbour in the fitting box |
| **EXT** | Best-fitting core radius |
| EXT_STAT | Extension statistic |
| DET_STAT[a] | Detection statistic |
| **X_IMA,Y_IMA**[a] | Best-fitting position in pixel |
| RA,DEC[a] | Best-fitting sky coordinates |
| **RATE_MOS**[a] | EPIC-MOS count rate |
| **RATE_PN**[a] | EPIC-PN count rate |
| SCTS_MOS[a] | Estimated source counts in MOS1+2 |
| SCTS_PN[a] | Estimated source counts in PN |
| BG_MAP_MOS[a] | Background level in MOS1+2 |
| BG_MAP_PN[a] | Background level in PN |
| PIX_DEV[a] | Distance between input/output position |
| N_ITER[a] | Number of AMOEBA iterations |

[a]Listed in the catalogues for both point-like and extended profile fits; [b]issued for each of the three EPIC detectors.
Notes. Free parameters of the fitting process are written in bold font.

*2.3.1 Likelihood model*

The statistic used to assess the reliability of a given model is a simplified version of the C-statistic (Cash 1979)

$$C = 2\sum_{i=1}^{N_{\rm pix}} (m_i - y_i \ln m_i)\,, \tag{2}$$

where $y_i$ is the number of observed photons in pixel $i$ and $m_i$ is the model value in that same pixel. In our specific case, the emission profile of a source is the product of its normalization $N_{\rm mod} = \sum_{i=1}^{N_{\rm pix}} m_i$ and its spatial distribution $d_i$, which are independent: $m_i = N_{\rm mod} \times d_i$. The C-statistic thus reads

$$C = 2\,(N_{\rm mod} - N_{\rm data} \ln N_{\rm mod}) - 2\sum_{i=1}^{N_{\rm pix}} (y_i \ln d_i)\,, \tag{3}$$

where $N_{\rm data} = \sum_{i=1}^{N_{\rm pix}} y_i$. Minimization of the C-statistic with respect to $N_{\rm mod}$ directly yields $N_{\rm mod} = N_{\rm data}$, and we consequently decided to fix $N_{\rm mod}$ and use the statistic

$$E = -2\sum_{i=1}^{N_{\rm pix}} (y_i \ln d_i), \tag{4}$$

which is equivalent to the C-statistic as far as parameter estimation is concerned. This formalism has the advantage of reducing the parameter space of the fit by one dimension (the overall normalization). However, it should be noted that the normalization term $2\,(N_{\rm mod} - N_{\rm data} \ln N_{\rm mod})$ that we have cancelled for the fit still impacts on the error budget, and has to be reintroduced while computing confidence ranges.

Here, we stress that, despite the common terminology, the $C$ (and $E$) statistics are not likelihood functions (which have the dimension of a probability or probability density), but are actually related to the true likelihood $\mathcal{L}$ by

$$C = -2 \times \log \mathcal{L} + B, \tag{5}$$

where $B$ is a constant.

As for the C-statistic, the increase of $E$ between its best-fitting value ($E_{\rm B.F.}$) and a model containing only background (i.e. uniform distribution of the photons), which is often improperly referred to as 'detection likelihood', quantifies the significance of a detection and is $\chi^2$ distributed in the limit of large number of counts (see Cash 1979). From now on, we refer to this parameter as the *detection statistic*:

$$\mathrm{DET\_STAT} = 2N_{\rm data} \ln(N_{\rm pix}) - E_{\rm B.F.}. \tag{6}$$

Similarly, the significance of the estimated extension, can be assessed using an *extension statistic* (improperly referred to as 'extension likelihood') which compares the value of $E$ for the best-fitting point-like and extended source models (once again $\chi^2$ distributed in the limit of large number of counts):

$$\mathrm{EXT\_STAT} = (E_{\rm B.F.})_{\rm point} - (E_{\rm B.F.})_{\rm extended}. \tag{7}$$

The interpretation of these statistics in terms of a detection/extension probability using the $\chi^2$ limit depends on the number of fitted parameters. All the statistics are thus ultimately converted into equivalent values that would correspond to a fit with two free parameters yielding the same probability. This provides a unique and well-defined link between our statistics and probability: for any statistic $S$, $P = \exp(-S/2)$.

*2.3.2 Source processing*

For each source, a fitting box is extracted, the size of which depends on the SExtractor inputs [start with three times the estimated full width at half-maximum (FWHM), with the added requirement to be always at least 35 arcsec]. The SExtractor pixel segmentation mask is used to flag out pixels belonging to neighbouring sources included in the box. This method works well both for source characterization and classification in our shallow exposures,[4] but one would ideally have to implement a simultaneous fit of blended sources in very crowded fields (in development).

The source models take into account all significant *XMM* instrumental effects: an image of the source emission profile is constructed[5] and normalized to the tested count rate;[6] this image is then multiplied by the exposure maps (taking into account vignetting, detection mask, quantum efficiency and the azimuthal sensitivity variations due to the anisotropic transmission from the Rating Grate Arrays) and a uniform background is added, whose level is set so as to match our normalization requirement ($N_{\rm mod} = N_{\rm data}$). Given the faint sources that we are analysing, this very simple background model is justified in the absence of small-scale variations of the *XMM* background in the soft bands. While the EPIC-PN detector is considered as an independent instrument, both the EPIC-MOS detectors are assumed to provide the same count rate for the source and are thus modelled as a single detector using the summed photon image and exposure map.

Starting from the SExtractor outputs as a first guess, the statistic $E$ is minimized using the simplex method amoeba (Press et al. 1992), for both the point source and extended emission models. It takes some 10 min for Xamin to process the average 120 detections per pointing found by SExtractor. The procedure output catalogue comprises 29 derived parameters in addition to the nine free parameters of the fits (four for the point-source model and five for the extended profile). These are listed in Table 2.

## 3 PERFORMANCE EVALUATION THROUGH MONTE CARLO SIMULATIONS

### 3.1 Description of the simulations

To assess the quality of our data analysis, we performed extensive Monte Carlo simulations of $10^4$ s *XMM* pointings with the software InstSimulation (Valtchanov et al. 2001). This procedure creates images from a source list taking into account the main instrumental characteristics (PSF, vignetting, detector masks, background, Poisson noise). In the following, cluster searches (in simulations as well as real pointings) are performed in the [0.5–2] keV band, and stated count rates or fluxes always refer to this band. Galaxy cluster emission is indeed barely detectable at higher energies in our low-exposure pointings, because of the combined effect of the redshifted bremsstrahlung exponential cut-off, the *XMM* drop in sensitivity and strong particle background above 2 keV.

The PSF of the simulations is obtained from the *XMM* calibration files medium model, while the azimuthally averaged off-axis dependency at 1 keV is used to model the vignetting. When simulated, the particle and photon background levels were taken from Read & Ponman (2003). In order to convert source fluxes to count rates, we assumed a constant EPIC-PN to EPIC-MOS count rate ratio regardless of the source spectrum. Note that in the following we will always refer to count rates as the sum of MOS1, MOS2 and PN rates after vignetting correction.[7] This means that for the same count rate, a source is more easily detected near the centre than on the border of the field of view (FOV).

**Table 3.** List of cluster simulations. For each cluster core radius, the number of simulated pointings performed for each count rate ($N_{\rm point}$) is given, as well as the number of simulated sources per pointing ($N_{\rm Src}$) in the central 10 arcmin.

| Radius (arcsec) | Count rate | $N_{\rm point}$ | $N_{\rm Src}$. |
|---|---|---|---|
| 10 | 0.005 | 10 | 8 |
| | 0.01 | 10 | 8 |
| | 0.02 | 10 | 8 |
| | 0.05 | 10 | 8 |
| | 0.1 | 10 | 8 |
| 20 | 0.005 | 10 | 8 |
| | 0.01 | 10 | 8 |
| | 0.02 | 10 | 8 |
| | 0.05 | 10 | 8 |
| | 0.1 | 10 | 8 |
| 50 | 0.01 | 15 | 6 |
| | 0.02 | 15 | 6 |
| | 0.05 | 15 | 6 |
| | 0.1 | 15 | 6 |
| 100 | 0.02 | 30 | 4 |
| | 0.05 | 30 | 4 |
| | 0.1 | 30 | 4 |

Four kinds of simulations were performed.

(i) 30 pointings of $10^4$ s containing only point sources. The flux distribution and source density is computed using the Log ($N$) − Log ($S$) from Moretti et al. (2003) down to $5 \times 10^{-16}$ erg s$^{-1}$ cm$^{-2}$. The background values from Read & Ponman (2003) were accordingly corrected for the contribution of point sources fainter than $4 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ (approximative flux limit of their analysis). We assumed a random spatial distribution of the sources (therefore neglecting the known angular correlation among AGNs).

(ii) 250 pointings of $10^4$ s containing extended sources only ($\beta$-model with fixed $\beta = 2/3$) with simulated background. We simulated core radii of [10, 20, 50, 100] arcsec with count rates in the range [0.005, 0.01, 0.02, 0.05, 0.1] count s$^{-1}$ (see Table 3 for the exact list). Spatial distribution of the sources was set at random so as to cover most of the area within 10 arcmin, with the extra requirement that sources do not overlap.

(iii) 250 simulations with the same extended sources as previously, but injected into a real *XMM–Newton* $10^4$ s pointing pertaining to the *XMM*-LSS (*XMM* id: 0037980501), in order to estimate how extended source characterization is affected by the point-source population.

[4]We detect some 0.1 source per arcmin$^2$ for a PSF FWHM of 6 arcsec.

[5]For both point-source and extended source profiles, we use the medium PSF model from the *XMM* calibration data which is the only one that reproduces the strong distortions of the PSF at large off-axis angles.

[6]For extended sources, this count rate is actually required to match the integral of the profile to infinity, as a significant amount of the source flux can fall outside the fitting box.

[7]In our $10^4$ s pointings, $10^{-2}$ count s$^{-1}$ roughly corresponds on the optical axis to 100-count spread over the three EPIC detectors and a flux of about $9 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ for both an active galactic nucleus (AGN) spectrum (a power-law spectral energy distribution with spectral index $\Gamma = 1.7$) and a local 2-keV cluster (thermal bremsstrahlung).

**Table 4.** Source selection criteria with the *XMM*-LSS pipeline.

| Classification | Criteria |
|---|---|
| Class 1 extended | Detection statistic>32, Extension statistic>33, Extension>5 arcsec |
| Class 2 extended | Extension statistic>15, Extension>5 arcsec |
| Point source | Neither C1 nor C2, Detection statistic>15 |

**Table 5.** Contamination statistics predicted from the simulations for each *XMM*-LSS pipeline source sample.

| Real source type | Classification | $N_{\rm Src}$/Pointing |
|---|---|---|
| Point like | Class 1 | 0.0 |
| Point like | Class 2 | 0.17 |
| Spurious | Class 1 | 0.0 |
| Spurious | Class 2 | 0.10 |
| Spurious | Point like | 0.53 |

(iv) 18 simulations containing close pairs of point-like sources (separated by 20 arcsec) injected into a typical real *XMM–Newton* pointing to test the deblending capabilities of the pipeline. In the first nine simulations, 10 pairs of $3 \times 10^{-3}$ count s$^{-1}$ were added in the pointing, while in the remaining ones, five pairs of $5 \times 10^{-2}$ count s$^{-1}$ were simulated. These simulations are also relevant for cluster false detection rate as blended point sources may be characterized as 'extended' sources.

Examples of simulated images are given in Figs 1–3.

All simulated images were analysed through steps (ii) and (iii) of our pipeline (see Section 2). Detected sources were then cross-identified with the simulation inputs using a correlation radius of 5 pixel for point sources and 15 pixel for extended ones. In the following sections, we will refer to *spurious* detections as those that could not be cross-identified with any input source.

### 3.2 Parameter estimation accuracy

#### *3.2.1 Extended sources*

Our simulations demonstrate that the mean photometry of extended sources with XAMIN is satisfactory in both pointings with or without point sources; it is unbiased for bright sources with a mean dispersion of about 20 per cent (see Fig. 4), while unavoidable Eddington bias and scatter increase appear for fainter ones. Count rates seem somewhat overestimated only for very faint or very extended sources. The scatter increases slowly as count rates decrease but also with increasing radius.

Even when the clusters are injected into real pointings, the performances remain correct up to 50-arcsec core radius. Knowing that, for a physical core radius of $180\,h_{70}^{-1}$ kpc, the apparent core radii span the range 55–22 arcsec for $0.2 < z < 1$ in Lambda cold dark matter ($\Lambda$CDM) cosmology, the goals of the pipeline are fully met. Above 50-arcsec core radii, the rates are somewhat overestimated. In addition, we note surprisingly a weak increase of the detection efficiency of these sources when adding AGNs. The simplest interpretation is that part of the very extended sources found their emission contaminated by faint AGNs (that fall below our detection

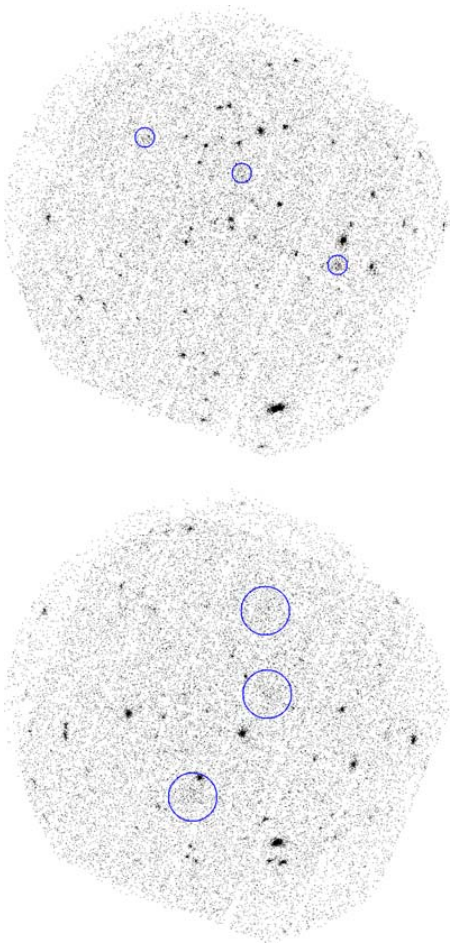

**Figure 1.** Examples of simulated $10^4$ s *XMM–Newton* images (co-addition of the EPIC cameras). Both contain point sources distributed following the X-ray Log($N$) − Log($S$). Blue circles show the position of simulated clusters. Top: clusters have core radii of 20 arcsec and on-axis count rates of (from top to bottom) 0.02, 0.01 and 0.03 count s$^{-1}$. Bottom: clusters have core radii of 50 arcsec and on-axis count rates of (from top to bottom) 0.03, 0.02 and 0.05 count s$^{-1}$. Displayed clusters are very faint (close to the detection limit) so as to illustrate the detection capabilities of the pipeline.

or deblending capacity), and thus tend to pass more easily the detection criteria of the pipeline, but with erroneous photometry and core radius.

A second point to note comparing the left- and right-hand panels of Fig. 4 is that the photometry seems correlated with extension measure accuracy. A poor modelling of the source

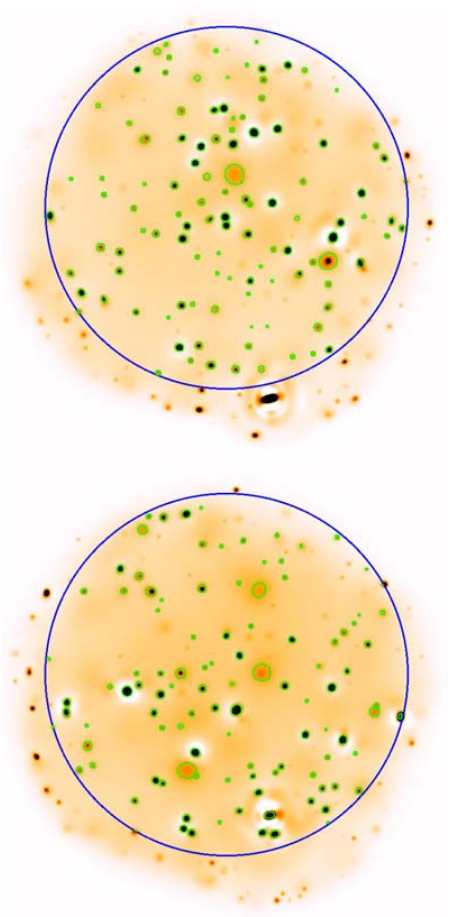

**Figure 2.** Wavelet images of Fig. 1 simulations, overlaid with SExtractor catalogues. The blue circle shows the central 13-arcmin radius of the FOV (centred on the mean optical axis) where SExtractor detections are performed.

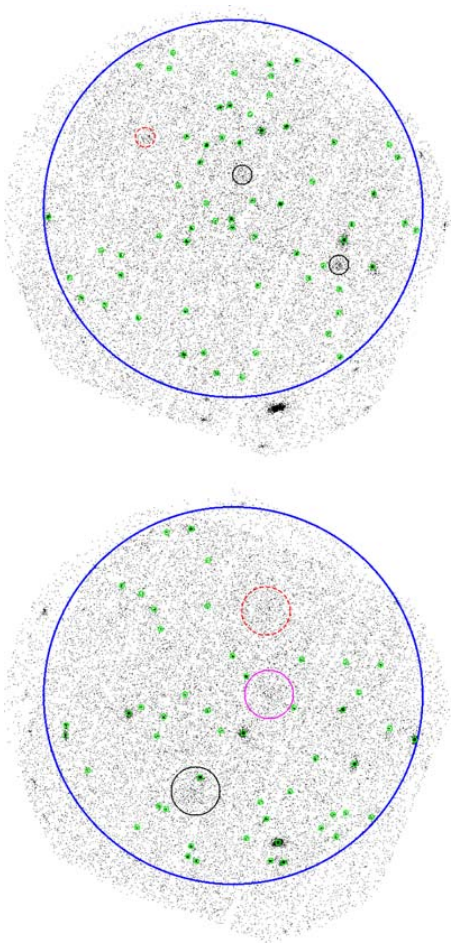

**Figure 3.** Raw images of Fig. 1 simulations, overlaid with Xamin catalogues. The green 10-arcsec radius circles show the detected point sources (see Fig. 4 for the selection criteria); black and magenta circles show the C1 and C2 clusters, respectively. Clusters not flagged by Xamin as C1 or C2 are indicated by red dashed line circles.

emission profile logically yields incorrect count rate estimates, particularly for very extended sources where the flux is extrapolated far outside of the fitting box. An inaccurate estimate of the source extensions is thus probably the reason for both the low count-rate and the high core radius photometry bias identified above.

### 3.2.2 *Point sources*

As shown in Fig. 5(b), the point-source photometric dispersion is basically comparable to the spread due to Poisson noise down to 20 count. At the faint tail of the distribution, a strong Eddington bias appears.

Another issue regards point-source confusion. We used our set of close pair simulations to test the deblending efficiency. The results are quite satisfactory: all $5 \times 10^{-2}$ count s$^{-1}$ pairs are deblended, while more than 65 per cent of the $3 \times 10^{-3}$ count s$^{-1}$ sources are also. This success rate cannot easily be reached by first step detection procedures based on sliding cells having a minimum size of 10 arcsec. This point is important not only for point-source

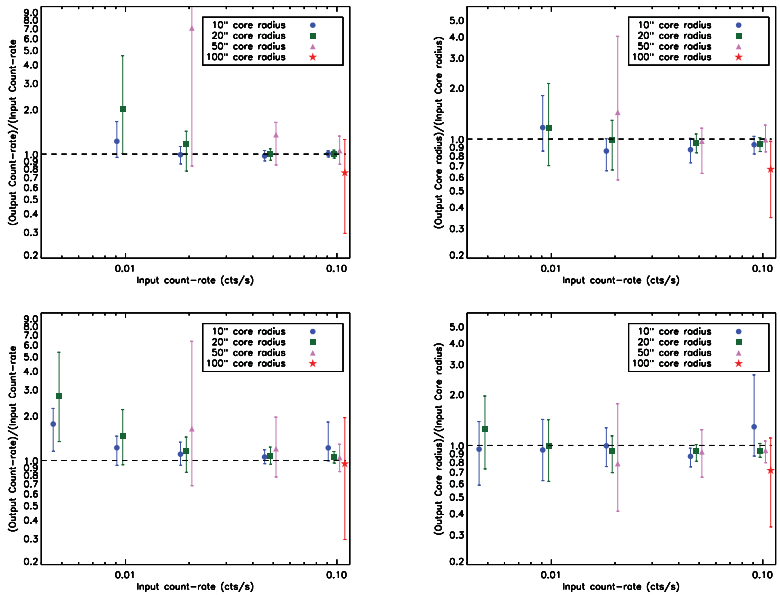


**Figure 4.** Extended source characterization with the *XMM*-LSS pipeline within 10 arcmin of the FOV for $10^4$ s exposures. For all plots, vertical bars show the standard deviation of measured points. Upper plots show the results for clusters in pointings without point sources, lower plots show the results when clusters are injected into real pointings. Left-hand panel: photometry as a function of on-axis counts and core radius. Right-hand panel: source extension measure as a function of on-axis counts and core radius. In all the plots, only the bins encompassing at least 10 recovered sources are shown.

statistics, but also for cluster detection, in order not to consider blended point sources as a single extended one.

### 3.3 Source classification using the simulations

Source selection and estimation of the selection function in surveys is always a complicated task and results from the necessary trade-off between sample completeness and contamination. For this purpose, we explored the XAMIN output parameter space by means of our simulations in order to set well-controlled extended/point-like source selection criteria, and to estimate contamination by spurious or misclassified sources.

#### *3.3.1 Point sources*

As AGNs represent more than 90 per cent of the extragalactic X-ray sources at our sensitivity, we restrict ourselves to the estimation of

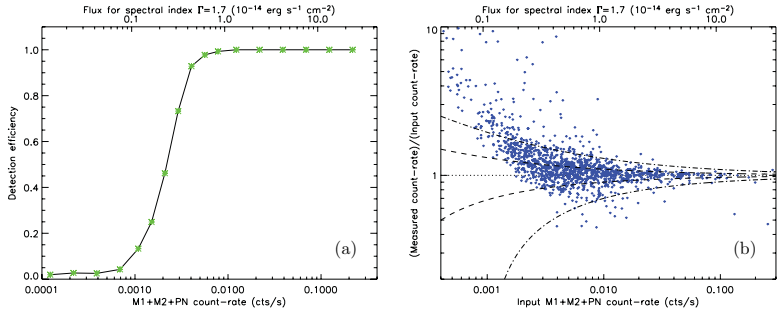


**Figure 5.** Point-source analysis with the *XMM*-LSS pipeline within 10 arcmin of the FOV for $10^4$ s exposures. (a) Detection probability as a function of count rate. (b) Photometry, dashed and dot–dashed lines show, respectively, the intrinsic $1\sigma$ and $3\sigma$ scatter expected from Poisson noise.

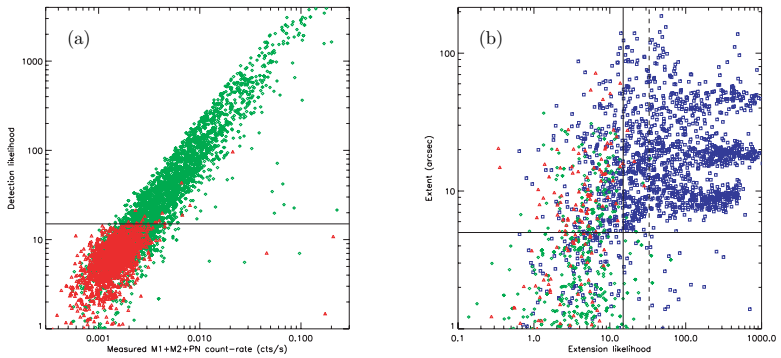


**Figure 6.** Determination of the *XMM*-LSS pipeline selection criteria. AGNs are displayed as green diamonds, galaxy clusters as blue squares. Red triangles stand for spurious detections. (a) Selection of point sources in the count rate – detection likelihood plane; the solid line at Likelihood = 15 defines the point-source sample. (b) Cluster selection in the Extent – extension Likelihood plane; the solid lines at Extent = 5 arcsec and Likelihood = 15 define the C2 sample; the dashed line shows the extension likelihood criteria of the C1 sample.

the spurious detection rate based on our point-source simulations. As can be seen in Fig. 6(a), a simple threshold of 15 in the detection statistic gives the best balance between contamination and completeness: at this threshold, some 40–50 real point-like sources are detected in each pointing within 10 arcmin of the FOV, for only 0.5 spurious ones.

The resulting detection efficiency as a function of count rate is shown in Fig. 5(a). The point-source flux limit (90 per cent completeness) is about $4 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ in [0.5–2] keV, but more than 50 per cent of the sources are detected down to $\sim 2.5 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$.

### *3.3.2 Extended sources*

Source selection is complicated for extended sources because these objects are generally of lower surface brightness (see e.g. Fig. 1), and one does not only have to deal with spurious detections, but also with contaminating misclassified point sources. This task requires special care, keeping in mind the very cosmological applications of the survey. Fig. 7 shows the fraction of extended sources that are detected by SExtractor in the primary catalogue as a function of flux and core radius. Our purpose is then to find a location in the Xamin output parameter space where the majority of these sources are recovered while keeping the contamination rate to a manageable level.

As a first step, we scanned the detection/extension statistic-extension space for the largest uncontaminated extended source sample. This is obtained for EXT > 5 arcsec, EXT_STAT > 33 and extended fit DET_STAT > 32 simultaneously[8] (see Table 2 for the definition of these parameters). From now on, we will refer to this sample as class 1 (C1) extended sources. Fig. 6(a) illustrates the main C1 selection process in the extension – extension statistic plane.

[8]Note that from the definition of EXT_STAT (see Section 2.3.1 and Table 2), it is very unlikely for DET_STAT to be lower than 32 if EXT_STAT is greater than 33, except in the few rare cases where the point-source fit crashed.

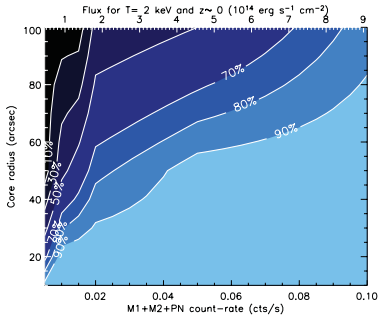


**Figure 7.** Detection probability for extended sources by SExtractor; this can be considered as the ultimate sensitivity with $10^4$ s *XMM* images, but the contamination is maximal.

Due to our non-contamination requirement, the C1 sample naturally excludes a number of extended sources (generally very low surface brightness or more compact sources). A less conservative sample (required by the *XMM*-LSS cluster search in order to detect as many valid sources as possible) can be obtained by relaxing the previous criteria to EXT > 5 arcsec, EXT_STAT > 15 and no DET_STAT constraint (see Fig. 6). From the number of detections matching these criteria in our point-source simulations, we can estimate that this class 2 sample (C2) contains less than one spurious detection or misclassified point source every three pointings.

The mean detection probabilities of extended sources within 10 arcmin of the FOV are presented for both C1 and C2 samples in Fig. 8, as a function of count rate and apparent core radius. As expected, this probability is higher within the C2 sample for low surface brightness and faint compact sources.

Note that detection efficiency is not a simple function of source flux as is sometimes assumed in X-ray cluster surveys (see

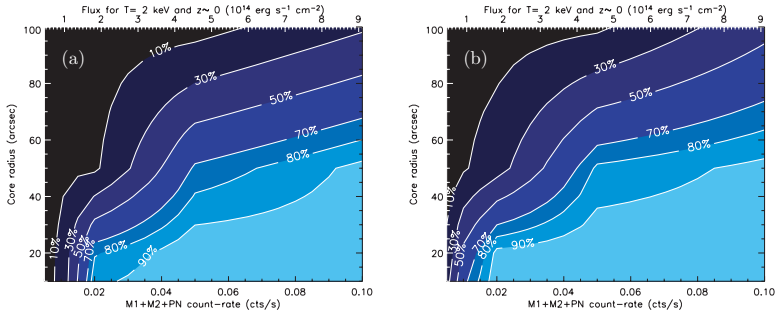


**Figure 8.** Extended source detection efficiency of the *XMM*-LSS pipeline in $10^4$ s exposures as a function of source counts and core radius inside 10 arcmin of the FOV. (a) C1 sample. (b) C2 sample.

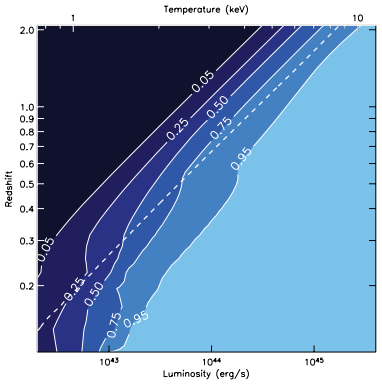


**Figure 9.** Probability of detecting a cluster located inside the central 10 arcmin of the FOV as a C2 source as a function of its redshift and luminosity given the cosmological model presented in Section 4.2 within $10^4$ s pointings. An indicative flux limit of $2 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ is shown by the thick dashed line.

e.g. Rosati et al. 1998), but it varies significantly when considering different source sizes, and this should be modelled to interpret correctly the results of the *XMM*-LSS. This impact of source extent on our detection capacity is illustrated by Fig. 9, where the detection probability as a function of luminosity and redshift is shown for the C2 sample, assuming a canonical core radius of $180\,h_{70}^{-1}$ kpc. At high redshift, where the angular distance is almost constant, our selection process closely resembles a flux limit, while the sensitivity drops at lower redshift. In this model, we find that roughly 90 per cent of the sources down to 3 keV are detected in C2 at $z = 0.5$. This number falls to 50 per cent at a redshift of 0.9–1. A cluster similar to Coma (∼8 keV) would always be detected at least as C2 up to a redshift of 1, and have more than 75 per cent probability of being detected at $z = 2$. The definition of the source classes and their contamination statistics are summarized in Tables 4 and 5 respectively.

### 3.4 Validation on real data

To further validate our selection criteria, we processed all available *XMM*-LSS pointings and compared the pipeline output with our simulation results. Our X-ray data currently consist of 51 *XMM–Newton* pointings; 19 of them (G pointings) were obtained from guaranteed-time observation as part of a joint Liège/Milan/Saclay program (XMDS; Chiappetti et al. 2005) and have $2 \times 10^4$ s exposure; the remaining 32 are $10^4$ s long and were obtained with guest-observer time. Among these, three pointings (one G and two B) are unusable due to very high background levels (probable solar flare contamination).

#### *3.4.1 Point sources*

In our 30 $10^4$ s exposure pointings, we obtain on average 45.8 sources per pointing with DET_STAT >15 for the point-source fit. As a comparison, taking into account the detection probabilities of Fig. 5 and integrating over the log $(N) - \log(S)$ of Moretti et al. (2003) between $5 \times 10^{-16}$ and $1 \times 10^{-11}$ erg s$^{-1}$ cm$^{-2}$ yields on average 47.4 sources per pointing. Though the matching is already satisfactory, the remaining difference mostly reflects the lack of very bright sources in the *XMM*-LSS area (probably due to cosmic variance) identified by Gandhi et al. (2006).

We additionally cross-identified our sources with those of the XMDS/VIMOS VLT Deep Survey $4\sigma$ catalogue (Chiappetti et al. 2005) which results from an alternate analysis of the G fields (that mainly uses the standard *XMM*-SAS procedures and is thus suitable only for point sources). We found a very good agreement with both detected sources and their characteristics. XAMIN count rates are always within the error bars of *XMM*-SAS EMLDETECT measurements for sources that do not fall on CCD gaps. Moreover, our detection statistic values are tightly correlated to their detection probability estimates.

#### *3.4.2 Extended sources*

Until now, the *XMM*-LSS optical spectroscopy follow-up program (see Pierre et al. 2004) enabled us to confirm about 60 cluster candidates. This allowed crosschecking the definition of our selection criteria obtained from the simulations against real data.

As regards the C1 sample, only genuine extended X-ray sources are detected, as expected, with no additional contamination. The majority (≈85 per cent) of the extended C1 sources are clusters, the remainder being nearby galaxies. A little contamination, around 0.5 false detections per pointing, is observed in the C2 sample (in the present data sets, this amounts to about 50 per cent of the newly detected sources, once the nearby galaxies and C1 clusters have been excluded). This false detection rate, while still acceptable for a survey with optical follow-up, is slightly higher than our estimates from simulations, and this is probably the result of neglecting the AGN correlation function (thus lowering the number of non-deblended close pairs of AGNs). Another possibility is that we are detecting some AGNs that are included in cosmological filaments with weak X-ray emission, which was not accounted for in the simulations.

#### *3.4.3 Example runs on $z > 1$ clusters*

We ran the pipeline on the archival *XMM–Newton* observation 0111790101, for which detection of the highest redshift X-ray cluster to date, XMMUJ2235.3-2557, at $z \sim 1.4$ was reported (Mullis et al. 2005). The observation was performed in the medium filter and EPIC-PN small window mode so that the source, located 7.7 arcmin from the optical axis, is only observed in the EPIC-MOS detectors. Using the full $4.5 \times 10^4$ s exposure of the pointing ($\approx 3 \times 10^4$ s at the source position), the cluster is easily identified as a C1 extended source. We further simulated the *XMM*-LSS observing conditions by analysing only the first $10^4$ s of the observation. XMMUJ2235.3-2557 is still detected with extended fit DET_STAT = 93.8, EXT_STAT = 31.1 and EXT = 9.8, as a C2 extended source, at the limit of the C1 parameter space, and would therefore have been detected as C1 in the exact *XMM*-LSS observing conditions (i.e. using thin filter and with EPIC-PN data available). The ease with which this high-redshift cluster is detected is mainly due to its apparent brightness: ∼220 (respectively, 70) photons were available in the $4.5 \times 10^4$ s ($10^4$ s) exposure. For comparison, we note that the $z = 1.22$ cluster XLSSJ022302.6-043621 detected by Bremer et al. (2006) in the *XMM*-LSS survey is classified as a C2 source (EXT = 5.4, EXT_STAT = 15.4 and DET_STAT = 51.4) with only 58 photons available for the fit.

## 4 THE *XMM*-LSS SELECTION FUNCTION

Our simulation programs provide us with tools to compute the *XMM*-LSS selection function. We can derive the detection probability as a function of source characteristic for any exposure time, background level and position on the detector.

Fig. 5 shows the point-source detection probability inside a radius of 10 arcmin from the mean optical axis as a function of flux. From this, a direct estimate of our mean sky coverage can be obtained.

For a given cosmology, a galaxy cluster of given luminosity, temperature, physical extent and redshift can be described by an angular core radius and a detected *XMM* count rate, for which Fig. 8 gives the detection probability for C1 and C2. We are therefore now able to properly describe our galaxy cluster selection process.

As an illustration, we compute below the expected redshift distribution of C1 and C2 clusters in ΛCDM cosmology.

### 4.1 Cosmological model

In the following, the cosmological parameters that determine the dynamics and content of the Universe are set to *Wilkinson Microwave Anisotropy Probe* (*WMAP*) values (Spergel et al. 2003), namely $H_0 = 71$ km s$^{-1}$ Mpc$^{-1}$, $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, $\Omega_b = 0.044$, $n = 0.93$ and $\sigma_8 = 0.84$.

#### *4.1.1 Mass function*

The shape of the linear power spectrum $P(k)$ is computed at $z = 0$ using the initial power-law dependency in $k^n$ and the transfer function from Bardeen et al. (1986). The influence of baryons on the transfer function was modelled using the shape parameter

$$\Gamma = \Omega_m h \exp[-\Omega_b(1 + \sqrt{2h}/\Omega_m)] \quad (8)$$

introduced by Sugiyama (1995), and the overall spectrum is normalized to $\sigma_8$.

Then at each redshift value on a fine grid: $P(k)$ is evaluated from its $z = 0$ value using the linear growth factor from Carroll, Press & Turner (1992) and $\sigma(M)$ is deduced. The comoving halo number density as a function of mass, $dn/dm(z)$, is computed using the Sheth & Tormen (1999) mass function.

This common procedure to determine the halo mass function has been largely tested on numerical simulations and is known to provide accurate predictions as long as one defines the mass of the haloes to be the one included inside $r_{200b}$, the radius that encloses an overdensity of 200 with respect to the mean background density.

#### *4.1.2 Applying the selection function*

Knowledge of the cluster scaling relations is needed to predict the temperature and luminosity of these haloes and compute *XMM* count rates. Unfortunately, one generally does not have access to the mass in $r_{200b}$ from the X-ray data, and a halo profile model is required in order to convert the mass function to another mass definition.

For this purpose, we used Navarro–Frenk–White profiles (Navarro, Frenk & White 1997) with scaling radius $r_s$ provided by the model of Bullock et al. (2001) which relates $r_s$ to the virial mass of the halo through the concentration parameter $c = r_{vir}/r_s$.[9] The conversion itself is performed using the formulae provided by the appendix of Hu & Kravtsov (2003).

The emission-weighted gas temperature is derived using the local $M_{200}$ –$T$ relation of Arnaud, Pointecouteau & Pratt (2005), i.e. a slope of $\alpha = 1.49$, valid for clusters with $T > 4$ keV. At lower temperatures, we added a gradual steepening of the correlation ($\alpha = 1.85$ below 4 keV and $\alpha = 2$ below 2 keV) as indicated by several recent works (see e.g. Finoguenov, Reiprich & Böhringer 2001). No evolution of the $M_{200}$ –$T$ relation with redshift was supposed. As an arbitrary condition to be considered as a group or cluster, we subsequently removed all haloes with $T < 1$ keV.

Bolometric luminosities are then computed using the $L_X$–$T$ relation of Arnaud & Evrard (1999) with no evolution. Though there is some evidence that the local $L_X$–$T$ relation also steepens at low $T$, this seems to be important only for $T \leqslant 1$ keV (Osmond & Ponman 2004) and is consequently ignored.

The total *XMM–Newton* EPIC count rate is estimated using an APEC[10] thermal plasma emission model (Smith et al. 2001) with neutral hydrogen absorption as modelled by Morrison & McCammon (1983) using fixed column density of $2.6 \times 10^{20}$ cm$^{-2}$ (representative of our field) folded through the EPIC response matrices for the thin filter in accordance with our observing mode.

[9] Note that we also tested the model of Eke, Navarro & Steinmetz (2001) and found a change in the redshift distribution of our C1/C2 samples lower than 10 per cent.

[10] http://cxc.harvard.edu/atomdb/

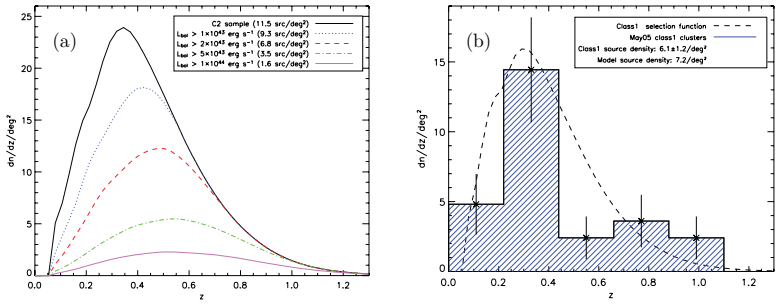


**Figure 10.** Cosmological expectations of the C1 and C2 samples for sources with $T > 1$ keV. (a) Luminosity and redshift distribution of the C2 sample. (b) Redshift distribution of the observed C1 sources (29 sources, 24 with redshifts) compared to the ΛCDM expectations.

The selection function is finally applied assuming a constant physical core radius of $180h_{70}^{-1}$ kpc.

### 4.2 Results

Using this simplified model and the selection functions obtained from the simulations, we find the following.

(i) The C2 sample should contain roughly 12 clusters per deg$^2$. When the *XMM*-LSS is complete, it will thus constitute the deepest X-ray selected galaxy cluster sample over a wide area.

(ii) The C1 sample should contain some 7 clusters per deg$^2$. While this source density is a bit lower than for C2, this selection process can be applied to the whole of the *XMM* archive, regardless of expensive and time consuming optical spectroscopy follow-up, as the sample is effectively uncontaminated.

The expected redshift distribution for both samples is shown in Fig. 10. Panel (a) of the figure also gives an idea of the luminosity distribution of the C2 sample.

To validate these results, we compared them with the redshift distribution of the observed C1 clusters. The sample contains 29 sources of which 24 have already been spectroscopically confirmed. Assuming that the five missing sources (∼17 per cent of the sample) will not alter significantly the current distribution, we find very good overall agreement with our prediction (Fig. 10b).

A further interesting result, already outlined in Section 3.3.2 and Fig. 8, is that our selection process does not reproduce a flux limited sample, especially at $z < 0.6$ where the change in angular distance is significant (see Fig. 11a).

This point is further illustrated in Fig. 11(b), where we investigate our detection efficiency as a function of cluster flux. This shows, for the assumed cluster population, the a priori impossibility of constructing a flux-limited sample from our primary catalogues, even accepting a substantial contamination level, unless a very high flux limit is set. In the present study, SExtractor is run on optimally filtered images (retaining only significant structures above $3\sigma$) with a very sensitive detection threshold and our results suggest that we have reached the limit of the data. This therefore challenges any further attempt aiming at defining deep flux limited samples with *XMM*.

### 4.3 Limitations of the present model

Although the matching between this simple model and our data (as shown in Fig. 10b) is impressive, one should keep in mind that some

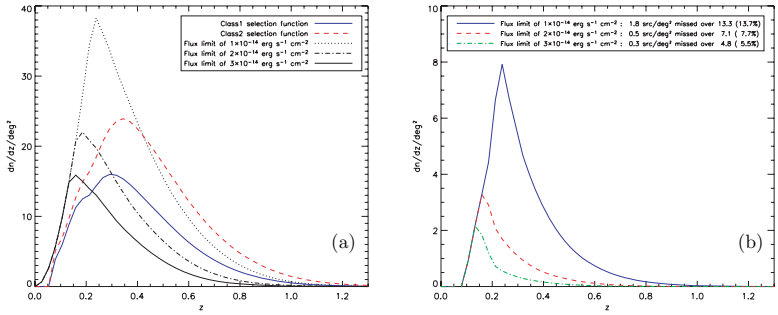


**Figure 11.** Comparison of our source selection process with the common assumption of a flux-limited sample for sources with $T > 1$ keV. (a) Expected $dn/dz$ for class 1 (blue) and class 2 (red) compared to flux-limited surveys; on average, we detect higher redshift clusters than flux-limited surveys with the same source density. (b) Redshift distribution of the sources not detected by the first pass (MR/1 + SExtractor) above several flux limits, assuming the source population generated by our simple cosmological model (we miss the low end of the luminosity function).

ingredients of the model are still uncertain (and this is precisely the purpose of the *XMM*-LSS to try to constrain them).

In particular, while the evolution of the $M_{200}$–$T$ relation is still unknown, there seems to be indication of a positive evolution as predicted by self-similar models (see e.g. Ettori et al. 2004; Maughan et al. 2006). However, none of these studies is probing our range of temperature and redshift, and the influence of non-gravitational processes can well alter this behaviour in the group regime, thus the use of the simplest non-evolving relations.

Also, in order to properly take into account the varying gas distribution with cluster mass, our assumption of a fixed core radius may seem too simple and one would have to consider lower $\beta$ values for the groups as indicated by observations (e.g. Osmond & Ponman 2004). However, such data are generally largely dominated by scatter and there is currently no well-established scaling relation for these global trends.

Finally, a large fraction of the observed scatter on all these scaling relations is intrinsic to the source properties and results from the complex process of hierarchical merging in CDM cosmologies and feedback from non-gravitational activity.

These are a number of caveats that need to be taken into account in the interpretation of such a small sample of low-temperature systems. In a forthcoming paper (Pacaud et al., in preparation), where we will present the full cluster catalogue, we will further discuss the effect of the various cluster scaling laws and evolution schemes on the d$n$/d$z$ using as input our $L$–$T$ relation for groups at redshift around 0.5.

## 5 SUMMARY AND CONCLUSIONS

We have described the procedure that we developed to analyse the $1 \times 10^4$–$2 \times 10^4$ s *XMM* images of the *XMM*-LSS survey. The main motivation of this work is the need for assembling a sample of clusters of galaxies out to a redshift of unity with controlled selection effects, suitable for cosmological and evolutionary studies. The resulting pipeline consequently combines multiresolution wavelet filtering (MR1) to reach the source detection limit, with a subsequent maximum likelihood analysis (XAMIN) to characterize the source properties.

The performances of the adopted procedure have been duly tested by means of extensive image simulations either reproducing all instrumental and astrophysical effects, or injecting extended and point-like sources into already existing pointings. This allowed us to investigate the ultimate capabilities, such as resolving power, cluster detectability and characterization, as a function of flux and apparent size, photometric accuracy. In this respect, our package constitutes a significant improvement over the standard SAS and the XMDS procedure (Chiappetti et al. 2005), especially for the extended source analysis.

Moreover, the XAMIN output parameter space, densely scanned by the simulations, provides a powerful means to interpret the detected sources. In this way, we are able to define two classes of extended sources: the C1 class which is basically uncontaminated by misclassified point-like sources and the C2 class allowing for some 50 per cent contamination. This selection process, derived from the simulations, has been subsequently checked and validated against the current *XMM*-LSS sample of spectroscopically confirmed galaxy clusters.

Finally, considering a canonical power spectrum combined with a simple halo model providing $n(M, z)$ and simple cluster scaling laws ($M$–$T$–$L$) in a ΛCDM cosmology allowed us to predict the d$n$/d$z$ distribution of the C1 cluster population. Comparison with our current C1 data sample shows a very good agreement. From this, we infer that our goal of producing a cluster sample with controlled selection effects is fulfilled at this stage. An important point to be further emphasized is that the resulting sample is not flux limited – a concept that is anyway not rigorously applicable when dealing with extended sources spanning a wide range in flux and size.

The way the C1 class is defined allows us to construct a purely X-ray selected cluster sample with a high number density of $\sim$7 deg$^{-2}$ in the redshift range [0–1.2]. Moreover, an unprecedented density of $\sim$12 deg$^{-2}$ can be obtained with the C2 sample which includes objects down to a flux of $\sim 5 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$. This opens the door to the routine construction of unbiased cluster samples from *XMM* images.

In the very near future, with the compilation of the full *XMM*-LSS cluster sample over the currently existing 5 deg$^2$, we will refine the cosmological modelling of the observed d$n$/d$z$ (Pacaud et al., in preparation). In particular, we will further investigate the effect of varied evolution schemes of the scaling relation, and assumptions on cluster sizes and shapes (including scatter on these average trends). Both aspects are especially relevant for the $T < 2$ keV groups out to $z \sim 0.5$, a population that the *XMM*-LSS is for the first time unveiling and that constitute the bulk of our sample. Noting that the C1 cluster sample is almost identical to the sample for which we can measure a temperature (Pierre et al. 2006), we will also be in a position to constrain the evolution of the $L_X$–$T$ relation.

The combined d$n$/d$z$, $L_X$–$T$ and shape modelling will provide very useful constrains on numerical simulations, the missing link between the theoretical parameter $M$ and the observable $L_X$, and consequently a self-consistent description of the building blocks of the present day clusters.

## ACKNOWLEDGMENTS

This paper was based on observations obtained with *XMM–Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA. The simulations were performed at the CNRS 'Centre de Calcul de l'IN2P3' located in Lyon, France. The authors would like to thank Jean Ballet for sharing with us his deep knowledge of statistics and *XMM–Newton* calibrations, and Pierrick Micout for his help regarding the use of the CC-IN2P3. We are also grateful to Jean-Paul Le Fèvre for developing and monitoring our cluster database (the L3SDB [11]) thus helping us to handle our large data sets, and to Monique Arnaud for useful discussions and advices regarding local X-ray cluster scaling relations.

## REFERENCES

Arnaud M., Evrard A. E., 1999, MNRAS, 305, 631
Arnaud M., Pointecouteau E., Pratt G. W., 2005, A&A, 441, 893
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bremer M. N. et al. 2006, MNRAS, 371, 1427
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Carroll S. M., Press W. H., Turner E. L., 1992, ARA&A, 30, 499
Cash W., 1979, ApJ, 228, 939
Cavaliere A., Fusco-Femiano R., 1976, A&A, 49, 137
Chiappetti L. et al., 2005, A&A, 439, 413
Eke V. R., Navarro J. F., Steinmetz M., 2001, ApJ, 554, 114
Ettori S., Tozzi P., Borgani S., Rosati P., 2004, A&A, 417, 13
Finoguenov A., Reiprich T. H., Böhringer H., 2001, A&A, 368, 749
Gandhi P. et al., 2006, A&A, in press

Hu W., Kravtsov A. V., 2003, ApJ, 584, 702
Infante L., 1987, A&A, 183, 177
Kron R. G., 1980, ApJS, 43, 305
Maughan B. J., Jones L. R., Ebeling H., Scharf C., 2006, MNRAS, 365, 509
Moretti A., Campana S., Lazzati D., Tagliaferri G., 2003, ApJ, 588, 696
Morrison R., McCammon D., 1983, ApJ, 270, 119
Mullis C. R., Rosati P., Lamer G., Böhringer H., Schwope A., Schuecker P., Fassbender R., 2005, ApJ, 623, L85
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Osmond J. P. F., Ponman T. J., 2004, MNRAS, 350, 1511
Pierre M. et al., 2004, J. Cosmol. Astropart. Phys., 9, 11
Pierre M. et al., 2006, MNRAS, this issue (doi:10.1111/j.1365-2966.2006.10886.x)
Pratt G., Arnaud M., 2002, A&A, 394, 375
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes. Cambridge Univ. Press, Cambridge
Read A. M., Ponman T. J., 2003, A&A, 409, 395
Rosati P., della Ceca R., Norman C., Giacconi R., 1998, ApJ, 492, L21
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Smith R. K., Brickhouse N. S., Liedahl D. A., Raymond J. C., 2001, ApJ, 556, L91
Spergel D. N. et al., 2003, ApJS, 148, 175
Starck J.-L., Pierre M., 1998, A&AS, 128, 397
Starck J.-L., Murtagh F., Bijaoui A., 1998, Image Processing and Data Analysis: The Multiscale Approach. Cambridge Univ. Press, Cambridge (MR/1)
Sugiyama N., 1995, ApJS, 100, 281
Valtchanov I., Pierre M., Gastaud R., 2001, A&A, 370, 689
Vikhlinin A., McNamara B. R., Forman W., Jones C., Quintana H., Hornstrup A., 1998, ApJ, 502, 558

This paper has been typeset from a TEX/LATEX file prepared by the author.

## 7.3 Premier sous-échantillon complet :

## "The XMM LSS survey : a well-controlled X-ray cluster sample over the D1 CFHTLS area"

### Contexte/motivation

Dans la continuité du travail réalisé pour Willis et al. (2005), j'ai développé des méthodes systématiques pour la mesure des températures et de la luminosité des amas XMM-LSS. Dès que le pipeline fut défini et validé, la chaîne complète de traitement (depuis la détection jusqu'à l'analyse des sources) fut donc testée sur le champ de 1 $\deg^2$ qui comportait les données les plus exhaustives et profondes.

### Résumé

L'article présente le catalogue d'amas du relevé XMM-LSS correspondant à la region D1 du CFHTLS. Cette liste contient 13 amas sélectionnés en X et confirmés spectroscopiquement sur 0.8 $\deg^2$ jusquà un redshift de 1. Elle constituait (à la date de parution de l'article) l'échantillon d'amas X de plus forte densité. La luminosité bolométrique des sources couvre un intervalle de 0.03 à $5 \times 10^{44}$ erg.s$^{-1}$. La méthode de construction du catalogue y est détaillée : depuis la détection en rayon X des candidats amas jusqu'à l'assemblage d'un échantillon d'amas confirmés spectroscopiquement avec une fonction de sélection précise. Cette procédure fournit des paramètres X basiques comme la température, le flux et la luminosité. Nous détectons légèrement plus d'amas avec des flux (0.5-2 keV) $> 2 \times 10^{-14}$ erg.s$^{-1}$.cm$^{-2}$qu'attendu d'après les prévisions des relevés profonds ROSAT. Nous présentons également la relation luminosité-température pour nos 9 objets les plus brillants qui possèdent une mesure de température crédible. La pente est en bon accord avec la relation locale. Les données semblent compatible avec une augmentation de luminosité pour les objets entre $0.15 < z < 0.35$ ayant des temperatures $1 < T < 2$ keV, une population que le XMM-LSS identifie pour la première fois de manière systématique. La présente étude permet la constitution d'échantillons d'amas à partir d'images XMM dont les effets de sélecion sont compris. Cela légitime, en plus de l'étude des structures à grande échelle, l'examen systématique de l'évolution des relations d'échelle des amas, particulièrement pour les groupes X de faible masse qui constituent la majorité de la population observée.

### Contribution personnelle

Outre la détection des sources (C1,C2) et une participation à la sélection des candidats C3, j'ai mesuré l'ensemble des propriétés X des sources ($F_X$, $L_X$, T) et généré les figures correspondantes.

Mon. Not. R. Astron. Soc. **372,** 591–608 (2006) doi:10.1111/j.1365-2966.2006.10886.x

# The *XMM* Large-Scale Structure survey: a well-controlled X-ray cluster sample over the D1 CFHTLS area⋆

M. Pierre,[1]† F. Pacaud,[1] P.-A. Duc,[1] J. P. Willis,[2] S. Andreon,[3] I. Valtchanov,[4]‡ B. Altieri,[5] G. Galaz,[6] A. Gueguen,[1] J.-P. Le Fèvre,[7] O. Le Fèvre,[8] T. Ponman,[9] P.-G. Sprimont,[10] J. Surdej,[10] C. Adami,[8] A. Alshino,[9] M. Bremer,[11] L. Chiappetti,[12] A. Detal,[10] O. Garcet,[10] E. Gosset,[10] C. Jean,[10] D. Maccagni,[12] C. Marinoni,[8] A. Mazure,[8] H. Quintana[6] and A. Read[13]

[1]*DAPNIA/SAp CEA Saclay, 91191 Gif sur Yvette*
[2]*Department of Physics and Astronomy, University of Victoria, Elliot Building, 3800 Finnerty Road, Victoria, BC, Canada V8P 1A1*
[3]*INAF – Osservatorio Astronomico di Brera, Milan, Italy*
[4]*Astrophysics Group, Blackett Laboratory, Imperial College of Science Technology and Medicine, London SW7 2BW*
[5]*ESA, Villafranca del Castillo, Spain*
[6]*Departamento de Astronomía y Astrofísica, Pontificia Universidad Católica de Chile, Santiago, Chile*
[7]*DAPNIA/SEDI CEA Saclay, 91191 Gif sur Yvette*
[8]*Laboratoire d'Astrophysique de Marseille, France*
[9]*School of Physics and Astronomy, University of Birmingham, Edgebaston, Birmingham B15 2TT*
[10]*Université de Liège, Allée du 6 Août, 17, B5C, 4000 Sart Tilman, Liège Belgium*
[11]*Department of Physics, University of Bristol, Tyndall Avenue, Bristol BS8 1TL*
[12]*INAF – IASF Milan, Italy*
[13]*Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH*

Accepted 2006 July 25. Received 2006 July 25; in original form 2005 December 20

**ABSTRACT**

We present the *XMM* Large-Scale Structure Survey (*XMM*-LSS) cluster catalogue corresponding to the Canada–France–Hawaii Telescope Legacy Survey D1 area. The list contains 13 spectroscopically confirmed, X-ray selected galaxy clusters over 0.8 deg$^2$ to a redshift of unity and so constitutes the highest density sample of clusters to date. Cluster X-ray bolometric luminosities range from 0.03 to $5 \times 10^{44}$ erg s$^{-1}$. In this study, we describe our catalogue construction procedure: from the detection of X-ray cluster candidates to the compilation of a spectroscopically confirmed cluster sample with an explicit selection function. The procedure further provides basic X-ray products such as cluster temperature, flux and luminosity. We detected slightly more clusters with (0.5–2.0 keV) X-ray fluxes of $>2 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ than we expected based on expectations from deep *ROSAT* surveys. We also present the luminosity–temperature relation for our nine brightest objects possessing a reliable temperature determination. The slope is in good agreement with the local relation, yet compatible with a luminosity enhancement for the $0.15 < z < 0.35$ objects having $1 < T < 2$ keV, a population that the *XMM*-LSS is identifying systematically for the first time. The present study permits the compilation of cluster samples from *XMM* images whose selection biases are understood. This allows, in addition to studies of large-scale structure, the systematic investigation of cluster scaling law evolution, especially for low mass X-ray groups which constitute the bulk of our observed cluster population. All cluster ancillary data (images, profiles, spectra) are made available in electronic form via the *XMM*-LSS cluster data base.

**Key words:** surveys – X-rays: galaxies: clusters.

## 1 INTRODUCTION

The question of cosmic structure formation is substantially more complicated than the study of the spherical collapse of a pure dark matter perturbation in an expanding universe. While it is possible to predict theoretically how the shape of the inflationary

⋆Based on data collected with *XMM*, VLT, Magellan, NTT and CFH telescopes; ESO programme numbers are 070.A-0283, 070.A-907 (VVDS), 072.A-0104, 072.A-0312, 074.A-0360 and 074.A-0476.

†E-mail: mpierre@cea.fr
‡Present address: ESA, Villafranca del Castillo, Spain.

fluctuation spectrum evolves until recombination, understanding the subsequent formation of galaxies, active galactic nucleus (AGN) and galaxy clusters is complicated by the physics of non-linear growth and feedback from star formation. Attempts to use the statistics of visible matter fluctuations to constrain the nature of dark matter and dark energy are therefore reliant upon an understanding of non-gravitational processes.

Clusters, as the most massive entities of the Universe, form a crucial link in the chain of understanding. They lie at the nodes of the cosmic network, possess virialized cores, yet are still growing by accretion along filaments. The rate at which clusters form, and the evolution of their space distribution, depends strongly on the shape and normalization of the initial power spectrum, as well as on the dark energy equation of state (e.g. Rapetti, Allen & Weller 2005). Consequently, both a three-dimensional map of the cluster distribution and an evolutionary model relating cluster observables to cluster masses and shapes (predicted by theory for the average cluster population) are needed to test the consistency of structure formation models within a standard cosmology with the properties of clusters in the low-$z$ Universe.

The main goal of the *XMM* Large-Scale Structure Survey (*XMM*-LSS) is to provide a well-defined statistical sample of X-ray galaxy clusters to a redshift of unity over a single large area, suitable for cosmological studies (Pierre et al. 2004). In this paper, we present the first sample of *XMM*-LSS clusters for which canonical selection criteria are uniformly applied over the survey area. In this way, we demonstrate the properties of the survey along with a description of data analysis tools employed in the sample construction; the aim being to provide a deep and well-controlled sample of clusters and to investigate evolution trends, in particular for the low end of the cluster mass function. The paper will therefore act as a reference for future studies using *XMM*-LSS data. The chosen region is located at $36^\circ < \mathrm{RA} < 37^\circ$, $-5^\circ < \mathrm{Dec.} < -4^\circ$. This region is known as D1, one of the four deep areas of the Canada–France–Hawaii Telescope Legacy Survey[1] (CFHTLS). It also includes one of the VIMOS VLT Deep Survey patches (VVDS; Ilbert et al. 2005) and was observed at 1.4 GHz down to the μJy level by the VLA-VIRMOS Deep Field (Bondi et al. 2003). The rest of the *XMM*-LSS survey surrounds D1 and corresponds to part of the wide W1 CFHTLS component (see Pierre et al. 2004 for a general layout and associated multi-λ surveys) for which the complete cluster catalogue will be published separately. The sample is the result of a fine tuned X-ray plus optical approach developed with the aim of understanding the various selection effects. We describe the catalogue construction procedure in tandem with a companion paper presenting a detailed description of the X-ray pipeline developed as part of the *XMM*-LSS survey (Pacaud et al. 2006).

The deepest published statistical samples of X-ray clusters over a contiguous sky area to date are all based on the *ROSAT* All-Sky Survey (RASS): REFLEX (Böhringer et al. 2001), NORAS (Böhringer et al. 2000), NEP (Henry et al. 2001). In parallel, a number of serendipitous cluster surveys were conducted using deep *ROSAT* pointings with the goal of investigating the evolution of the cluster luminosity function, for example, Southern SHARC (Burke et al. 1997), RDCS (Rosati et al. 1998), 160 deg$^2$ (Vikhlinin et al. 1998), Bright SHARC (Romer et al. 2000), BMW (Moretti et al. 2004). The advent of the *XMM* satellite has provided an X-ray imaging capability of increased sensitivity and angular resolving power compared to *ROSAT*. The *XMM*-LSS employs 1–20 ks pointings and samples the cluster population to a depth of $\sim 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ – a flux sensitivity comparable to the deepest serendipitous *ROSAT* surveys (Rosati, Borgani & Norman 2002). However, *XMM* observations possess a narrower point spread function [PSF; full width at half-maximum (FWHM) ∼6 arcsec for *XMM* versus ∼20 arcsec for the *ROSAT* PSPC] which suggests that the reliable identification of extended sources can be performed for apparently smaller sources. Instrumental characteristics such as background noise and the complex focal plane configuration are also quite different. In this context, our dual aim of optimizing the *XMM*-LSS sensitivity and of quantifying the many selection biases led us to develop a dedicated source detection pipeline as well as specific optical identification and spectroscopic confirmation procedures: special attention is given to extended, X-ray faint sources whose identification requires deep optical/IR multicolour imaging. These steps are described in Section 2 along with the presentation of the D1 cluster catalogue. Section 3 presents the X-ray properties of the newly assembled sample and some optical characteristics. Section 4 summarizes the global properties of our sample within the context of a 'concordance' cosmological model. We conclude with a discussion of our cluster selection function in comparison with earlier works as well as the scaling laws for the low end of the cluster mass function.

Throughout the paper we assume $\Omega_\mathrm{M} = 0.27, \Omega_\Lambda = 0.73$ and $H_0 = 71\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. All X-ray flux measures are quoted in the [0.5–2] keV band. The generic name 'cluster' refers to the entire population of gravitationally bound galaxy systems, while we use the term 'groups' for those systems whose potential corresponds to an X-ray temperature lower than 2 keV.

## 2 THE X-RAY CLUSTER CATALOGUE

### 2.1 X-ray observations

The *XMM*-LSS D1 region consists of a mosaic of 10 *XMM* pointings that form part of the *XMM* Medium Deep Survey (XMDS; Chiappetti et al. 2005). The pointing layout is displayed in Fig. 1 and the properties of individual pointings are shown in Table 1. The nominal exposure per pointing is 20 ks for this subregion,[2] with the exception of pointing G07, whose nominal exposure time of 40 ks was reduced to ∼20 effective ks as a result of solar activity. The raw X-ray observations (ODFs) were reduced using the standard *XMM* Science Analysis System (XMMSAS version v6.1) tasks emchain and epchain for the MOS and PN detectors, respectively. High background periods, related to soft protons, were excluded from the event lists following the procedure outlined by Pratt & Arnaud (2002). Raw photon images in different energy bands were then created with a scale of 2.5 arcsec pixel$^{-1}$. A complete discussion of the image analysis and source characterization procedures are provided by Pacaud et al. (2006). Cluster detection was performed in the [0.5–2] keV band and was limited to the inner 11 arcmin of the *XMM* field. The total scanned area is 0.81 deg$^2$. Information regarding the individual pointings is summarized in Table 1 and the layout of the pointings on the sky displayed in Fig. 1.

[1] http://cdsweb.u-strasbg.fr:2001/Science/CFHLS/

[2] Outside the XMDS, the nominal exposure per pointing for the rest of the *XMM*-LSS is 10 ks.

**Table 1.** Properties of individual *XMM* pointings. Quoted exposures are effective exposures computed after filtering high background periods.

| Internal ID | *XMM* ID | RA (J2000) | Dec. (J2000) | MOS1, MOS2, pn exposure times (ks) |
|---|---|---|---|---|
| G01 | 0112680201 | 02:27:20.0 | −04:10:00.0 | 24.6, 25.3, 21.4 |
| G02 | 0112680201 | 02:26:00.0 | −04:10:00.0 | 10.1, 9.7, 6.7 |
| G03 | 0112680301 | 02:24:40.0 | −04:10:00.0 | 21.8, 21.7, 17.3 |
| G05 | 0112680401 | 02:28:00.0 | −04:30:00.0 | 23.5, 23.9, 12.5 |
| G06 | 0112681301 | 02:26:40.0 | −04:30:00.0 | 16.4, 16.6, 10.5 |
| G07 | 0112681001 | 02:25:20.0 | −04:30:00.0 | 22.5, 25.1, 18.6 |
| G08 | 0112680501 | 02:24:00.0 | −04:30:00.0 | 21.2, 21.3, 15.9 |
| G10 | 0109520201 | 02:27:20.0 | −04:50:00.0 | 24.7, 24.6, 18.5 |
| G11 | 0109520301 | 02:26:00.0 | −04:50:00.0 | 21.7, 21.8, 16.1 |
| G12 | 0109520401 | 02:24:40.0 | −04:50:00.0 | Not usable because of very high flare rate |

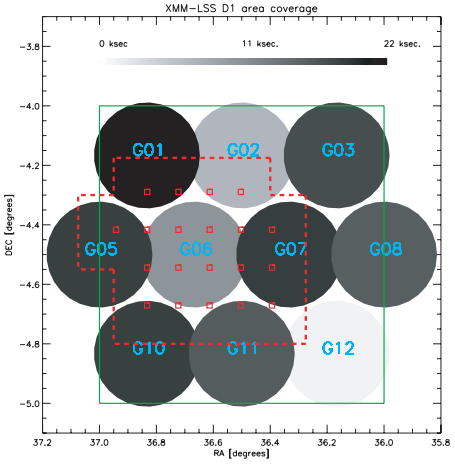


**Figure 1.** The *XMM* pointing mosaic over the D1 area (green square). The radius of the displayed pointings is 11 arcmin. The grey-scale indicates effective mean exposure time per detector, after removal of high background periods. The red squares show the centres of the VVDS pointings (Ilbert et al. 2005) and the red dotted line indicates the total area covered by the VVDS. The VLA-VIRMOS Deep Field encompasses exactly the D1 region.

### 2.2 Cluster X-ray detection and optical identification procedure

The compilation of an X-ray cluster sample featuring positional and redshift data ultimately requires the input of optical and/or near-infrared (NIR) data in order to select putative cluster galaxies for which precise redshifts can be obtained. Therefore, although X-ray selection is employed to better avoid projection effects, to provide direct clues about cluster masses and to provide more easily tractable selection criteria, optical/NIR data for each cluster must be assessed in order to obtain cluster redshift data. The goal of *XMM*-LSS is to produce a faint, statistical cluster catalogue over a wide spatial area (several tens of square degrees) and a large redshift interval (zero to unity). Cluster identification procedures must therefore identify robustly a wide range of cluster properties at both X-ray and optical/NIR wavelengths. Given the above requirements the *XMM*-LSS has developed over the last three years from initially simple and very robust cluster selection procedures to a refined, quantitative approach focusing on key cluster selection parameters.

Developing the X-ray pipeline was an essential part of the procedure as is reflected in the successive publications. We summarize these developments below.

(i) Spectroscopic observations during 2002 were performed for a number of cluster candidates identified following the method developed by (Valtchanov, Pierre & Gastaud 2001); extended X-ray sources were accepted as candidate clusters if associated with a spatial overdensity of galaxies displaying a uniform red colour sequence determined using either CFHT/CFH12K *BVRI* or CTIO/MOSAIC $Rz$ imaging. This approach maximized the success rate of the first *XMM*-LSS spectroscopic observations (conducted in the last quarter of 2002) which demonstrated that clusters to a redshift of 1 are detectable with 1–20 ks *XMM* observations (Valtchanov et al. 2004; Willis et al. 2005).

(ii) In order to proceed towards a purely X-ray selected sample – that is, to reduce contamination by spurious extended sources – a maximum likelihood procedure named XAMIN was combined with the wavelet-based detection algorithm developed previously (Pierre et al. 2004). The sample of candidate clusters thus generated was investigated during the spectroscopic observations conducted in 2003 and 2004.

(iii) Finally, the combination of spectroscopic results for the above cluster sample with a detailed study of the simulated performance of the X-ray pipeline led to the definition of three clearly defined classes of X-ray cluster candidates.

The cluster identification procedure described above satisfies the goal of generating a relatively uncontaminated sample of X-ray clusters with a well-defined selection function. A detailed description of the X-ray parameters employed to generate each class of cluster candidate is described in the following section.

### 2.3 The cluster classification and sample

The ability to detect faint, extended sources in X-ray images is subject to a number of factors. Although the apparent size of a typical cluster ($R_c = 180$ kpc) is significantly larger than the *XMM* PSF (on-axis FWHM ∼6 arcsec) at any redshift of interest,[3] it is incorrect to assume that all clusters brighter than a given flux will be detected – unless the flux limit is set to some high value. Cluster

[3] $100 > R_c > 20$ arcsec for $0.1 < z < 1$.

detectability depends not only on the instrumental PSF, object flux and morphology but also upon the background level and the detector topology (e.g. CCD gaps and vignetting), in addition to the ability of the pipeline to separate close pairs of point-like sources – all of which are a function of the specified energy range (Scharf 2002). We thus stress that the concept of 'sky coverage', that is, the fraction of the survey area covered at a given flux limit, is strictly valid only for point sources because, for faint extended objects, the detection efficiency is surface brightness limited (rather than flux limited). Moreover, since the faint end of the cluster luminosity function is poorly characterized at $z > 0$, it is not possible to estimate a posteriori what fraction of groups remain undetected, unless a cosmological model is assumed, along with a thorough modelling of the cluster population out to high redshift; the lower limit of the mass or luminosity function being here a key ingredient. To our knowledge, this has never been performed in a fully self-consistent way so far for any deep X-ray cluster survey ($F_{\rm lim} \sim 2$–$5 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$). It is also important to consider that the flux recorded from a particular cluster represents only some fraction of the total emitted flux and must therefore be corrected by integrating an assumed spatial emission model to large radius.

Consequently, with the goal of constructing deep controlled samples suitable for cosmology we define, rather than flux limits, classes of extended sources. These are defined in the extension and significance parameter space and correspond to specific levels of contamination and completeness. As shown by Pacaud et al. (2006), extensive simulations of various cluster and AGN populations generate detection probabilities as a function of sources properties. This enables a simultaneous estimate of the source completeness levels and of the frequency of contamination by misclassified point-like sources or spurious detections. X-ray source classification was performed using XAMIN and employs the output parameters: extent, likelihood of extent, likelihood of detection. The reliability of the adopted selection criteria has been checked against the current sample of 60 spectroscopically confirmed *XMM*-LSS clusters. We have defined three classes of extended sources as follows.

(i) The C1 class is defined such that no point sources are misclassified as extended and is described by extent $> 5$ arcsec, likelihood of extent $>33$ and likelihood of detection $>32$. The C1 class contains the highest surface brightness extended sources and inevitably includes a few nearby galaxies – these are readily discarded from the sample by inspection of optical overlays.

(ii) The C2 class is described by extent $>5$ arcsec and $15 <$ likelihood of extent $<33$ and typically displays a contamination rate of 50 per cent. The C2 class includes clusters fainter than C1, in addition to a number of nearby galaxies. Contaminating sources include saturated point sources, unresolved pairs, and sources strongly masked by CCD gaps, for which not enough photons were available to permit reliable source characterization. Contaminating sources were removed after a visual inspection of the optical/NIR data for each field and in some cases as a result of follow-up spectroscopy.

(iii) The C3 class was constructed in order to investigate clusters at the survey sensitivity limit, particularly clusters at high redshift. Sources within the C3 class typically display $2 <$ extent $< 5$ arcsec and likelihood of extent $> 4$. Selecting such faint, marginally extended sources generates a high contamination rate. However, low selection thresholds are required to identify extended sources at the survey limit: faint sources will never be characterized by high likelihood values. When refining the C3 class, the X-ray, optical and NIR appearance was examined thoroughly. Generally speaking, C3 sources display low surface brightness or extended emission affected by a point source. Additional constraints included that the detection should be located at an off-axis angle $<10$ arcmin and that the total detection should generate 30 photons or greater (stronger constraint on the off-axis value is necessary because weak objects are subject to strong distortions beyond 10 arcmin, thus hardly measurable). The most plausible C3 candidates were investigated spectroscopically and confirmed clusters are presented.

The analysis of simulated cluster and AGN data permits the computation of selection probabilities for the C1 and C2 cluster samples.[4] The extent to which the C1 and C2 classes are comparable to flux-limited samples is analysed in detail by Pacaud et al. (2006) and further discussed in the last section of this paper. The selection probability for C3 clusters has not been determined.

### 2.4 Determination of cluster redshifts

The *XMM*-LSS spectroscopic Core Programme aims at the redshift confirmation of the X-ray cluster candidates; velocity dispersion may subsequently be obtained for a subsample of confirmed clusters as a second step programme. Spectroscopic observations were performed using a number of telescope and instrument combinations and are summarized in Table 2. Details of which observing configuration was employed for each cluster are presented in Table 3.

The minimum criterion required to confirm a cluster was specified to be three concordant redshifts ($\pm 3000$ km s$^{-1}$) within a projected scale of approximately 500 kpc of the X-ray emission centroid, computed at the putative cluster redshift. For nearby X-ray clusters of temperature $T_{\rm X} = 2$ keV, a radius of 500 kpc corresponds to approximately 50 per cent of the virial radius and encloses 66 per cent of the total mass – with both fractions being larger for higher temperature clusters (Arnaud, Pointecouteau & Pratt 2005). The final cluster redshift was computed from the non-weighted mean of all galaxies within this projected aperture and within a rest-frame velocity interval $\pm 3000$ km s$^{-1}$ of the interactively determined redshift peak. Potential cluster galaxies are selected for spectroscopic observation by identifying galaxies displaying a uniformly red colour distribution within a spatial aperture centred on the extended X-ray source (see Andreon et al. 2005, for more details). Cluster members flagged via this procedure are then allocated spectroscopic slits in order of decreasing apparent magnitude (obviously avoiding slit overlap).

The exact observing conditions for each cluster form a heterogeneous distribution. However, each cluster was typically observed with a single spectroscopic mask featuring slitlets of 8–10 arcsec in length and 1–1.4 arcsec in width. The use of a different telescope and instrument configuration generally restricts the available candidate cluster member sample to a different limiting *R*-band magnitude (assuming an approximately standard exposure time of 2 h per spectroscopic mask). Typical apparent *R*-band magnitude limits generating spectra of moderate ($S/N > 5$) quality for each telescope were found to be the following: VLT/FORS2 (23), Magellan/LDSS2 (22) and NTT/EMMI (21.5). Spectroscopic data reduction followed standard IRAF[5] procedures. Redshift determination was performed by cross-correlating reduced, one-dimensional spectra with suitable

[4] The C1 and C2 classes are defined from simulations representative of a mean exposure time of 10 ks. In the present paper, we keep the same definition as the signal-to-noise ratio (S/N) increase is only $\sqrt{2}$.

[5] IRAF is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.

**Table 2.** Observing resources employed to determine cluster spectroscopic redshifts.

| Telescope | Instrument | Grism + filter | Approximate resolving power ($R$) | Identifier |
|---|---|---|---|---|
| VLT | FORS2 | 300V + GC435 | 500 | 1 |
| VLT | FORS2 | 600RI + GC435 | 1000 | 2 |
| VLT | FORS2 | 600z | 1300 | 3 |
| NTT | EMMI | Grism #3 | 700 | 4 |
| Magellan (Clay) | LDSS-2 | Medium red | 500 | 5 |
| VLT | VIMOS | LRRED | 220 | 6 |

**Table 3.** Spectroscopically confirmed X-ray clusters within the D1 area.

| Source | XLSSC | RA (°) | Dec. (°) | *XMM* pointings | Off-axis angle[a] (arcmin) | Redshift | # of members[b] | Observed (see Table 2) |
|---|---|---|---|---|---|---|---|---|
| C1 | | | | | | | | |
| XLSS J022404.1−041329[c] | 029 | 36.0172 | −4.2247 | G03 | 9.0 | 1.05 | 5 | 3 |
| XLSS J022433.5−041405 | 044 | 36.1410 | −4.2376 | G03 | 3.8 | 0.26 | 9 | 4 |
| XLSS J022524.5−044042 | 025 | 36.3526 | −4.6791 | G07 | 10.3 | 0.26 | 10 | 5 |
| XLSS J022530.6−041419 | 041 | 36.3777 | −4.2388 | G02 | 9.1 | 0.14 | 9 | 4 |
| XLSS J022609.7−045804 | 011 | 36.5403 | −4.9684 | G11 | 8.1 | 0.05 | 7 | 4 |
| XLSS J022709.1−041759[d] | 005 | 36.7877 | −4.3002 | G01 | 7.8 | 1.05 | 5 | 2 |
| XLSS J022725.8−043213[e] | 013 | 36.8588 | −4.5380 | G05 | 8.1 | 0.31 | 11 | 5 |
| XLSS J022739.9−045129 | 022 | 36.9178 | −4.8586 | G10 | 5.6 | 0.29 | 5 | 5 |
| C2 | | | | | | | | |
| XLSS J022725.0−041123 | 038 | 36.8536 | −4.1920 | G01 | 1.9 | 0.58 | 7 | 4 |
| C3 | | | | | | | | |
| XLSS J022522.7−042648 | a | 36.3454 | −4.4468 | G07 | 3.9 | 0.46 | 4 | 2 |
| XLSS J022529.6−042547 | b | 36.3733 | −4.4297 | G07 | 5.8 | 0.92 | 7 | 6 |
| XLSS J022609.9−043120 | c | 36.5421 | −4.5226 | G06 | 8.0 | 0.82 | 8 | 6 |
| XLSS J022651.8−040956 | d | 36.7164 | −4.1661 | G01 | 6.6 | 0.34 | 5 | 1 |

[a]The off-axis angle is computed from the barycentre of the optical axes of the three telescopes using XMMSAS variables XCEN YCEN weighted by the mean detector sensitivity (see Pacaud et al. 2006). [b]Only galaxies within a projected distance <500 kpc of the cluster centre are counted.[c]Listed by Andreon et al. (2005). [d]Already published by Valtchanov et al. (2004). [e]Already published by Willis et al. (2005).

templates within the IRAF procedure xcsao (Kurtz & Mink 1998) and confirmed via visual inspection. A more detailed description of the spectroscopic techniques employed by the *XMM*-LSS survey can be found in Valtchanov et al. (2004) and Willis et al. (2005). In addition to the above spectroscopic observations, cluster redshift information was supplemented where available by spectra contributed from the VVDS (see Table 3).

The D1 X-ray clusters with confirmed redshifts are presented in Table 3. C1 and C2 confirmed clusters constitute a controlled sample (following Section 2.3) and are associated with the label 'XLSSC' and a three digit identifier.[6] This nomenclature is used to identify individual clusters in any later discussion. The completeness of C3 sources is not addressed. In the case where a particular cluster is present in two separate *XMM* pointings, only the pointing where the cluster is the closest to the optical centre has been used to measure its properties. Note that the off-axis restriction imposed on the C3 clusters excludes two faint clusters located at the very border of the D1 area and reported in the Willis et al. (2005) initial sample (XLSSUJ022633.9−040348, XLSSUJ022628.2−045948). Cluster redshift values given in Table 3 are the unweighed mean of relatively small member samples and observed, in a few cases, using different spectrographs. As this approach may result in relatively large (several hundred km s$^{-1}$) uncertainties in the computed redshift, the cluster redshift precision reported in Table 3 has been set to two decimal points (3000 km s$^{-1}$). X-ray/optical overlays of each cluster field are displayed in Fig. B1.

[6]The acronym is defined at CDS at the following URL http://vizier.u-strasbg.fr/cgi-bin/Dic?XLSSC.

## 3 X-RAY PROPERTIES OF CONFIRMED CLUSTERS

The spectral and spatial X-ray data for each spectroscopically confirmed cluster were analysed to determine the temperature, spatial morphology and total bolometric luminosity of the X-ray emitting gas.

### 3.1 Spectral modelling and temperature determination

A complete description of the spectral extraction and analysis procedures as applied to X-ray sources with low signal levels using the XSPEC package (Arnaud 1996), together with a discussion on the accuracy of the computed temperatures, are presented in Willis et al. (2005). We summarize the principal steps below.

Spectral data were extracted within an aperture of specified radius (see Table 4) and a corresponding background region was defined by a surrounding annulus. Photons were extracted over the energy range [0.3–10] keV, excluding the energy range [7.5–8.5] keV due to emission features produced by the pn detector support. Analyses of

**Table 4.** Spectral X-ray parameters. The radius of the circular aperture used for the spectral extraction is denoted by $R_{\rm spec}$. Computed source counts are summed over the three detectors. The suffix 'F' applied to temperature values indicates that a reliable temperature fit was not achieved and a gas temperature of 1.5 keV was assumed for the computation of bolometric luminosity.

| Cluster | $R_{\rm spec}$ (arcscec) | Source counts in $R_{\rm spec}$ [0.3–7.5]+[8.5–10] keV | $T$ (keV) | C-statistics (per degree of freedom) | $r_{500}$ (Mpc) | $r_{500}$ (arcsec) |
|---|---|---|---|---|---|---|
| XLSSC 029 | 33 | 311 | $4.1^{+1.7}_{-1.0}$ | 1.08 | 0.52 | 67 |
| XLSSC 044 | 55 | 234 | $1.3^{+0.2}_{-0.2}$ | 1.15 | 0.40 | 100 |
| XLSSC 025 | 35 | 661 | $2.0^{+0.5}_{-0.3}$ | 1.06 | 0.53 | 129 |
| XLSSC 041 | 45 | 523 | $1.3^{+0.3}_{-0.1}$ | 1.00 | 0.43 | 172 |
| XLSSC 011 | 68 | 425 | $0.6^{+0.2}_{-0.1}$ | 1.04 | 0.28 | 272 |
| XLSSC 005 | 35 | 164 | $3.7^{+3.5}_{-1.4}$ | 1.02 | 0.49 | 60 |
| XLSSC 013 | 30 | 161 | $1.0^{+0.2}_{-0.2}$ | 0.92 | 0.33 | 73 |
| XLSSC 022 | 39 | 1304 | $1.7^{+0.2}_{-0.2}$ | 0.91 | 0.47 | 109 |
| XLSSC 038 | 33 | 118 | 1.5F | – | 0.37 | 56 |
| cluster a | 24 | 160 | 1.5F | – | 0.40 | 69 |
| cluster b | 30 | <100 | 1.5F | – | 0.30 | 38 |
| cluster c | 30 | <100 | 1.5F | – | 0.32 | 42 |
| cluster d | 25 | 157 | $0.9^{+0.2}_{-0.2}$ | 0.74 | 0.31 | 65 |

simulated spectral data with less than 400 total counts indicated that using C-statistics on unbinned spectral data produced a systematic offset in the computed temperature. This bias is significantly reduced for such faint spectra when the data are resampled such that at least five photons are present in each spectral bin corresponding to the background spectrum. We determined that this approach produces reliable temperature measures for low temperature ($T < 3$ keV), low count level (<400 photons) spectral data. The assumed fitting model employs an absorbed APEC plasma (Smith et al. 2001) with a fixed metal abundance ratio given by Grevesse & Sauval (1999) and set to 0.3 of the solar value. Absorption due to the Galaxy is modelled using the WABS function (Morrison & McCammon 1983), fixing the hydrogen column density to the value given by Dickey & Lockman (1990) at the cluster position (typically $\sim 2.6 \times 10^{20}$ cm$^{-2}$). Where the temperature fitting procedure failed to converge to a sensible model (due to low signal levels), the source temperature was fixed at 1.5 keV. Results of the X-ray spectral analysis are presented in Table 4. An example of cluster spectrum and fit is shown in Fig. 2.

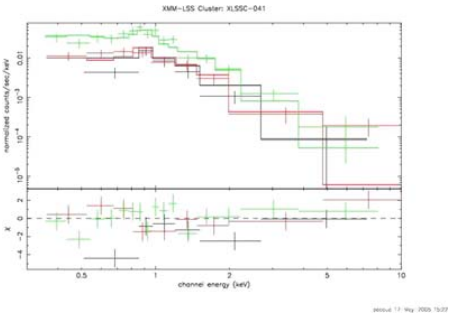


**Figure 2.** Fitted spectrum and residuals for cluster XLSSC 041 (MOS1: black, MOS2: red, pn: green).

### 3.2 Source morphology and spatial modelling

Sources detected using XAMIN are initially compared to two surface brightness models describing the two-dimensional photon distribution: a point source and a circular $\beta$-profile of the form

$$S(r) = \frac{A}{\left[1 + (r/R_{\rm c})^2\right]^{3\beta-1/2}}, \quad (1)$$

where $\beta = 2/3$ is fixed while the core radius, $R_{\rm c}$, and profile normalization, $A$, are permitted to vary (Cavaliere & Fusco-Femiano 1976). Each profile is convolved with the mean analytical PSF at the corresponding off-axis location and a comparison of the statistical merit achieved by each profile provides an effective discriminator of point and extended sources in addition to an initial estimate of the source flux (see Pacaud et al. 2006).

The photometric reliability of this procedure when applied to faint, extended sources is affected by the presence both of gaps in the *XMM* CCD array and by nearby sources (although both are described within the fitting procedure) – largely due to variations in the true source morphology and the fact that a larger fraction of the total emission is masked by the background when compared to brighter sources. Although such effects are naturally incorporated into the selection function appropriate for each cluster class via simulations (Pacaud et al. 2006), a further interactive spatial analysis was performed on each spectroscopically confirmed cluster in order to optimize the measure of the total emission (i.e. flux and luminosity) within a specified physical scale.

The photon distribution for each confirmed cluster is modelled using a one-dimensional circular $\beta$-profile in which $\beta$, $R_{\rm c}$ and $A$ are permitted to vary. Photons from the three *XMM* detectors are co-added applying a weight derived from the relevant exposure map and pixels associated with nearby sources are excluded. Photons are binned in concentric annuli of width 3 arcsec centred on the cluster X-ray emission. Radial data bins are subsequently resampled to generate a minimum S/N > 3 per interval. The background is computed at large radius assuming a constant particle contribution plus vignetted cosmic emission. The above resampling procedure is then applied to the circular $\beta$-profile convolved with the mean

analytical PSF[7] computed at the corresponding cluster off-axis angle (Ghizzardi 2002). Model cluster profiles are realized in this manner over a discrete grid of $\beta$ and $R_c$ values with the best-fitting model for each cluster determined by minimizing the $\chi^2$ statistic over the parameter grid. Finally, the best-fitting spatial profile (at this point in units of photon count rate) is integrated out to a specified physical radius and converted into flux and luminosity units using standard procedures within XSPEC.

The majority of confirmed clusters are apparently faint, displaying total photon counts of order a few hundred, and the observed photon distribution in many cases represents only a fraction of the extended X-ray surface brightness distribution. Under such conditions the parameters $\beta$ and $R_c$ are degenerate when fitted simultaneously, limiting the extent to which 'best-fitting' parameters can be viewed as a physically realistic measure of the cluster properties, although providing a useful ad hoc parametrization. For this reason we do not quote best-fitting values of $\beta$ and $R_c$ derived for each confirmed cluster. The uncertainty associated with the procedure is evaluated using a large suite of simulated observations (and subsequent analyses) of clusters of specified surface brightness properties (i.e. $\beta, R_c$ and apparent brightness – Pacaud et al. in preparation). The fractional uncertainty can then be quoted as a function of the number of photons collected within the fitting radius $R_{\rm fit}$ (the maximum radius out to which the resampling criterion S/N > 3 was achieved) and the radius to which the profile is calculated (possibly, extrapolated). Note that, as shown in the next section, almost all clusters have $R_{\rm fit}$ greater than the physical integration radius, hence requiring no profile extrapolation. For the very faintest clusters (those with total photon counts less than ∼100) a simple sum within a circular aperture was applied.

As an example of the above spatial fitting procedure, Fig. 3 shows the data analysis regions applied to the cluster XLSSC 041. Fig. 4 displays the resulting one-dimensional radial profile and the best-fitting surface brightness model for the same cluster.

### 3.3 Determination of cluster flux and luminosity

Values of flux and luminosity for confirmed clusters are obtained by integrating the cluster emission model, described by the appropriate XSPEC plasma emission and surface brightness models, out to a specified physical radius. We use a different physical radius for flux measures as opposed to luminosity – mainly because tabulated flux values for cluster surveys present in the literature prefer an estimate of the 'total' flux within a limited energy interval whereas luminosity values are computed as the bolometric emission within a physical radius corresponding to a constant overdensity in an evolving universe (e.g. $r_{500}$).

Flux values are computed by integrating the best-fitting $\beta$-profile to a radius of 500 kpc. The specified aperture includes a substantial fraction of the total flux (approximately 2/3 of the flux from $\beta$-profile described by $\beta = 2/3$ and $R_c = 180$ Mpc) yet avoids uncertainties associated with the extrapolation of the profile to large radii.

In order to obtain cluster luminosities within a uniform physical radius, we have integrated the best-fitting $\beta$-profile for each cluster to $r_{500}$, that is, the radius at which the cluster mass density reaches 500 times the critical density of the Universe at the cluster redshift. Values of $r_{500}$ for each cluster were computed using the mass–temperature data of Finoguenov, Reiprich & Böhringer (2001) which, when converted to our assumed cosmology and fitted with an orthogonal regression line, yield the expression $r_{500} = 0.391\ T^{0.63}\ h_{70}(z)^{-1}$ Mpc for clusters in the $0.7 < T < 14$ keV interval. As reported in Willis et al. (2005), values of $r_{500}$ from this formula agree well with those derived from assuming an isothermal $\beta$-profile for clusters displaying $T \lesssim 4$ keV.

For each cluster in the confirmed sample, with the exception of XLSSC 005, the computed values of $r_{500}$ lie within radius of, or close to, the region employed to fit the $\beta$-profile (hence, for this cluster alone, the photometric errors are large as reported in Table 5).

[7]The convolution of the two profiles models the photon distribution factor introduced by the two-dimensional convolution (Arnaud et al. 2002).

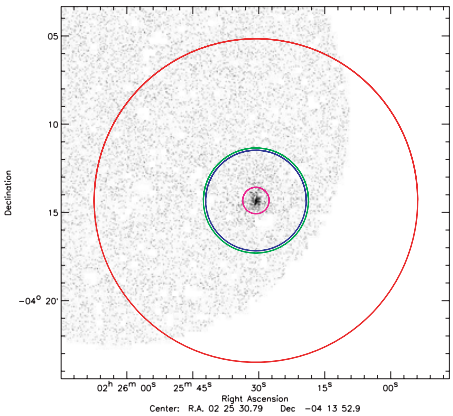


**Figure 3.** An example of the spatial and spectral analysis regions applied to cluster XLSSC 041. The photon image is displayed. The purple circle indicates the spectral extraction region. The green circle indicates $r_{500}$. The blue circle indicates $R_{3\sigma} = R_{\rm fit}$. All X-ray sources, except the cluster of interest, are masked. The external red circle delineates the region used for the fit.

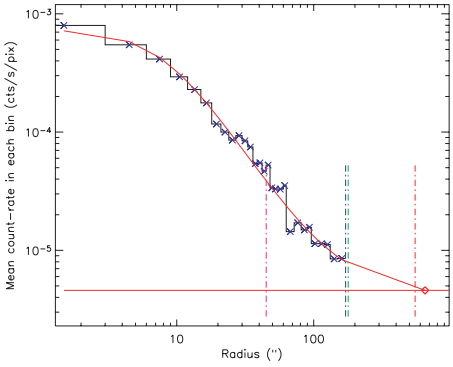


**Figure 4.** The radially averaged X-ray emission profile for cluster XLSSC 041. The red curve indicates the best-fitting model compared to the data (black histogram with blue crosses). The vertical lines follow the same colour coding as Fig. 3. The red horizontal line marks the background level.

**Table 5.** Results of the spatial analysis for confirmed clusters. See text for the definition of the fitting radius $R_{\rm fit}$. The net source counts are computed within $R_{\rm fit}$ and are uncorrected for vignetting. A value of 'NF' indicates that no reliable spatial fit was possible for the cluster: the source counts, flux and luminosity were computed applying a circular aperture of radius 500 kpc. Flux values are computed by integrating the best-fitting cluster $\beta$-profile out to a radius of 500 kpc. The photometric precision indicates the mean $1\sigma$ errors estimated from analyses of simulated cluster data and accounts for the profile fitting uncertainties only (see text for details).

| Cluster | $R_{\rm fit}$ (arcsec) | Source counts in $R_{\rm fit}$ in [0.5–2] keV | $F_{[0.5-2]}$ ($10^{-14}$ erg s$^{-1}$ cm$^{-2}$) | $L_{\rm bol}(r_{500})$ ($10^{44}$ erg s$^{-1}$) | Photometric accuracy (per cent) |
|---|---|---|---|---|---|
| XLSSC 029 | 60 | 361 | 3.1 | 4.8 | 20 |
| XLSSC 044 | 129 | 318 | 2.5 | 0.11 | 15 |
| XLSSC 025 | 123 | 905 | 9.4 | 0.52 | 15 |
| XLSSC 041 | 171 | 819 | 20.6 | 0.24 | 15 |
| XLSSC 011 | 354 | 972 | 16.4 | 0.015 | 15 |
| XLSSC 005 | 39 | 128 | 1.1 | 1.5 | 60 |
| XLSSC 013 | 234 | 383 | 2.7 | 0.15 | 20 |
| XLSSC 022 | 171 | 1785 | 9.8 | 0.65 | 10 |
| XLSSC 038 | NF | [60] | 0.3 | 0.09 | – |
| cluster a | NF | [108] | 0.7 | 0.1 | – |
| cluster b | NF | [52] | 0.4 | 0.4 | – |
| cluster c | NF | [29] | 0.3 | 0.3 | – |
| cluster d | 432 | 417 | 0.83 | 0.078 | 20 |

Bolometric X-ray luminosities were calculated for each cluster by extrapolating the APEC plasma code corresponding to the best-fitting temperature to an energy of 50 keV. Values of $r_{500}$, flux and luminosity for each confirmed cluster are listed in Tables 4 and 5.

In Appendix A, we analyse the impact of further sources of uncertainty affecting the luminosity and temperature measurements.

Appendix B gathers notes on the individual clusters and investigate, among others, the possibility that the computed clusters brightness values may be contaminated by AGN emission. We conclude that none of the C1 and C2 clusters (i.e. those used for detailed population statistics) contained in the D1 sample are significantly contaminated by AGN emission.

### 3.4 Trends in the $L_{\rm X}$ versus $T_{\rm X}$ correlation

Fig. 5 displays the $L_{\rm X}$ versus $T_{\rm X}$ distribution of the D1 clusters for which it was possible to measure a temperature (eight C1 and one C3 objects). Although the D1 area represents only a subset of the anticipated *XMM*-LSS area, the C1 sample is complete and reliable temperature information is available for all systems. It is therefore instructive to consider trends in the $L_{\rm X}$ versus $T_{\rm X}$ distribution in anticipation of a larger sample of C1 clusters from the continuing survey. The location of C1 clusters in the $L_{\rm X}$ versus $T_{\rm X}$ plane is compared to a regression line computed for a combined sample of local sources based upon the group data of Osmond & Ponman (2004) and cluster data of Markevitch (1998). The computed regression line takes the form, $\log L_{\rm X} = 2.91 \log T_{\rm X} + 42.54$, for bolometric luminosity computed within $r_{500}$. A complete discussion of the regression fit will be presented by Helsdon & Ponman (in preparation).

One issue of interest concerns the properties of intermediate redshift ($z \sim 0.3$) X-ray groups (i.e. $T_{\rm X} \sim 1$–2 keV). Such systems dominate the *XMM*-LSS numerically and, when compared to higher temperature, higher mass clusters, are expected to demonstrate to a greater degree the effects of non-gravitational physics in the evolution of their X-ray scaling relations with respect to self-similar evolution models. The luminosity of X-ray sources in *XMM*-LSS may be compared to those of local sources at the same temperature by computing the luminosity enhancement factor, $F = L_{\rm obs}/L_{\rm pred}$, where $L_{\rm obs}$ is the observed cluster X-ray luminosity within a radius, $r_{500}$, and $L_{\rm pred}$ is the luminosity expected applying the fitted $L_{\rm X}$ versus $T_{\rm X}$ relation computed for the local sources and the *XMM*-LSS measured temperature. Sources XLSSC 013, 022 and 041 have a luminosity enhancement factors $F \approx 3$–4, compared to a value 1.15 expected from self-similar[8] luminosity scaling within $r_{500}$.

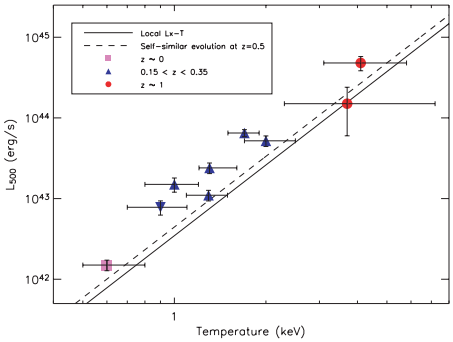


**Figure 5.** $L_{\rm X}(r_{500})$ versus $T_{\rm X}$ relation for the clusters for which it was possible to derive a temperature; all of them but *cluster d* (displayed as an upside-down triangle) are C1. The solid line gives the mean local $L_{\rm X} - T_{\rm X}$ relation (see text), while the dotted line is the expected luminosity enhancement assuming self-similar luminosity scaling within $r_{500}$ at $z = 0.5$. Different plotting symbols indicate clusters located within three different redshift intervals.

From the local Universe, we know that low-temperature groups show a larger dispersion in the *L–T* relation than massive clusters (Helsdon & Ponman 2003). This reflects their individual formation histories, since they are particularly affected by non-gravitational effects, as well as the possible contributions from their member

[8]Self-similar implies that the luminosity scales as the Hubble constant when integrated within a radius corresponding to a fixed ratio with respect to the critical density of the Universe as a function of redshift (Voit 2005).

galaxies. The apparent biasing towards more luminous objects and/or cooler system could come from the fact that we detect more easily objects having a central cusp, that is, putative cool-core groups. This has the effect of both decreasing the temperature and increasing the luminosity. The bias could also simply reflect the fact that we can measure a temperature only for the brightest objects.

In order to test this hypothesis, we have considered a few orders of magnitude. The local $L$–$T$ relation predicts a luminosity of $1.1 \times 10^{43}$ and $2.6 \times 10^{43}$ erg s$^{-1}$ for $T = 1.5$ and 2 keV groups, respectively. A factor of 2 under luminosity for such objects would thus correspond to $5.6 \times 10^{42}$ and $1.3 \times 10^{43}$ erg s$^{-1}$. In Table 5, we note that (i) the lowest flux cluster (XLSSC 005) is detected with some 150 photons in $R_{\rm fit}$; the X-ray image appears moreover to be quite flat; (ii) group XLSSC 044 ($z = 0.26$) has a luminosity of $1.1 \times 10^{43}$ erg s$^{-1}$, for some 300 photons in $R_{\rm fit}$. From this, we infer that we could have detected groups around $z \sim 0.25$, having $1.5 < T < 2$ keV that are under luminous by a factor of 2, if any were present in our sample. Below 1.5 keV, these objects are likely to remain undetected.

The coming availability of the larger *XMM*-LSS sample will permit a more reliable assessment of such effects on the morphology–luminosity–temperature plane for such groups (Pacaud et al. in preparation).

## 4 DISCUSSION AND CONCLUSION

We have used 20 ks *XMM* images to construct a deep sample of galaxy clusters. The total cluster surface density of 15.5 deg$^{-2}$ is approximately five times larger than achieved previously with the deepest *ROSAT* cluster surveys (e.g. RDCS, Borgani et al. 2001). On the one hand, from the optical point of view, we note that none of the detected clusters shows strong lensing features, hence the likely absence of massive clusters in the D1 area.[9] This is consistent with the fact that the highest cluster X-ray temperature is only 4 keV: this temperature corresponds to a mass of $\sim 3 \times 10^{14}$ M$_\odot$ and, in a standard Lambda cold dark matter halo model (Pacaud et al. 2006), the density of clusters more massive than this limit, that is, those most likely to produce strong lensing, is $\sim$0.8 deg$^{-2}$. On the other hand, it is indeed a salient property of the *XMM*-LSS to unveil for the first time the bulk of the $1 < T < 2$ keV group population in the $0.1 < z < 0.4$ range along with its capability of detecting $z \geqslant 1$ clusters. We further review below the main properties of the sample.

The C1 cluster subsample corresponds to a purely X-ray selected sample (zero contamination) and displays a surface density of $\sim$9 deg$^{-2}$. Reliable temperature information is available for all C1 sources and optical spectroscopic observations are required only to confirm the cluster redshift. Relaxing the selection criteria used to generate the C1 sample creates additional samples labelled C2 and C3. However, optical imaging and spectroscopic data are required to identify bona fide clusters within these samples. The C2 sample possesses a well-defined X-ray selection function (approximately 50 per cent of sources are confirmed as clusters) and the surface density of C1 + C2 clusters is $\sim$11 deg$^{-2}$. Sources labelled C3 represent significant detections out with the C1 and C2 selection criteria and, given the high contamination rate, we do not compute a selection function for these sources. The C3 sample contains potentially interesting objects and points to our ultimate sensitivity for cluster detection which appears to be around $5 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$; however, we stress that the full selection function for the C1 and C2 samples is multidimensional (see below). Noting this caveat – for comparison purposes only – the quoted flux sensitivity corresponds to a cluster of $\sim 7 \times 10^{43}$ erg s$^{-1}$ ($T \sim 3$ keV) at $z = 1$ and to a group of $3.5 \times 10^{42}$ erg s$^{-1}$ ($T = 1$ keV) at $z = 0.3$. We further note that no C1 or C2 cluster emission appears to be significantly contaminated by AGN activity – partly a result of the high threshold put on the extent likelihood for these samples.

Having the D1 *XMM*-LSS sample now assembled, it is instructive to examine in what manner it differs from a purely flux-limited sample. This question is phenomenologically addressed by Pacaud et al. (2006) as the answer depends on two major ingredients: (1) the pipeline efficiency (involving itself the many instrumental effects) – this is quantified by means of extensive simulations; (2) the characteristics of the cluster population out to a redshift of unity at least – this latter point being especially delicate as the low-end of the cluster mass function, critical for the survey sensitivity, is poorly known and cluster scaling law evolution, still a matter of debate. Hence the need for a self-consistent approach basically involving a cosmological model, a halo mass spectrum and some $L_{\rm X}(M, T, R_{\rm c}, z)$ function[10] describing the evolution of the cluster intrinsic properties that directly impact on the cluster detection efficiency. The principal conclusion regarding the use of a single flux limit is that, to obtain a cluster sample displaying a high level of completeness and reasonably low contamination, the flux limit has to be set to some high value, for example, $F > 5 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ in the case of *XMM*-LSS. The present study demonstrates on real data the advantage of using the C1 set of criteria as a well-defined sample that includes groups down to $T = 1$ keV and fluxes as low as $10^{-14}$ erg s$^{-1}$ cm$^{-2}$ and, consequently, significantly increases the size of the purely X-ray selected sample. Although the C1 criteria, even fully controlled, might at first sight appear more pipeline dependent and, thus, less physical than a simple flux limit, we stress that any X-ray detection algorithm is bound to miss low luminosity clusters in a way that is pipeline dependent – the loss of efficiency not being a simple flux limit. For these reasons, we favour the concept of *controlled sample* (in the C1 selection sense) rather than of *complete flux-limited sample*.

Our D1 sample contains seven clusters displaying a flux in excess of $2 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ (all C1). This corresponds to a surface density of $\sim$8 deg$^{-2}$ and is larger than the 4–5 clusters deg$^{-2}$ implied by the RDCS log N–log S relation (Rosati et al. 1998) – the only *ROSAT* cluster sample complete to $2 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ (Rosati et al. 2002) – in addition to the shallow *XMM*/2dF survey (Gaga et al. 2005), reporting $7/2.3 = 3$ deg$^{-2}$ at the same flux limit. The probability to obtain 8 clusters deg$^{-2}$ from the RDCS number density is 1–10 per cent, assuming simple Poisson statistics. Given the relatively small fields covered in each case, the effects of cosmic variance upon any such comparison may well be important (the *XMM*/2dF survey and deepest regions of the RDCS cover 2.3 and 5 deg$^2$, respectively). For comparison, the simple cluster evolution plus cosmological model presented by Pacaud et al. (2006) predicts a surface density of $\sim$7.5 clusters deg$^{-2}$ displaying $T > 1$ keV and a flux $F > 2 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$.

We have computed reliable temperature values for nine of the 13 confirmed clusters – in particular, the C1 sample is 'temperature'

[9]With the caveat that the presence of giant arcs requires not only a large mass concentration but also a specific lens/source configuration.

[10]The function is normalized from local Universe observations and its evolution constrained by available high-$z$ data, numerical simulations and other possible prescriptions such as self-similarity evolution; one of the main unknowns being the role of non-gravitational physics in cluster evolution.

complete. This is important as it displays the potential for survey quality *XMM* observations to investigate the evolution of X-ray (and additional waveband) cluster scaling relations in a statistical manner over a wide, uniformly selected redshift interval. Together with the initial sample presented by Willis et al. (2005), the D1 sample is the first sample to investigate the $L_X$ versus $T_X$ relation for $1 < T < 2$ keV groups at $0.15 < z < 0.35$, for the simple reason that this population was previously undetected. *XMM*-LSS therefore samples a relatively complete, high surface density population of clusters displaying temperatures $T > 1$ keV at redshifts $z \lesssim 1$ and provides an important new perspective for the study the cluster and group evolution employing only moderate *XMM* exposure times.

All data presented in this paper – cluster images taken at X-ray and optical wavebands in addition to detailed results for the spectral and spatial analyses – are available in electronic form at the *XMM*-LSS cluster online data base: http://l3sdb.in2p3.fr:8080/l3sdb/login.jsp.

## ACKNOWLEDGMENTS

We are grateful to M. Arnaud for providing her profile convolution routine. We thank J. Ballet, R. Gastaud and J.-L. Sauvageot for useful discussions. AG acknowledges a CNES CDD position. GG and HQ thank the support of the FONDAP Centre for Astrophysics # 15010003. SA acknowledges financial contribution from contract ASI-INAF I/023/05/0. The results presented here are based on observations obtained with *XMM–Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA. The cluster optical images were obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the CFHT which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France and the University of Hawaii. This work is based in part on data products produced at TERAPIX and at the Canadian Astronomy Data Centre as part of the CFHTLS, a collaborative project of NRC and CNRS.

## REFERENCES

Andreon S., Valtchanov I., Jones L. R., Altieri B., Bremer M., Willis J., Pierre M., Quintana H., 2005, MNRAS, 359, 1250
Arnaud K. A., 1996, in Jacoby G. H., Barnes J., eds, ASP Conf. Ser. Vol. 101, Astronomical Data Analysis Software and Systems V. Astron. Soc. Pac., San Francisco, p. 17
Arnaud M., Majerowicz S., Lumb D., Neumann D. M., 2002, A&A, 390, 27
Arnaud M., Pointecouteau E., Pratt G. W., 2005, A&A, 441, 893
Böhringer H. et al., 2000, ApJS, 129, 435
Böhringer H. et al., 2001, A&A, 369, 826
Bondi M. et al., 2003, A&A, 403, 857
Borgani S. et al., 2001, ApJ, 561, 13
Burke D. J., Collins C. A., Sharples R. M., Romer A. K., Holden B. P., Nichol R. C., 1997, ApJ, 488, L83
Cavaliere A., Fusco-Femiano R., 1976, A&A, 49, 137
Chiappetti L. et al., 2005, A&A, 439, 413
Dickey J. M., Lockman F. J., 1990, ARA&A, 28, 215
Finoguenov A., Reiprich T. H., Böhringer H., 2001, A&A, 368, 749
Gaga T., Plionis M., Basilakos S., Georgantopoulos I., Georgakakis A., 2005, MNRAS, 363, 811
Ghizzardi S., 2002, *XMM*-SOC-CAL-022 & 029
Grevesse N., Sauval A. J., 1999, A&A, 347, 348
Helsdon S. F., Ponman T. J., 2003, MNRAS, 340, 485
Henry J. P., Gioia I. M., Mullis C. R., Voges W., Briel U. G., Böhringer H., Huchra J. P., 2001, ApJ, 553, L109
Ilbert O. et al., 2005, A&A, 439, 863
Kurtz M. J., Mink D. J., 1998, PASP, 110, 934
Markevitch M., 1998, ApJ, 504, 27
Moretti A., Campana S., Lazzati D., Tagliaferri G., 2003, ApJ, 588, 696
Moretti A. et al., 2004, A&A, 428, 21
Morrison R., McCammon D., 1983, ApJ, 270, 119
Osmond J. P. F., Ponman T. J., 2004, MNRAS, 350, 1511
Pacaud F., Pierre M. et al., 2006, MNRAS, in press (doi:10.1111/j.1365-2966.2006.10881.x) (this issue)
Pierre M. et al., 2004, J. Cosmology Astropart. Phys., 9, 11
Pratt G. W., Arnaud M., 2002, A&A, 394, 375
Rapetti D., Allen S. W., Weller J., 2005, MNRAS, 360, 555
Romer A. K. et al., 2000, ApJS, 126, 209
Rosati P., della Ceca R., Norman C., Giacconi R., 1998, ApJ, 492, L21
Rosati P., Borgani S., Norman C., 2002, ARA&A, 40, 539
Scharf C., 2002, ApJ, 572, 157
Smith R. K., Brickhouse N. S., Liedahl D. A., Raymond J. C., 2001, ApJ, 556, L91
Valtchanov I., Pierre M., Gastaud R., 2001, A&A, 370, 689
Valtchanov I. et al., 2004, A&A, 423, 75
Vikhlinin A., McNamara B. R., Forman W., Jones C., Quintana H., Hornstrup A., 1998, ApJ, 498, L21
Voit M., 2005, Rev. Mod. Phys., 77, 207
Willis J. P. et al., 2005, MNRAS, 363, 675

## APPENDIX A: ERROR BUDGET ON THE CLUSTER LUMINOSITY AND TEMPERATURE MEASUREMENTS

This appendix investigates the impact of specific sources of uncertainty on the cluster temperature and flux measurements. The calculations are performed for three clusters representative of the flux and temperature ranges covered by the present sample, namely XLSSC 41, 29, 13.

(1) Cluster luminosities are computed within $r_{500}$, a quantity empirically derived from the temperature (Section 3.3), while the photometric errors quoted in Table 5 results from the spatial fit only. In Table A1 we provide the uncertainty on the luminosity induced by the errors from the temperature measurements, as propagated through the derivation of $r_{500}$ (the spatial fit is then assumed to be exact). The results show that they are negligible compared to the accuracy level of the spatial fits.

(2) Cluster temperatures are estimated fixing the metal abundance to 0.3 solar (Section 3.1). Since many of our clusters have $T < 2$ keV, for which the contribution from emission lines is significant, some temperature-abundance degeneracy could occur in the spectral fitting – all the more so since the number of photons involved in the spectral fit is small. In Table A1 we provide further temperature measurements fixing the abundance to 0.1 or 0.6 solar. The results show that the impact on the derivation of the temperature and associated errors is negligible. In all these trials, the Galactic column density is held fixed.

(3) Finally, we investigate whether the contribution of unresolved AGNs to the integrated cluster emission is statistically significant at our sensitivity. We proceed here assuming the *XMM* $\log N$–$\log S$ for point sources (Moretti et al. 2003) since no information on the AGN environment of low-luminosity clusters is available to date. For each ring of the cluster spatial profile, we compute the limiting flux for which a point source is to be detected at the $3\sigma$ level. We then integrate the $\log N$–$\log S$ over the cluster area out to $R = 500$ kpc, between the varying flux limit and the background flux level. Results are gathered in Table A1 and show that, statistically, the point source contribution is negligible. In Appendix B we further inspect, for each cluster, the possibility that the core of the cluster emission could be contaminated by an AGN.

**Table A1.** Error budget for secondary effects. $r_{500}^{-}$ and $r_{500}^{+}$ are the extreme possible values for $r_{500}$ derived from the temperature uncertainties quoted in Table 4.

| XLSSC | 41 | 29 | 13 |
|---|---|---|---|
| source counts in $R_{\rm fit}, R_{\rm spec}$ | 819, 523 | 361, 311 | 383, 161 |
| $r_{500}^{-}, r_{500}, r_{500}^{+}$ | 171, 179, 196 arcsec | 66, 55, 82 arcsec | 76, 69, 84 arcsec |
| $\frac{L_{\rm X}(r_{500}^{-/+})-L_{\rm X}(r_{500})}{L_{\rm X}(r_{500})}$ | −2 per cent /+ 4 per cent | −1 per cent/+ 0.6 per cent | −5 per cent/+ 6 per cent |
| Ab = 0.1 | $T = 1.3$ (1.1, 1.5) keV | $T = 4.2$ (3.1, 6.1) keV | $T = 1.0$ (0.8, 1.2) keV |
| Ab = 0.3 | $T = 1.3$ (1.2, 1.6) keV | $T = 4.1$ (3.0, 5.8) keV | $T = 1.0$ (0.9, 1.2) keV |
| Ab = 0.6 | $T = 1.5$ (1.3, 1.7) keV | $T = 4.0$ (3.0, 5.5) keV | $T = 1.0$ (0.9, 1.3) keV |
| Undetected AGN contribution | | | |
| $R < 500$ kpc | 0.6 per cent | 0.5 per cent | 0.8 per cent |

## APPENDIX B: NOTES ON INDIVIDUAL CLUSTERS

In this appendix, we provide additional information for confirmed clusters. We have paid particular attention to whether the X-ray spectra and spatial emission of each cluster displays evidence for contamination by AGN emission. For each cluster we have compared the results generated by fitting the spatial emission having either included or excluded the central radial bin (radius of 3 or 6 arcsec). Although in some cases the best-fitting values of $\beta$ and $R_{\rm c}$ varied significantly, in all cases the integrated count rate within $r_{500}$ showed variations less than 10 per cent. The small number of photons contributed by the central few arcsec of each cluster prevented a separate spectral analysis of the central regions. Finally, we have investigated all examples of spatial coincidence between radio sources presented in Bondi et al. (2003) and the D1 cluster sample. We discuss individual clusters below.

In parallel, we have discovered a point-like source, XLSS J022528.2−041536, which is associated, at least in projection, with a group at redshift $z \sim 0.55$ (more than five concordant redshifts). However, the group emission, which is embedded in that of XLSSC 041, appears totally dominated by the point source and consequently, this object does not enter the C1 or C2 or C3 classes. A radio source ($S_{\rm tot} = 0.28$ mJy, $\sim 5$ arcsec extent) is present at 2 arcsec from the X-ray centre. XLSS J022528.2−041536 is located in the field of XLSSC 041 and the group member galaxies are indicated in the overlay of Fig. B1. In a further study involving optical, radio and IR data, we will assess the fraction of clusters that remain unclassified because of strong AGN contamination.

**XLSSC 029** The presence of a radio source ($S_{\rm tot} = 1.5$ mJy, $\sim$ 2 arcsec extent) some 11 arcsec from the cluster centre might suggest that the X-ray emission is contaminated by an AGN. However, examination of a recent 100 ks *XMM* pointing on this object (obs 0210490101, PI L. Jones) shows that the astrometry of the original *XMM*-LSS survey image is correct within 1 arcsec, and does not reveal a secondary maximum coincident with the radio source.

**XLSSC 044** This cluster is of very low surface brightness and displays elongated emission. Its X-ray morphology, galaxy distribution and temperature of $\sim$1 keV suggest a group in formation.

**XLSSC 025** A weak radio source ($S_{\rm tot} = 1.0$ mJy, $\sim$5 arcsec extent) lies at the centre of the cluster emission, however, the central galaxy spectrum displays no strong emission lines (rest-frame wavelength interval 3000–6000 Å is sampled). The X-ray spatial profile is peaked yet exclusion of the central 3 arcsec radial bin, does not change significantly the fitted value of $R_{\rm c}$ (it changes from 5 to 6 arcsec while $\beta = 0.44$ remains constant). We thus exclude any strong contamination by a central AGN, and favour the cool-core hypothesis.

**XLSSC 041** No radio source was identified within 30 arcsec of the cluster emission centroid.

**XLSSC 011** The X-ray and optical appearance of this system are suggestive of a compact group of galaxies. All point sources within a radius of 350 arcsec have been removed from the spatial analysis. However, it was not possible to estimate the extent of any contribution from the central galaxy to the group emission.

**XLSSC 005** No radio sources are identified within the projected area covered by the X-ray emission. This system displays double-peaked X-ray emission morphology, with each peak associated with a distinct velocity component. The photometric uncertainty for this system is large as less than 200 photons were available for the spatial fit which was limited to $R_{\rm fit} \sim r_{500}/2$. The system is further discussed in Valtchanov et al. (2004).

**XLSSC 013** No radio sources are identified within 1 arcmin of the X-ray centroid. A bright X-ray point source is present at 1 arcmin distance yet does not affect the cluster classification procedure and the point source is removed from the subsequent spatial and spectral analysis.

**XLSSC 022** The X-ray profile is peaked and here is a weak radio source within 2 arcsec of the X-ray emission centroid ($S_{\rm tot} =$ 0.15 mJy, no extent). The spectrum of the central galaxy displays no significant emission features consistent with AGN activity. The cluster morphology is very similar to XLSSC 025 in that $R_{\rm c}$ and $\beta$ values do not vary significantly depending upon the inclusion of the central bin in the spatial analysis. We therefore favour the cool-core hypothesis for this system.

**cluster a** The X-ray centroid coincides with a radio source ($S_{\rm tot} =$ 0.21 mJy, no extent) within 2 arcsec. The optical spectrum obtained for this object is faint and did not allow us to secure its redshift; however, no emission line is apparent in the spectrum. The photometric redshift is 0.98, with a SBI starburst as best-fitting spectrum. The flux at 24 μm is 0.4 mJy which is rather high. This suggests that the coincidence between this red object and the X-ray centroid might be fortuitous, but the contamination by an active nucleus cannot be excluded.

**cluster b** An X-ray point source is located some 30 arcsec from the X-ray emission centroid and was subsequently removed from the spatial and spectral analyses.

**cluster d** A radio source is located within 3 arcsec of the X-ray emission centroid ($S_{\rm tot} = 0.08$ mJy, no extent). The optical spectrum of the central galaxy does not show significant emission features. With an extent of 2.5 arcsec and a extent likelihood = 21 this marginal source is classified as C3.

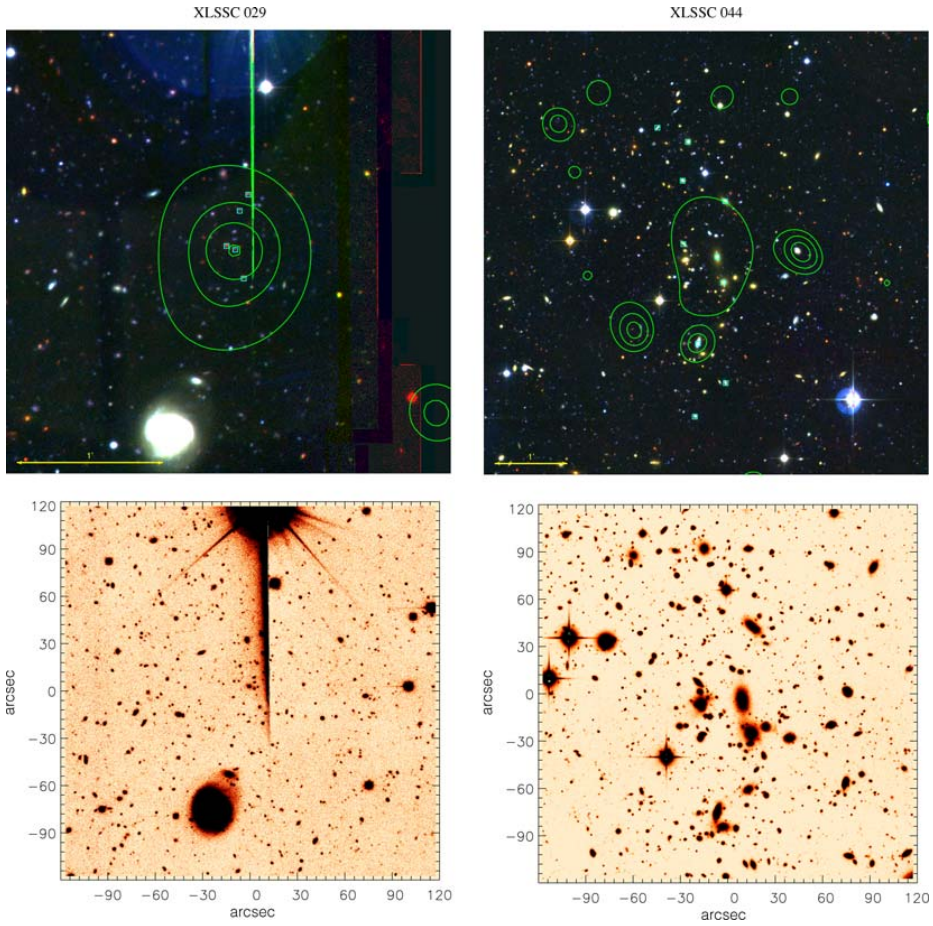


**Figure B1.** Colour overlays of the D1 clusters. Upper panel: cluster X-ray contours overlaid on a $u, r, z$ CFHTLS composite. Contours are drawn from the co-added [0.5–2] keV MOS1 + MOS2 + pn mosaic filtered in the wavelet space using a significance threshold of $10^{-3}$ for Poisson statistics (not corrected for vignetting). The intensity scale is logarithmic and the same for all images; values are: 2.6e−06, 5e−06, 9.6e−06, 1.8e−05, 3.5e−05, 6.7e−05, 0.00013 counts pixel$^{-1}$ s$^{-1}$. Boxes indicate galaxies whose redshift is in the $[z_{\rm cl} - 0.01, z_{\rm cl} + 0.01]$ range. The size of the images is 1.5 Mpc at the cluster redshift, except for the nearby group XLSSC011, for which the image is 0.6 Mpc aside. Lower panel: CFHTLS cluster I images, 4 arcmin aside.

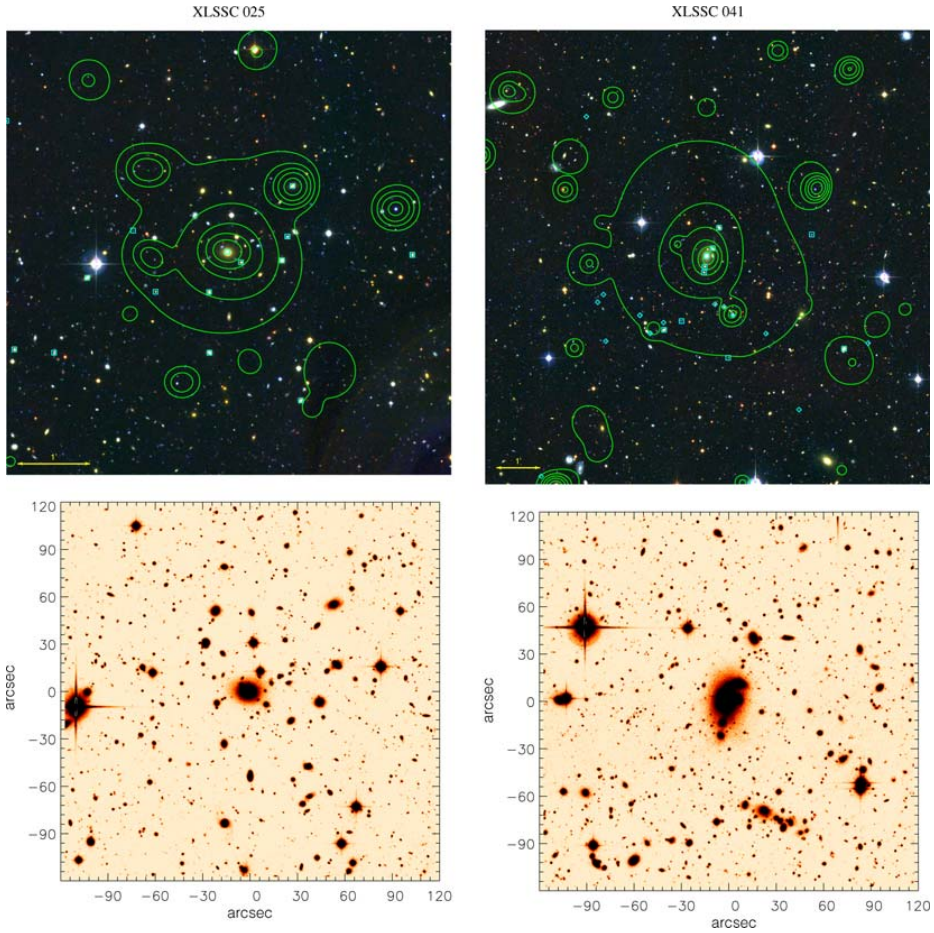


**Figure B1** – *continued*

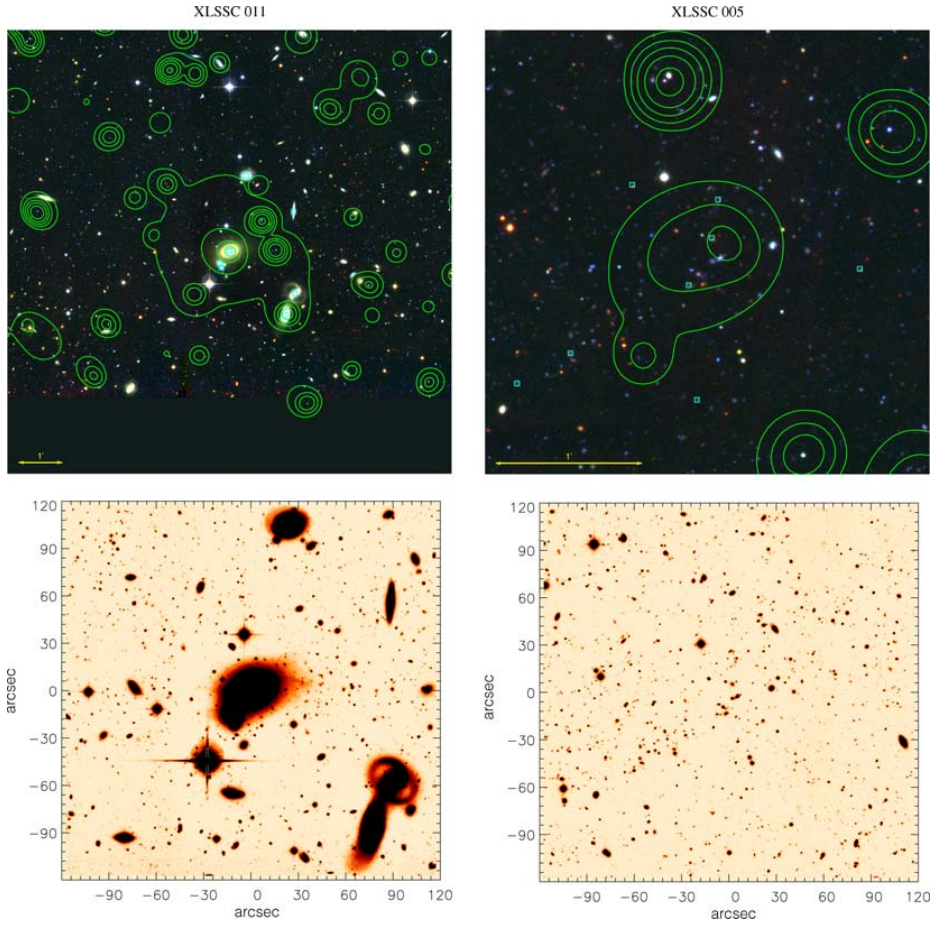


**Figure B1** – *continued*

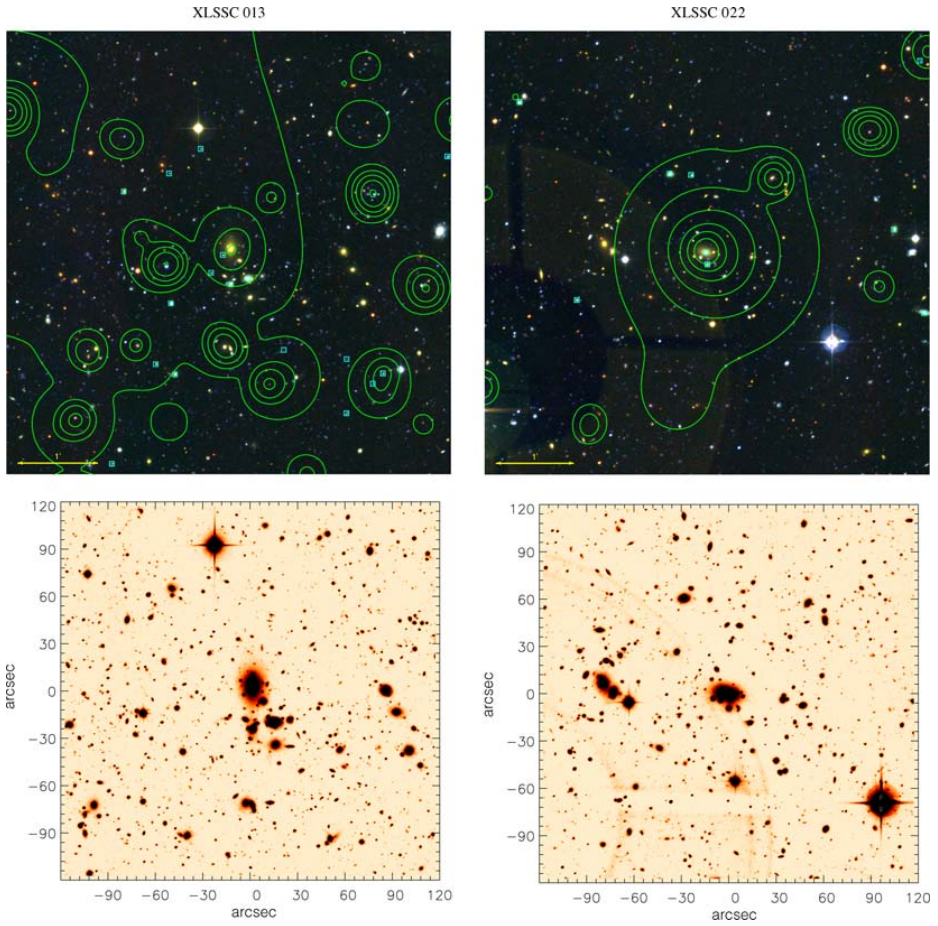


**Figure B1** – *continued*

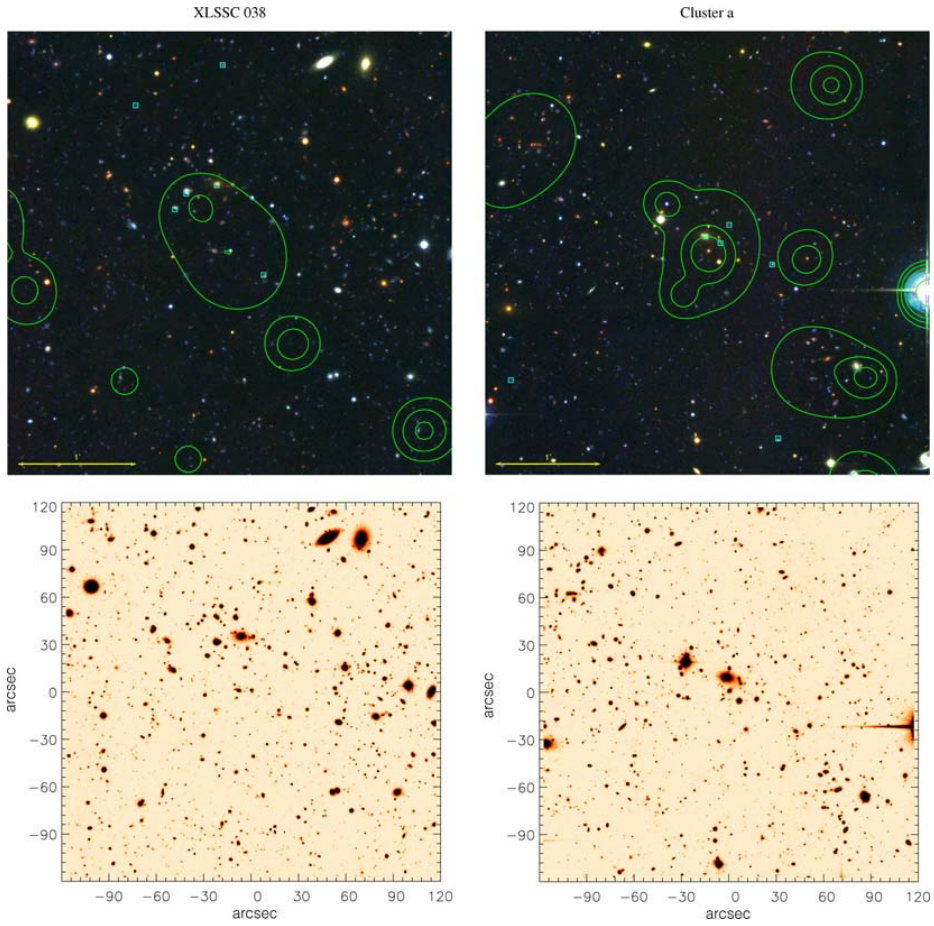


**Figure B1** – *continued*

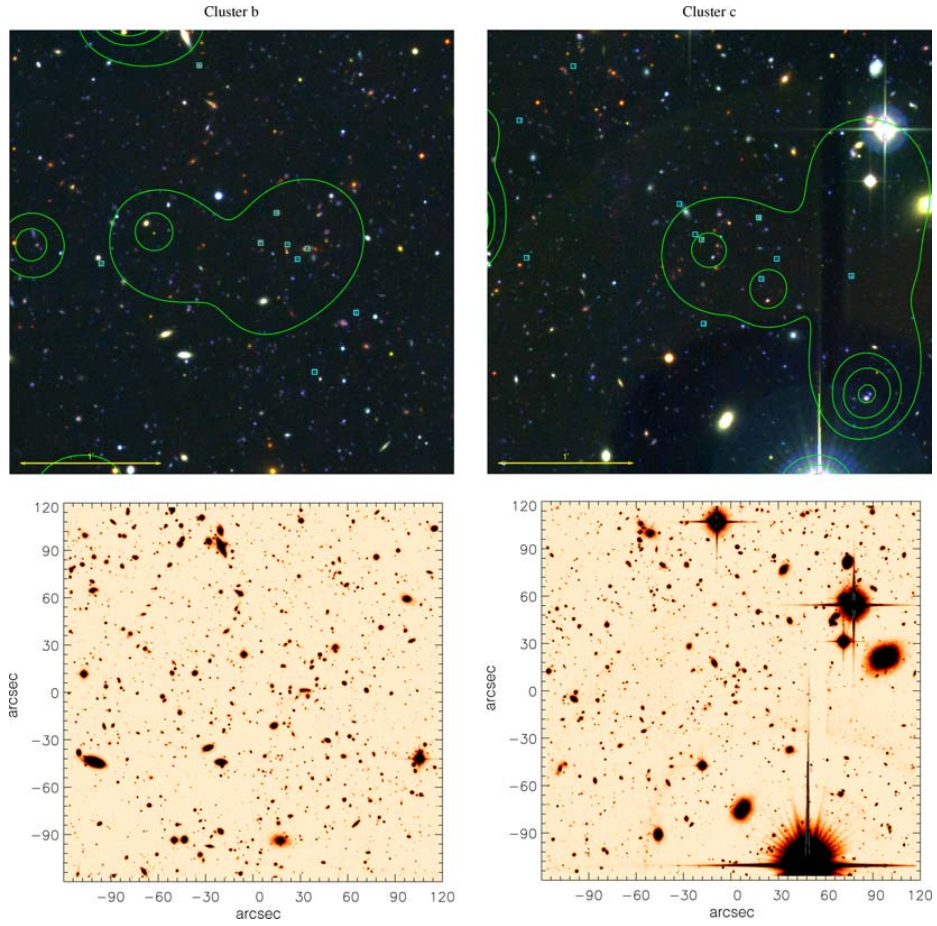


**Figure B1** – *continued*

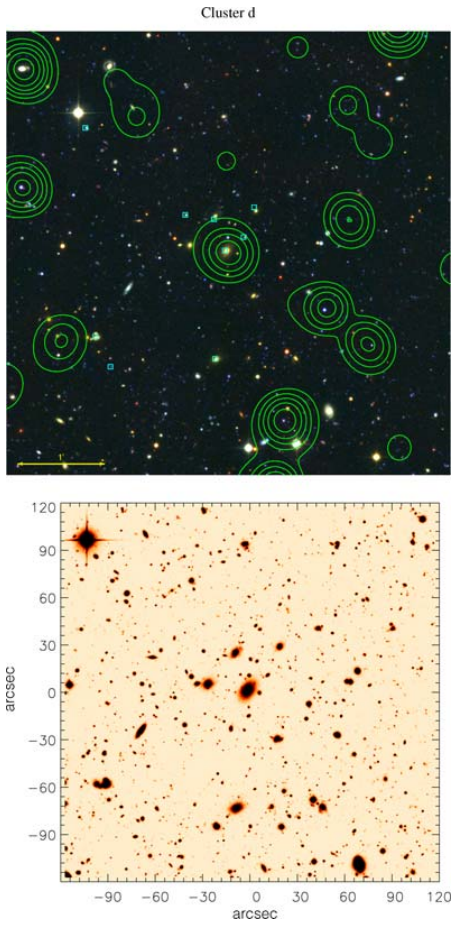


**Figure B1** – *continued*

This paper has been typeset from a TEX/LATEX file prepared by the author.

## 7.4 Le catalogue XMM-LSS :

## "The XMM-LSS catalogue : X-ray sources and associated optical data. Version I"

### Contexte/motivation

Les catalogues constitués sur l'ensemble des données AO1/AO2/GT sont rendus publics. A cet occasion, un article décrit les propriétés de l'échantillon complet ainsi que la manière de l'utiliser.

### Résumé

Faisant suite à la présentation des méthodes de détection de sources X du XMM-LSS (Pacaud et al., 2006), nous fournissons les listes de sources pour les premiers 5.5 $deg^2$ du relevé. Ces catalogues comprennent les bandes [0.5-2] et [2-10] keV et contiennent au total 3386 sources ponctuelles ou étendues ayant une vraissemblance de détection supérieure à 15 dans au moins une des bandes. La comparaison avec les Log(N)-Log(S) des relevés profonds est excellente. Les principaux paramètres mis à disposition sont la position, le taux de comptage, et l'extension de la source, ainsi que les valeurs de vraissemblance associées. Un ensemble de données additionelles telles que des corrections astrométriques et des flux sont de plus calculées, tandis que les erreurs sur la position et les taux de comptage sont estimées à l'aide de simulations. Nous décrivons la construction du catalogue multi-bandes, permettant une sélection rapide de sous-échantillon et une corrélation croisée aisée avec des catalogues multi-mongueurs d'onde externes. Un petit sous-échantillon optique multi-bandes du CFHTLS est associé à chaque source ainsi qu'une image superposée X/optique.

### Contribution personnelle

Outre la création des catalogues X sur lesquels l'article se base, j'ai interagi avec les administrateurs afin de les insérer dans les différentes bases de données électroniques du consortium XMM-LSS. J'ai également estimé les différents effets de sélection mentionnés dans l'article : complétude / biais photométrique / précision photométrique et astrométrique. Finalement, j'ai analysé la distribution en flux des sources [Log(N)-Log(S)] en la comparant aux travaux antérieurs.

Mon. Not. R. Astron. Soc. **382,** 279–290 (2007) doi:10.1111/j.1365-2966.2007.12354.x

# The *XMM*-Large Scale Structure catalogue: X-ray sources and associated optical data. Version I

M. Pierre,[1⋆] L. Chiappetti,[2] F. Pacaud,[1] A. Gueguen,[1] C. Libbrecht,[1] B. Altieri,[3] H. Aussel,[1] P. Gandhi,[4†] O. Garcet,[5] E. Gosset,[5] L. Paioro,[2] T. J. Ponman,[6] A. M. Read,[7] A. Refregier,[1] J.-L. Starck,[1,8] J. Surdej,[5] I. Valtchanov,[3] C. Adami,[9] D. Alloin,[1] A. Alshino,[6] S. Andreon,[10] M. Birkinshaw,[11] M. Bremer,[11] A. Detal,[5] P.-A. Duc,[1] G. Galaz,[12] L. Jones,[6] J.-P. Le Fèvre,[8] O. Le Fèvre,[9] D. Maccagni,[2] A. Mazure,[9] H. Quintana,[12] H. J. A. Röttgering,[13] P.-G. Sprimont,[5] C. Tasse,[13] G. Trinchieri[2] and J. P. Willis[14]

[1]*Laboratoire AIM, CEA/DSM - CNRS - Universit Paris Diderot, DAPNIA/Service d'Astrophysique, CEA-Saclay, F-91191 Gif-sur-Yvette Cedex, France*
[2]*INAF, IASF Milano, via Bassini 15, I-20133 Milano, Italy*
[3]*ESA, Villafranca del Castillo, Spain*
[4]*Institute of Astronomy, Madingley Road, Cambridge CB3 0HA*
[5]*Institut d'Astrophysique et de Géophysique, Université de Liège, Allée du 6 Août, 17, B5C, 4000 Sart Tilman, Belgium*
[6]*School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT*
[7]*Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH*
[8]*DAPNIA/SEDI CEA Saclay, 91191 Gif-sur-Yvette Cedex, France*
[9]*Laboratoire d'Astrophysique de Marseille, France*
[10]*INAF, Osservatorio Astronomico di Brera, Milan, Italy*
[11]*Department of Physics, University of Bristol, Tyndall Avenue, Bristol BS8 1TL*
[12]*Departamento de Astronomía y Astrofísica, Pontificia Universidad Católica de Chile, Santiago, Chile*
[13]*Leiden Observatory, PO Box 9513, 2300 RA Leiden, the Netherlands*
[14]*Department of Physics and Astronomy, University of Victoria, Elliot Building, 3800 Finnerty Road, Victoria, BC V8P 1A1, Canada*

Accepted 2007 August 15. Received 2007 June 20; in original form 2007 February 5

**ABSTRACT**

Following the presentation of the *XMM*-Large Scale Structure (*XMM*-LSS) survey X-ray source detection package by Pacaud et al., we provide the source lists for the first surveyed 5.5 deg$^2$. The catalogues pertain to the [0.5–2] and [2–10] keV bands and contain in total 3385 point-like or extended sources above a detection likelihood of 15 in either band. The agreement with deep log $N$– log $S$ is excellent. The main parameters considered are position, count rate, source extent with associated likelihood values. A set of additional quantities such as astrometric corrections and fluxes are further calculated while errors on the position and count rate are deduced from simulations. We describe the construction of the band-merged catalogue allowing rapid subsample selection and easy cross-correlation with external multiwavelength catalogues. A small optical Canada–France–Hawaii Telescope Legacy Survey multiband subset of objects are associated with each source along with an X-ray/optical overlay. We make the full X-ray images available in FITS format. The data are available at the Centre de Données de Strasbourg and, in a more extended form, at the Milan *XMM*-LSS survey data base.

**Key words:** catalogues – surveys – X-rays: general.

## 1 INTRODUCTION

The *XMM*-Large Scale Structure (*XMM*-LSS) survey has been designed to provide a well-defined statistical sample of X-ray galaxy clusters out to a redshift of unity, over a single large area, suitable for cosmological studies (Pierre et al. 2004). This requires the

⋆E-mail: mpierre@cea.fr
†Present address: RIKEN Cosmic Radiation Lab, 2-1 Hirosawa, Wakoshi, Saitama 351-0198, Japan.

**Table 1.** The individual *XMM*-LSS survey pointings. Quoted exposures are effective exposures computed after filtering high background periods; pointings with too low exposures (which have been re-executed during the XMM AO-5) are indicated by a star and are not included in the source list presented in this paper. Shifts are the astrometric corrections: $\Delta\mathrm{RA}/\cos(\mathrm{Dec.}) = \mathrm{RA}_{XMM} - \mathrm{RA}_{\mathrm{correct}}$, $\Delta\mathrm{Dec.} = \mathrm{Dec.}_{XMM} - \mathrm{Dec.}_{\mathrm{correct}}$ (see Section 2.3.3); the last column indicates whether the correction is based on CFHTLS W1 data (1), see Section 3, or on the USNO-A2 catalogue (2). The letter G refers to the Liège/Milan/Saclay Guaranteed Time (nominal exposure time: 20 ks) and the letter B to the observations performed during the Guest Observer phases (nominal exposure time: 10 ks).

| Field Id | | RA (J2000) | Dec. (J2000) | Exposure times (ks) | | | Astrometric correction | | |
|---|---|---|---|---|---|---|---|---|---|
| Internal | XMM ID | | | MOS1 | MOS2 | pn | ΔRA (arcsec) | ΔDec. (arcsec) | Origin |
| G01 | 112680101 | 02:27:25.4 | −04:11:06.4 | 24.6 | 25.3 | 21.4 | −1.60 | +0.53 | 1 |
| G02 | 112680201 | 02:25:54.2 | −04:09:05.6 | 10.1 | 9.7 | 6.7 | −1.07 | +0.00 | 1 |
| G03 | 112680301 | 02:24:45.6 | −04:11:00.8 | 21.8 | 21.7 | 17.3 | −1.07 | +1.07 | 1 |
| G04 | 109520101 | 02:23:25.3 | −04:11:07.6 | 25.5 | 25.8 | 19.5 | −1.60 | +0.53 | 1 |
| G05 | 112680401 | 02:28:05.1 | −04:31:08.1 | 23.5 | 23.9 | 12.5 | −1.07 | +1.07 | 1 |
| G06 | 112681301 | 02:26:34.4 | −04:29:00.8 | 16.4 | 16.6 | 10.5 | −0.53 | +0.53 | 1 |
| G07 | 112681001 | 02:25:25.3 | −04:31:07.1 | 22.5 | 25.1 | 18.6 | −0.53 | +0.00 | 1 |
| G08 | 112680501 | 02:23:54.6 | −04:29:00.1 | 21.2 | 21.3 | 15.9 | +0.00 | +1.07 | 1 |
| G09 | 109520601 | 02:22:45.2 | −04:31:11.1 | 22.5 | 22.7 | 16.4 | +0.00 | +0.53 | 1 |
| G10 | 109520201 | 02:27:25.4 | −04:51:04.4 | 24.7 | 24.6 | 18.5 | −1.07 | +1.07 | 1 |
| G11 | 109520301 | 02:26:05.1 | −04:51:06.1 | 21.7 | 21.8 | 16.1 | −1.06 | +0.53 | 1 |
| G12* | 109520401 | 02:24:45.4 | −04:51:11.2 | Not used | | | −0.54 | +0.53 | 1 |
| G13 | 109520501 | 02:23:13.1 | −04:49:03.1 | 23.6 | 23.9 | 17.8 | −2.67 | −0.53 | 1 |
| G14 | 112680801 | 02:22:04.1 | −04:51:09.7 | 14.4 | 14.1 | 8.3 | +1.06 | +0.53 | 1 |
| G15 | 111110101 | 02:27:54.1 | −05:09:02.3 | 20.8 | 21.8 | 14.0 | −2.67 | +0.00 | 1 |
| G16a* | 111110201 | 02:26:34.2 | −05:09:03.1 | Not used | | | −1.60 | +0.53 | 1 |
| G16b | 111110701 | 02:26:35.2 | −05:08:46.6 | 11.9 | 11.9 | 11.5 | +0.00 | +0.53 | 1 |
| G17 | 111110301 | 02:25:14.3 | −05:09:08.4 | 22.4 | 22.2 | 17.5 | −2.67 | −0.53 | 1 |
| G18 | 111110401 | 02:23:54.1 | −05:09:09.7 | 27.7 | 28.0 | 19.2 | −2.67 | +0.00 | 1 |
| G19 | 111110501 | 02:22:34.0 | −05:09:02.1 | 23.2 | 23.8 | 16.8 | −2.67 | +0.00 | 1 |
| B01 | 037980101 | 02:22:45.5 | −03:50:58.8 | 14.1 | 14.2 | 8.3 | −2.13 | +0.00 | 1 |
| B02 | 037980201 | 02:24:05.6 | −03:51:00.0 | 13.2 | 13.2 | 7.8 | −2.13 | +0.00 | 1 |
| B03 | 037980301 | 02:25:25.7 | −03:50:59.2 | 13.3 | 13.0 | 7.9 | −1.07 | +0.53 | 1 |
| B04* | 037980401 | 02:26:45.4 | −03:51:00.1 | Not used | | | −1.07 | +0.53 | 1 |
| B05 | 037980501 | 02:28:05.4 | −03:51:00.5 | 15.7 | 15.7 | 10.5 | −0.53 | +1.06 | 1 |
| B06 | 037980601 | 02:22:05.6 | −03:31:00.2 | 13.2 | 13.2 | 7.7 | −1.60 | +0.53 | 2 |
| B07 | 037980701 | 02:23:25.7 | −03:30:56.7 | 12.3 | 12.3 | 6.9 | −3.73 | +0.00 | 2 |
| B08 | 037980801 | 02:24:34.3 | −03:29:05.1 | 10.6 | 11.5 | 6.3 | −0.53 | +1.07 | 2 |
| B09 | 037980901 | 02:26:05.4 | −03:31:01.1 | 13.9 | 13.8 | 8.9 | −0.54 | +0.53 | 2 |
| B10 | 037981001 | 02:27:14.2 | −03:28:58.7 | 13.0 | 13.2 | 8.6 | −1.07 | +2.13 | 2 |
| B11 | 037981101 | 02:22:34.2 | −03:09:02.5 | 12.4 | 12.3 | 7.9 | −3.20 | +1.07 | 2 |
| B12 | 037981201 | 02:23:54.4 | −03:09:04.1 | 11.4 | 11.3 | 7.0 | −3.20 | +0.00 | 2 |
| B13* | 037981301 | 02:25:14.4 | −03:08:57.4 | Not used | | | −0.53 | +1.60 | 2 |
| B14 | 037981401 | 02:26:34.4 | −03:08:57.6 | 12.6 | 12.5 | 6.6 | −0.53 | +3.73 | 2 |
| B15 | 037981501 | 02:27:54.6 | −03:08:59.3 | 11.4 | 11.5 | 8.4 | −2.13 | +1.07 | 2 |
| B16 | 037981601 | 02:23:14.4 | −02:48:56.3 | 12.3 | 12.7 | 7.9 | −1.60 | +2.13 | 2 |
| B17* | 037981701 | 02:24:34.8 | −02:48:50.0 | Not used | | | -2.67 | +3.73 | 2 |
| B18 | 037981801 | 02:25:55.0 | −02:48:49.1 | 16.0 | 16.2 | 11.4 | −1.07 | +2.13 | 2 |
| B19 | 037981901 | 02:27:14.9 | −02:48:49.5 | 13.1 | 13.1 | 8.7 | −1.07 | +2.13 | 2 |
| B20 | 037982001 | 02:20:34.8 | −04:48:46.3 | 16.1 | 16.8 | 11.7 | −1.60 | −0.53 | 1 |
| B21 | 037982101 | 02:21:14.9 | −04:28:45.9 | 15.6 | 15.6 | 9.9 | −2.13 | +0.53 | 1 |
| B22* | 037982201 | 02:22:05.5 | −04:11:03.9 | Not used | | | +0.00 | +0.00 | 1 |
| B23 | 037982301 | 02:20:05.5 | −04:31:03.9 | 9.0 | 9.6 | 6.5 | −1.07 | +1.60 | 1 |
| B24 | 037982401 | 02:20:45.6 | −04:11:01.2 | 18.1 | 18.2 | 10.4 | −1.07 | +0.53 | 1 |
| B25 | 037982501 | 02:21:25.5 | −03:51:02.4 | 11.5 | 11.9 | 6.8 | −1.07 | −0.53 | 1 |
| B26 | 037982601 | 02:20:34.8 | −03:28:50.4 | 14.7 | 14.6 | 8.3 | −2.13 | +0.53 | 2 |
| B27 | 037982701 | 02:21:14.8 | −03:08:49.1 | 16.3 | 16.2 | 10.7 | −1.60 | +1.07 | 2 |
| B28 | 147110101 | 02:21:55.1 | −02:49:02.4 | 11.4 | 11.6 | 9.1 | −1.60 | +0.00 | 2 |
| B29 | 147110201 | 02:20:35.1 | −02:49:00.1 | 10.6 | 10.6 | 8.2 | −1.60 | +0.00 | 2 |
| B30 | 147111301 | 02:21:15.1 | −05:09:00.7 | 12.4 | 12.5 | 9.9 | −1.07 | +0.53 | 1 |
| B31 | 147111401 | 02:20:36.3 | −05:28:59.5 | 10.8 | 11.0 | 8.4 | −0.53 | +0.53 | 1 |
| B32* | 147111501 | 02:21:56.0 | −05:28:56.4 | Not used | | | −3.20 | +0.53 | 1 |

ability to detect and characterize faint extended sources, in such a way as to control both the selection effects and the contamination by spurious or misclassified point-like sources. For this purpose, we have developed a dedicated X-ray image processing package, XAMIN, which is adapted to the complex characteristics of the *XMM* focal plane (Pacaud et al. 2006). It is a two-step procedure combining wavelet multi-resolution analysis and maximum-likelihood fits, both using Poisson statistics. The package has been extensively

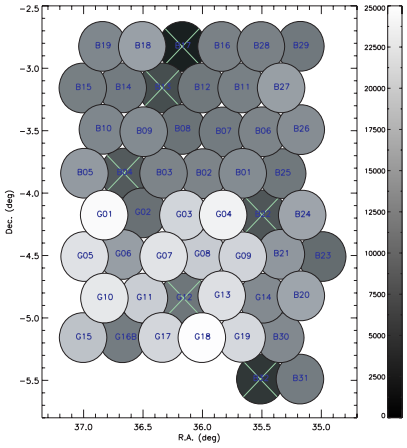


**Figure 1.** Layout of the Guaranteed Time, AO-1 and AO-2 pointings. The grey-scale indicates the effective exposure times. Pointings not included in the present release (see Table 1) are flagged by a cross. The total resulting geometrical area, taking a radius of 13 arcmin for the pointing useful radius (Section 2.3.1) amounts to 5.5 deg$^2$.

**Table 2.** XAMIN output parameters (Pacaud et al. 2006). See Pacaud et al. (2006) for the statistical definition of the derived likelihood values.

| Parameter | Notes | Content |
|---|---|---|
| CUTRAD | | Size of the fitting box |
| EXP | *b* | Mean exposure time in the box |
| GAPFLAG | *b* | Distance to the nearest CCD gap |
| GAP_NEIGHBOUR | | Distance to the nearest detected neighbour in the fitting box |
| EXT | | Best-fitting core radius |
| EXT_LH | | Extension likelihood |
| DET_LH | *a* | Detection likelihood |
| X_IMA,Y_IMA | *a* | Best-fitting position in pixel |
| RA,DEC | *a* | Best-fitting sky coordinates |
| RATE_MOS | *a* | EPIC-MOS count rate |
| RATE_PN | *a* | EPIC-pn count rate |
| SCTS_MOS | *a* | Estimated source counts in MOS1 + 2 |
| SCTS_PN | *a* | Estimated source counts in pn |
| BG_MAP_MOS | *a* | Background level in MOS1 + 2 |
| BG_MAP_PN | *a* | Background level in pn |
| PIX_DEV | *a* | Distance between input/output positions |
| N_ITER | *a* | Number of iterations in the fit |

*Notes*. [a]Computed for both point-like and extended profile fits. [b]Issued for each of the three EPIC detectors.

tested and its parameters adjusted by means of simulations, for the extended (clusters) and point-like [active galactic nucleus (AGN)] source populations; the latter representing some 95 per cent of the X-ray sources at our sensitivity of $\sim 4\,10^{-15}$ erg s$^{-1}$ cm$^{-2}$ in the [0.5–2] keV band.

The first reports on cluster and AGN populations, based on the XAMIN products and an associated selection function, were published by Pierre et al. (2006) and Gandhi et al. (2006), respectively.

**Table 3.** Mean total background level (MOS1 + MS2 + pn) for the pointings included in the present catalogue, as a function of off-axis angle. Units are $10^{-6}$ count s$^{-1}$ pixel$^{-1}$ for a pixel size of 2.5 arcsec.

| Off-axis (arcmin) | [0.5–2] keV | [2–10] keV |
|---|---|---|
| 0–5 | 7.1 | 12.7 |
| 5–10 | 6.3 | 12.4 |
| 10–13 | 5.8 | 12.2 |

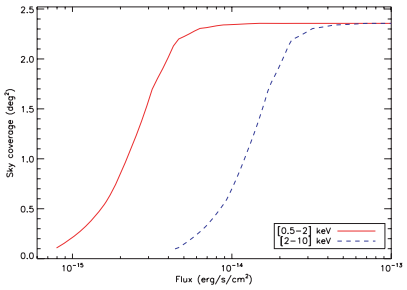


**Figure 2.** Sky coverage for the [0.5–2] and [2–10] keV bands corresponding to 27 pointings of 10 ks. The curve in the soft band is derived from extensive simulations of point sources and includes only sources having a detection likelihood >15 within off-axis distances less than 10 arcmin. The hard curve is scaled from the soft one on the basis of an equivalent S/N, using adequate background, vignetting and PSF characteristics.

The guaranteed time pointings (G fields in Table 1) were previously analysed with an independent, more traditional pipeline analogous to the one used for the HELLAS2XMM survey (Baldi et al. 2002), as described in Chiappetti et al. (2005), providing also the standard reference for the Milan *XMM*-LSS survey data base.

The *XMM*-LSS survey, located around $2^h30^m$, $-5^\circ$, is associated in the optical with the Wide Synoptic component of the Canada–France–Hawaii Telescope Legacy Survey[1] (CFHTLS – W1).

In this paper, we present the source lists obtained for the first 45 *XMM*-LSS survey pointings processed by XAMIN along with the relevant information on the catalogues (Section 2) as well as the X-ray images (Section 3.1). Section 3.2 describes the associated optical data that we make publicly available. The online facilities and plans for future releases are presented in the last two sections.

## 2 THE X-RAY SOURCE LISTS

### 2.1 List of available pointings

The *XMM*-LSS survey pointings pertaining to the Guaranteed Time, AO-1 and AO-2 periods are listed in Table 1 and displayed in Fig. 1. The raw X-ray observation data files (ODFs) were reduced using the standard *XMM* Science Analysis System (XMMSAS, v6.1) tasks EMCHAIN and EPCHAIN for the MOS and pn detectors, respectively. High-background periods, related to soft protons, were excluded

[1] http://cfht.hawaii.edu/Science/CFHLS/.

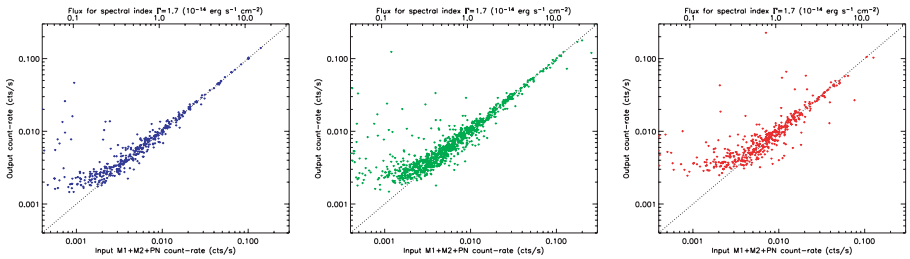


**Figure 3.** Photometric accuracy for three ranges of off-axis values from 10-ks simulated images in the soft band: 0–5, 5–10 and 10–13 arcmin. 'Count-rate' is the measured MOS1 + MOS2 + pn rate, normalized to the on-axis value.

from the event lists following the procedure outlined by Pratt & Arnaud (2002). The resulting light curves were visually inspected. After this operation, a number of pointings appeared to be not usable for the purpose of the *XMM*-LSS survey (too-low effective exposure times) and, consequently, are not included in the present version of the catalogue. The fields in question were re-observed during AO-5 (2006 July–2007 January) and will be published in a subsequent data release.

### 2.2 Processing and XAMIN output source parameters

Photon images in different energy bands are then created with a scale of 2.5 arcsec pixel$^{-1}$. Images from the three detectors (pn, MOS1, M0S2), pertaining to the same band, are co-added. The resulting image is in turn filtered in wavelet space[2] to remove the noise at a given significance level and subsequently scanned by a source detection algorithm set to a low threshold to obtain a primary source list. Detailed properties of each detected source are further assessed from the individual photon images using XAMIN, a maximum-likelihood profile-fitting procedure[3] designed for the *XMM*-LSS survey. The specific goal of this second pass is to monitor in a clean and systematic way the characterization of extended X-ray sources and associated selection effects. The principle and performances of XAMIN, for extended and point-like sources, are described in detail by Pacaud et al. (2006) and we recall here the main lines of the procedure. Basically, for each source, two spatial emission models convolved by the *XMM* point spread function (PSF) are tested: (pnt) a point source and (ext) a $\beta$-profile assuming a constant slope of 2/3. The main parameters returned are the position and the fitted count rates and, for model ext, the best estimate of the core radius of the $\beta$-model. Further, for each model, the likelihood[4] of the source is computed, as well as the likelihood of the extension for model ext. The fits are simultaneously performed on the pn and on the MOS1 + MOS2 sum images only requiring the source centre to be the same in both images.[5] The procedure takes into account all main technical characteristics such as: the blurring of the PSF and the vignetting as a function of off-axis angle and photon energy, the gaps between the CCD of the pn and MOS arrays, and the various background contributions. Results of the ext fit allow us to define two classes of extended objects in the XAMIN parameter space: the C1 and C2 classes (see Section 2.3.4). Table 2 summarizes the output parameters of the pipeline. Table 3 provides the background level, averaged over all pointings included in the present catalogue. The measurements were performed on the X-ray images (obtained after removing bad time intervals) by masking the detected source regions, for three ranges of an off-axis distance. We recall that the background consists of two components: (1) the cosmic background which is affected by the vignetting and (2) the particle background, which is significant at high energy and not affected by the vignetting. As a rule of thumb, a point-like source detected with 15 net photons in any of the three rings shows a signal-to-noise ratio (S/N) of $\sim$3.6 and 3.3 for the [0.5–2] and [2–10] keV bands, respectively (when considering a circular aperture having a radius of 10 arcsec).

### 2.3 Source list description

In this paper, we present the source lists for two bands, [0.5–2] and [2–10] keV, named B and CD, respectively. Tables 4 and 5 display the parameters we make available: in addition to the XAMIN output, a number of parameters are a posteriori calculated in order to facilitate the interpretation of the data set. In its present state, XAMIN does not perform error calculations. Mean statistical errors were estimated by means of extensive simulations. Pacaud et al. (2006) presented a detailed account of uncertainty estimates for the extended sources (count rate and core radius). In this paper, further error information is provided for the point-source population; we note that only the first two digits are to be considered significant for the count rate and for the core radius as well as for the derived quantities.

#### *2.3.1 Thresholds*

XAMIN processes sources only out to an off-axis distance of 13 arcmin by applying a detection mask centred on the mean optical axis of the three telescopes, considering in this way only sources that are visible by the three detectors. The total geometrical area of the present catalogue corresponding to the validated pointings indicated in Fig. 1 amounts to 5.5 deg$^2$. The present version of the B and CD band catalogues contains the extended sources identified as C1 and C2 (see Section 2.3.4) to which are added sources having a point-source detection likelihood (LH) greater than 15 (so-called *non-spurious*). As shown in Fig. 2, this ensures a 90 per cent completeness limit of $4 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ in the B band within

[2] À TROU algorithm, combined with a Poisson noise model; see Starck & Pierre (1998).
[3] The likelihood is computed with respect to a flat image.
[4] The values actually provided are the natural logarithm of the likelihoods but, to follow the widespread usage, we will call them simply 'likelihood' throughout this paper. We refer to Pacaud et al. (2006, section 2.3.1) for clarification of this terminology issue.
[5] In the case of model ext, an additional fitting condition imposes that the core radii inferred from the pn and MOS1 + MOS2 images are the same.

**Table 4.** List of parameters provided in the public *XMM*-LSS survey catalogue. All are available at the *XMM*-LSS survey Milan data base in separate tables XLSSB for the soft band and XLSSCD for the hard band (the column name has an appropriate prefix: when there are two column names given, one with the prefix B and one with the prefix CD, only the one applicable to the given band appears in the relevant table but both may show up in the band-merged table; column names without prefix are relevant to the individual band only). The last four columns indicate respectively: (X) whether a parameter is natively computed by XAMIN; (m) whether a parameter is available also in the band-merged table; (o) whether a parameter is present in the XLSSOPT table together with those described in Table 10; and (C) whether a parameter is present in the catalogue stored at CDS.

| Column name | Units | Meaning and usage | X | m | o | C |
|---|---|---|---|---|---|---|
| Bseq or CDseq | – | Internal sequence number (unique) | | X | X | X |
| Bcatname or CDcatname | – | IAU catalogue name XLSS*x* Jhhmmss.s-ddmmss, *x*=B or CD | | X | X | X |
| Xseq | – | Pointer to merged entry see Table 5 | | X | X | X |
| Xcatname | – | Pointer to merged entry see Table 5 | | X | X | X |
| Xfield | – | *XMM* pointing 1-32 for B01-B32 1001-1019 for G01-G19 | | X | | |
| expm1 | s | MOS1 camera exposure in the band | X | | | |
| expm2 | s | MOS2 camera exposure in the band | X | | | |
| exppn | s | pn camera exposure in the band | X | | | |
| gapm1 | arcsec | MOS1 distance to nearest gap | X | | | |
| gapm2 | arcsec | MOS2 distance to nearest gap | X | | | |
| gappn | arcsec | pn distance to nearest gap | X | | | |
| Bnearest or CDnearest | arcsec | Distance to nearest detected neighbor | X | X | | |
| Bc1c2 | 0–1–2 | 1 for class C1, 2 for C2, 0 for undefined | | X | X | X |
| CDc1c2 | 0–1–2 | 1 for class C1, 2 for C2, 0 for undefined | | | | |
| Bcorerad or CDcorerad | arcsec | Core radius EXT (for extended sources) | X | X | | X |
| Bextlike or CDextlike | – | Extension likelihood EXT_LH | X | X | | X |
| Bdetlik_pnt or CDdetlik_pnt | – | Detection likelihood DET_LH for pointlike fit | X | | | |
| Bdetlik_ext or CDdetlik_ext | – | Detection likelihood DET_LH for extended fit | X | | | |
| Boffaxis or CDoffaxis | arcmin | off-axis angle | | X | | X |
| Brawra_pnt or CDrawra_pnt | ° | Source RA (not astrometrically corrected) for point-like fit | X | | | |
| Brawdec_pnt or CDrawdec_pnt | ° | Source Dec. (not astrometrically corrected) for point-like fit | X | | | |
| Brawra_ext or CDrawra_ext | ° | Source RA (not astrometrically corrected) for extended fit | X | | | |
| Brawdec_ext or CDrawdec_ext | ° | Source Dec. (not astrometrically corrected) for extended fit | X | | | |
| Bra_pnt or CDra_pnt | ° | Source RA (astrometrically corrected) for point-like fit | X | | | |
| Bdec_pnt or CDdec_pnt | ° | Source Dec. (astrometrically corrected) for point-like fit | X | | | |
| Bra_ext or CDra_ext | ° | Source RA (astrometrically corrected) for extended fit | X | | | |
| Bdec_ext or CDdec_ext | ° | Source Dec. (astrometrically corrected) for extended fit | X | | | |
| Bposerr or CDposerr | arcsec | Error on coordinates according to Table 8 | | X | | X |
| Bratemos_pnt or CDratemos_pnt | counts $s^{-1}$ | MOS count rate for pointlike fit | X | | | |
| Bratepn_pnt or CDratepn_pnt | counts $s^{-1}$ | pn count rate for pointlike fit | X | | | |
| Bratemos_ext or CDratemos_ext | counts $s^{-1}$ | MOS count rate for extended fit | X | | | |
| Bratepn_ext or CDratepn_ext | counts $s^{-1}$ | pn count rate for extended fit | X | | | |
| countmos_pnt | counts | MOS number of counts for point-like fit | X | | | |
| countpn_pnt | counts | pn number of counts for point-like fit | X | | | |
| countmos_ext | counts | MOS number of counts for extended fit | X | | | |
| countpn_ext | counts | pn number of counts for extended fit | X | | | |
| bkgmos_pnt | counts $pixel^{-1}$ | MOS local background for point-like fit | X | | | |
| bkgpn_pnt | counts $pixel^{-1}$ | pn local background for point-like fit | X | | | |
| bkgmos_ext | counts $pixel^{-1}$ | MOS local background for extended fit | X | | | |
| bkgpn_ext | counts $pixel^{-1}$ | pn local background for extended fit | X | | | |
| Bflux or CDflux | erg $cm^{-2}$ $s^{-1}$ | source flux (undefined i.e. −1 for extended) | | X | | X |
| Bfluxflag or CDfluxflag | 0 to 2 | 0 if MOS-pn difference <20 per cent, 1 between 20–50 per cent, 2 above 50 per cent | | X | | X |

an off-axis distance of 10 arcmin for the 10-ks pointings (beyond 10 arcmin, the sensitivity falls below 50 per cent of the value at the centre); in average, this corresponds to 85 and 70 sources per 20 and 10 ks pointing, respectively, judged to be real within the inner 13 arcmin.

### 2.3.2 *Count rate and flux*

XAMIN fits the MOS and pn count rates independently[6] (assuming that the two MOS have the same response). We performed a number of point-source tests by simulating a set of 30 images in which the $\log N$– $\log S$ distribution was injected. We display in Fig. 3 the full range of simulations detailing the photometric accuracy out to an off-axis distance of 13 arcmin (Pacaud et al. 2006, presented averaged performances). The Eddington bias is obvious for faint sources and increases with the off-axis distance; in particular, the photometry appears to be unreliable for a number of sources detected with a total count rate (pn + MOS1 + MOS2) between 0.002 and 0.005 count $s^{-1}$ at an off-axis angle of 10–13 arcmin. These sources are, however, detected with a likelihood greater than 15, indicating that they are real. They are thus left in the public source catalogue but their photometry should be handled with caution. Mean values for the photometric bias and accuracy are summarized in Table 6.

[6] While the source position and extent are fixed to be the same for the three detectors.

**Table 5.** List of data base parameters, as Table 4, but for the additional columns present only in the merged catalogue table XLSS. When there are two column names given, one with the prefix B and one with the prefix CD, they relate to the given band, and both show up in the band-merged table. Column names with the prefix X are relevant to merged properties.

| Column name | Units | Meaning and usage | X | m | o | C |
|---|---|---|---|---|---|---|
| Xseq | – | Internal sequence number (unique) | | X | X | X |
| Xcatname | – | IAU catalogue name XLSS Jhhmmss.s-ddmmss*c*, see Section 2.3.7 | | X | X | X |
| Bspurious and CDspurious | 0 or 1 | Set to 1 when soft/hard component has DET_LH<15 | | X | | |
| Bdetlike and CDdetlike | – | Detection likelihood EXT_LH (pnt or ext according to source class) | X | X | | X |
| Xra | ◦ | Source RA (astrometrically corrected) (pnt or ext according to source class and in best band) | | X | X | X |
| Xdec | ◦ | Source Dec. (astrometrically corrected) (pnt or ext according to source class and in best band) | | X | X | X |
| Bra and CDra | ◦ | Source RA (astrometrically corrected) (pnt or ext according to source class) | | X | X | X |
| Bdec and CDdec | ◦ | Source Dec. (astrometrically corrected) (pnt or ext according to source class) | | X | X | X |
| Xbestband | 2 or 3 | Band with highest likelihood : 2 for B, 3 for CD | | X | | |
| Xastrocorr | 1 or 2 | Astrometric correction from CFHTLS (1) or USNO (2) | | X | | |
| Xmaxdist | arcsec | Distance between B and CD positions | | X | | |
| Xlink | – | Pointer to Xseq of secondary association, see Section 2.3.7 | | X | | |
| Bratemos and CDratemos | counts $s^{-1}$ | MOS count rate (pnt or ext according to source class) | X | X | | X |
| Bratepn and CDratepn | counts $s^{-1}$ | pn count rate (pnt or ext according to source class) | X | X | | X |

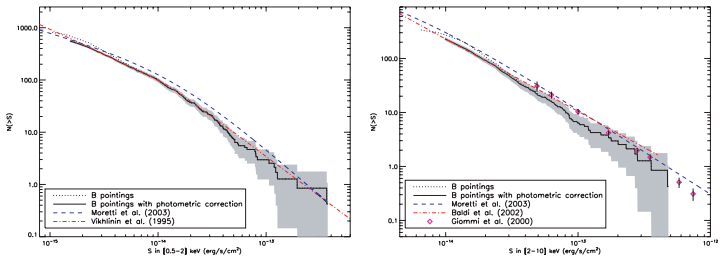


**Figure 4.** The log $N$– log $S$ distributions for the soft (left-hand panel) and hard (right-hand panel) bands, involving the 27 pointings with ∼10 ks. Only sources having a detection likelihood >15 and an off-axis distance less than 10 arcmin are considered. The dotted line corresponds to the raw counts and the solid line is corrected for the Eddington bias (as conspicuous in fig. 7 of Pacaud et al. 2006). The dashed region indicates the $1\sigma$ fluctuation level.

Count-rate values are in turn converted into fluxes assuming a standard power-law spectrum (photon index of 1.7) and the mean $N_{\rm H}$ value of the region ($2.6\,10^{20}$ cm$^{-2}$). The energy conversion factors are given in Table 7. The observed log $N$–log $S$ distributions are presented in Fig. 4; they are in good agreement with the Vikhlinin et al. (1995) (*ROSAT*), Moretti et al. (2003) (*ROSAT*, *Chandra* and *XMM*), Baldi et al. (2002) (*XMM*) and Giommi, Perri & Fiore (2000) (*BeppoSAX*) data points.

Note that the resulting pn and MOS fluxes may be quite different for some sources. In most of the cases, this is due to the fact that a part of the source is occulted by a CCD gap in one of the detectors to the point where the information is not recoverable. Such cases can be identified from the gap-related columns (Tables 2 and 4) and it is suggested to use the fitted parameters obtained from the detectors on which the source is not affected by a gap (if any). For the sake of simplicity, a single mean flux value is provided ([FLUX(MOS) + FLUX(pn)]/2)[7] along with column fluxflag indicating the difference between the fluxes inferred from the MOS1 + MOS2 combination and the pn (0: less than 20 per cent; 1: between 20 and 50 per cent; 2: greater than 50 per cent). For the two single-band catalogues, the sources are roughly equally distributed between the three categories of flux quality. Fluxes assuming a thermal spectrum as well as temperature and luminosity information for the extended sources classified as C1 and C2, confirmed as clusters, can be found in the *XMM*-LSS survey cluster data base.[8]

### 2.3.3 *Positional accuracy and astrometric corrections*

The positional accuracy for the point-source population was also estimated from the simulations. Results are displayed in Fig. 5 and summarized in Table 8. We recall here that for the point-like fit, as explained by Pacaud et al. (2006), positions are fixed to those of the first pass catalogue (SExtractor detections on the wavelet filtered image).

In parallel, in order to compensate for possible inaccuracies in the *XMM* pointing positions, a global rigid astrometric correction was estimated using the SAS task EPOSCORR. We generated, for each *XMM* pointing, a reference X-ray list with all 'non spurious' point-like sources along with a reference optical list containing all CFHTLS W1 objects within 6 arcsec of the X-ray objects, being brighter than

[7] As the MOS count rate is about 1/3 of the pn count rate for cluster spectra, the averaged MOS flux involves a number of photons comparable to that of the pn, thus justifying the simple mean.

[8] http://l3sdb.in2p3.fr:8080/l3sdb/.

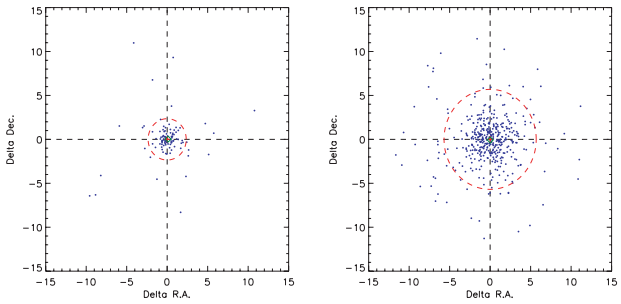


**Figure 5.** Example of positional accuracy diagrams resulting from the log $N$– log $S$ point-source simulations of 10-ks exposures in the soft band. Left-hand panel: input–output positions for the 0< off-axis angle <5 arcmin and CR > 0.01 sources. Right-hand panel: input–output positions for the 5< off-axis angle <10 arcmin and 0.002 < CR < 0.005 sources, where CR is the measured MOS1 + MOS2 + pn count rate, normalized to the on-axis value. Axes are in units of arcseconds. The circle indicates the $3\sigma$ rejection radius used to compute the mean positional error.

**Table 6.** Mean photometric bias ($b$) and $1\sigma$ error ($e$), as a function of count-rate and off-axis distance for point sources, for each of the B and CD bands. Values (in percentile) are derived from 10 ks simulations considering sources having a detection likelihood > 15 (cf. Fig. 3); below count-rates of 0.003, the output locus is degenerate and the bias is too large to estimate meaningful errors. In order to lower the flux at which the bias arises, only sources having gap* > 12 arcsec (cf. Table 4) are used for the error calculation. The true count-rate is related to the observed quantity by: $\mathrm{CR} = \mathrm{CR}_{\mathrm{obs}}\,[100/(b + 100) \pm e/100]$.

| Band<br>Count-rate (count $s^{-1}$) | B<br>$b, e$ | CD<br>$b, e$ |
|---|---|---|
| 0 < off-axis < 5′ | | |
| 0.003 < CR < 0.005 | 9, 21 | 9, 22 |
| 0.005 < CR < 0.0075 | 7, 15 | 8, 16 |
| 0.007 < CR < 0.01 | 6, 12 | 6, 12 |
| 0.01 < CR < 0.02 | 4, 8 | 4, 8 |
| CR > 0.02 | 1, 4 | 1, 4 |
| 5 < off-axis < 10′ | | |
| 0.003 < CR < 0.005 | 10, 26 | 11, 30 |
| 0.005 < CR < 0.0075 | 9, 19 | 9, 21 |
| 0.007 < CR < 0.01 | 7, 15 | 5, 11 |
| 0.01 < CR < 0.02 | 5, 10 | 2, 5 |
| CR > 0.02 | 2, 5 | 1.3 |
| 10 < off-axis < 13′ | | |
| 0.003 < CR < 0.005 | – | – |
| 0.005 < CR < 0.0075 | 10, 24 | 11, 29 |
| 0.007 < CR < 0.01 | 9, 19 | 9, 21 |
| 0.01 < CR < 0.02 | 6, 13 | 7, 14 |
| CR > 0.02 | 3, 7 | 3, 7 |

$i' = 25$ and having a 'good' or 'fair' chance probability $p < 0.03$ as defined in Section 3.3 (in the case where an X-ray source had more than one optical counterpart candidate, we retained only the one with the smallest chance probability). For the three topmost rows of *XMM* pointings in Fig. 1, for which the CFHTLS data are not yet processed, the optical list was made from the USNO-A2 catalogue. The reference files were fed into EPOSCORR using the parameter maxsig = 2 to force removal of spurious matches. The offsets computed by EPOSCORR are plotted in Fig. 6. Offsets computed using USNO-A2 and CFHTLS objects are usually compatible within errors (with the nominal EPOSCORR errors being larger in the USNO-A2 case). The offsets given in Table 1 were applied to all coordinate sets for each source in the data base. The data base column XASTROCORR indicates whether the origin of the astrometric correction is the CFHTLS W1 or the USNO-A2 catalogue. Astrometrically corrected positions are used in the subsequent operations: removal of the redundant sources, source naming and cross-identification with the optical catalogue. Simulations of pairs of point sources separated by 20 arcsec show that 67 per cent of the pairs are resolved for sources having 30 photons each; all pairs being resolved when the sources contain 500 photons. As shown by Pacaud et al. (2006), a detection likelihood threshold of 15 eliminates almost all spurious detections (the corresponding contamination level is ∼1 per cent).

**Table 7.** The energy conversion factors for the individual EPIC cameras and energy bands, stated in units of $10^{-12}$ erg $s^{-1}$ $cm^{-2}$ for a rate of one count $s^{-1}$. A photon index power law of 1.7 and a mean $N_{\mathrm{H}}$ value of 2.6 $10^{20}$ $cm^{-2}$) are supposed. The two MOS cameras are assumed to be identical.

| Detector | B band | CD band |
|---|---|---|
| MOS | 5.0 | 23 |
| pn | 1.5 | 7.9 |

### 2.3.4 *Extended source classification*

Pacaud et al. (2006) and Pierre et al. (2006) presented in detail the criteria for defining galaxy cluster candidates. The selection is performed in the XAMIN output parameter space obtained in the soft band. This band presents the highest S/N at any redshift for typical cluster spectra (as well as for galaxy thermal haloes) thus ensuring the highest completeness level for the extended source detection. The cluster candidate sample consists of the following two classes.

(i) The C1 class is defined such that no point sources are misclassified as extended and is described by extent > 5 arcsec, likelihood of extent > 33 and likelihood of detection[9] >32. Note that while the C1 class is meant to be uncontaminated in terms of point-like

[9] For the ext fit.

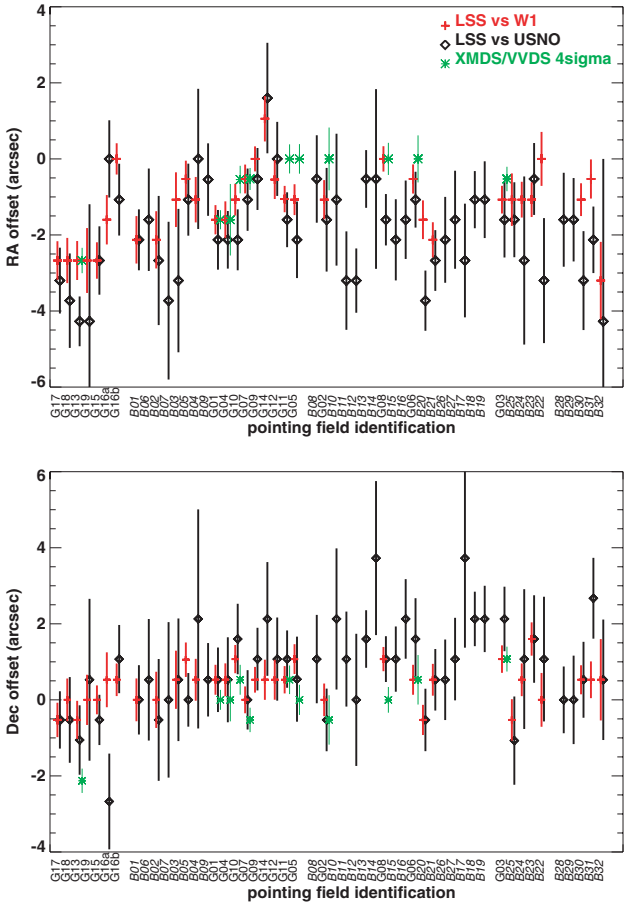


**Figure 6.** Astrometric correction offsets in RA (top panel) and Dec. (bottom panel), with nominal error bars computed by EPOSCORR. The diamonds indicate the offsets computed using the USNO-A2 catalogue, crosses those using CFHTLS W1 data, and asterisks indicate the offsets used in Chiappetti et al. (2005), when applicable. Different offsets for the same field are displaced horizontally for clarity. The *x*-axis gives the *XMM* pointings in chronological order of execution, with different observation 'seasons' separated by a blank space.

sources, it contains a few nearby galaxies whose X-ray emission is unambiguously extended (Pacaud et al., 2007).

(ii) The C2 class is described by extent $> 5$ arcsec and $15 <$ likelihood of extent $< 33$ (no condition on the likelihood of detection) and typically displays a contamination rate of 50 per cent.

There are 73 C1 and C2 objects flagged in the catalogue, column Bc1c2. Note that, for the unique purpose of band merging (see below), a similar classification has been applied in the hard band. In this band, there are only 21 sources flagged as C2 and one flagged as C1, which are not detected in the soft one. A subsequent X-ray/optical inspection of these latter sources does not reveal, as expected, any new cluster. They are mostly weak sources at extreme off-axis angle, or unresolved doubles, or cluster X-ray centroids displaced by more than 6 arcsec from the soft-band position.

### 2.3.5 *Band merging*

Because X-ray colours serve as a useful piece of information for numerous science studies, we provide, in addition to the individual catalogues, the two-band merged catalogue. Band merging is a delicate operation, given the rather large *XMM* PSF (full width at half-maximum $\sim$6 arcsec at 1 keV on axis) and its variation with energy and an off-axis angle. It may also happen that a close pair is

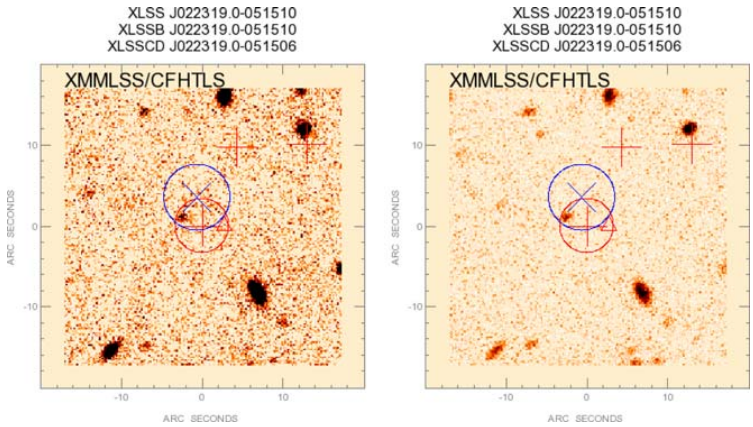


**Figure 7.** An example of postage stamp image centred on source XLSS J022319.0−051510. The CFHT *i*- and *g*-band 40 × 40-arcsec$^2$ images are shown in the left- and right-hand panels, respectively. Soft sources are indicated by a red +, and hard ones by a blue ×; all with astrometrically corrected positions. The original uncorrected position of the central source in the *best band* is indicated with a triangle of the corresponding colour. The $2\sigma$ radius mean error circles are indicated for the central source according to the values of Table 8. All sources within the 40 × 40-arcsec$^2$ field having LH >15 are plotted. The north is up and east is to the left-hand side.

resolved in one band only, because of too low a photon number in the other band. In order to provide the community with an efficient and reliable tool, the construction of the merged catalogue received special attention and is described below.

As an X-ray source can be detected in one or two bands and, for each band, is independently fitted by the extended and point-source models with the coordinates free, we adopt the following merging procedure. For each band, a source is classified as extended (E) if it satisfies extent > 5 arcsec and likelihood of extent > 15 (i.e. C1 or C2 class); if not, it is classified as point-like (P). Pointing per pointing, we then flag associations between the two bands within a search radius of 6 arcsec. Note that we allow associations involving spurious sources (LH < 15) at most in one band. We kept the information (rate, flux, etc.) about entries below this threshold in the merged catalogue, since it could be more useful (e.g. for upper limits) than no information at all, but we flag those cases with Bspurious = 1 or CDspurious = 1. Finally, for each soft–hard couple in the merged catalogue, we define the *best band*, that is, the band in which the detection likelihood of the source is the highest and from which the coordinates are taken. Details of the merging process are summarized in Table 9.

Starting from 2980 non-spurious sources in the soft band and 1255 non-spurious sources in the hard band, the resulting merged catalogue contains 3385 sources: out of them 50.5 per cent are detected as point-like in the soft band only, 36.3 per cent as point-like in both bands, 10.2 per cent as point-like in the hard band only; 2.1 per cent are candidate clusters of galaxies (soft band only). The remaining 0.8 per cent are mostly borderline cases (cf. end of Section 2.3.4). We note that in an extremely limited number of cases (four couples of entries) the merging process gives ambiguous results, that is, a detection in one band can be associated with two different detections in the other band. The flagging and naming of such cases is described in Section 2.3.7.

For *all* entries, we provide in the data base the distance between the positions found by XAMIN in the two energy bands (which in most cases could be used to solve the above ambiguities). Such an interband distance, for all the cases present in both bands, except for the few ambiguities, is within 2 arcsec in 35 per cent of the cases, within 4 arcsec in 79 per cent and above 5 arcsec only in 9 per cent. These percentages change to 39, 83 and 7 per cent if we exclude detections flagged as spurious in one band. If we compare the interband distance with the combination of the position errors computed according to the prescription in Table 8, we have that in 64 per cent of the cases the interband distance is less than the $1\sigma$ error, and in 1.4 per cent of the cases it is above $2\sigma$.

### 2.3.6 *Removal of redundant sources*

Finally, in the case of redundant objects detected in the regions where the pointings overlap, we keep in the catalogues only the detection pertaining to the pointing where the source is the closest to the optical centre (cf. columns Boffaxis, CDoffaxis in Table 4). Except for a few ambiguous cases described in Section 2.3.7, a value of 6 arcsec is found to be the adequate search radius to identify redundant detections. In this way, 280 doublets and a few triplets were identified and reduced to a single source.

### 2.3.7 *Source naming*

Objects in the merged catalogue are labelled following the IAU style convention, that is, XLSS Jhhmmss.s-ddmmss. The coordinates used in assigning the name are the ones deduced after the rigid astrometric correction, and chosen as official, that is, those for the *best band* (see Table 9).

In the individual band catalogues, the sources are assigned supplementary names, following the same standard – XLSSx JHHMMSS.S-DDMMSS – where x is B or CD and stands for the B and CD bands, respectively. In this case, the coordinates correspond to the extended (E) or point-like (P) fit in the relevant band (Table 9). As a consequence of the merging procedure (Section 2.3.5), the

**Table 8.** Positional accuracy ($1\sigma$ error on RA or Dec.) for point sources derived from simulations of 10-ks pointings and having a detection likelihood >15 (Fig. 5). Values are given for the B and CD bands, as a function of the summed measured count-rate: CR = MOS1 + MOS2 + pn. No selection is applied on the gap* parameters (cf. Table 4) but a $3\sigma$ rejection is performed in the calculation of the errors. Because of the strong Eddington bias for faint sources located beyond $R > 10$ arcmin (Fig. 3), no positional errors are provided for output CR below 0.002 count $s^{-1}$.

| Band<br>Count-rate (count $s^{-1}$) | B<br>Error (arcsec) | CD<br>Error (arcsec) |
|---|---|---|
| $0 <$ off-axis $< 5'$ | | |
| $0.001 <$ CR $< 0.002$ | 2.0 | 2.0 |
| $0.002 <$ CR $< 0.005$ | 1.7 | 1.7 |
| $0.005 <$ CR $< 0.01$ | 1.3 | 1.3 |
| CR $> 0.01$ | 0.8 | 0.8 |
| $5 <$ off-axis $< 10'$ | | |
| $0.001 <$ CR $< 0.002$ | 2.0 | 2.0 |
| $0.002 <$ CR $< 0.005$ | 1.8 | 1.9 |
| $0.005 <$ CR $< 0.01$ | 1.5 | 1.5 |
| CR $> 0.01$ | 1.0 | 1.0 |
| $10 <$ off-axis $< 13'$ | | |
| $0.001 <$ CR $< 0.002$ | – | – |
| $0.002 <$ CR $< 0.005$ | 1.9 | 2.0 |
| $0.005 <$ CR $< 0.01$ | 1.6 | 1.7 |
| CR $> 0.01$ | 1.2 | 1.3 |

XLSS coordinate designation will coincide with one of the XLSSB or XLSSCD (i.e. for the *best band*).

In a limited number of cases (eight entries), a source in a band happens to be associated with two different objects in the other band. These couples of catalogue entries are flagged by a non-zero value in column xlink, the registered value being a pointer to the other 'ambiguous' entry in the couple. For those cases, the ambiguity in the XLSS name is resolved (when necessary i.e. in six out of eight cases) by the addition of a suffix: the two members of a couple will appear as XLSS JHHMMSS.S-DDMMSSa and XLSS JHHMMSS.S-DDMMSSb.

## 3 ASSOCIATED DATA PRODUCTS

### 3.1 X-ray images

For each pointing, we make available – via the *XMM*-LSS survey data base in Milan – the following images in the FITS format. They are accessible as 'data products': for every source in the catalogue, a clickable link points to the images of the relevant pointings.

(i) The B and CD band photon images for the three detectors (after the event filtering).

(ii) Exposure maps for the three detectors and the two bands. All images have a pixel size of 2.5 arcsec. Note that the World Coordinate System (WCS) of the X-ray images is the one generated by the SAS; therefore, it does not take into account the astrometric correction described in Section 2.3.3 and quoted in Table 1. Consequently, when overlaying X-ray source positions exactly on the X-ray images, one should use the coordinates labelled as 'raw' in Table 4, although this does not make much difference for most of the sources, given the pixel size.

### 3.2 Optical data

For each X-ray source of the merged catalogue, we provide the following.

(i) The list of optical objects within a radius of 6 arcsec around each X-ray source, extracted from the CFHTLS catalogue.[10] These data are available in the *XMM*-LSS survey data base in Milan only, through the XLSSOPT table. Queries on X-ray lists are allowed and return the $u^*, g', r', i', z'$ magnitudes and further optical information derived from the Terapix merged panchromatic catalogues.

(ii) CFHTLS postage stamp images $40 \times 40$ arcsec$^2$ in the $i$ and $g$ bands (PNG format); in the case where a CFHTLS image is not yet processed, we provide a low-exposure CFH12K image when available. An example is shown in Fig. 7. The parameter list is given in Table 10.

### 3.3 A few statistics

Out of our 3385 X-ray sources, 2208 have at least one optical candidate closer than 6 arcsec, 1071 fall in the area without CFHTLS coverage[11] (corresponding to the topmost three rows of *XMM* pointings in Fig. 1), while 106 have no optical association. The total number of associated optical objects is 6275 (the radius used here is 6 arcsec irrespective of the fact the object is extended or point-like). A total of 472 X-ray sources have a single candidate within 6 arcsec, 613 have two and the rest have more. The number of optical counterparts brighter than $i' = 22.5$ is 1990 of which 55, 36 and 14 per cent are closer than, respectively, 3, 2 and 1 arcsec to an X-ray source (detected in any band). If we consider only the soft (B) band sources, the numbers are 1826, 56 per cent, 37 per cent and 14 per cent, respectively, while for hard (CD) band sources they are 1094, 59 per cent, 40 per cent and 17 per cent.

We further provide the distance $d$ between the X-ray and optical positions as well as an estimate of the probability of chance coincidence

$$p = 1 - \exp[-\pi n(< m)d^2],$$

where $n(< m)$ is the sky density of optical objects having an $i'$ magnitude brighter than the magnitude $m$ of the candidate counterpart, computed from the full CFHTLS W1 catalogue. Defining as 'good' or 'fair, respectively, the X-ray/optical matches having $p < 0.01$ and $0.01 < p < 0.03$ (as described by Chiappetti et al. 2005), we find 945 good associations and 637 fair ones. All the optical objects brighter than $i' = 22.5$ and within 1 arcsec and 87 per cent of those within 2 arcsec have a good probability, while for a distance within 3 arcsec the percentage lowers to 69 per cent for soft sources and 73 per cent for hard sources.

## 4 SUMMARY OF ONLINE AVAILABILITY

Online data access is summarized in Table 11. Namely, we provide the following.

(i) The raw XAMIN results in individual catalogues for the B [0.5–2] keV and CD [2–10] keV bands. Only sources above a detection likelihood of 15 are made available. Redundant sources detected in overlapping regions of different pointings are removed; data from

[10] Currently, from the T0003 W1 field release by Terapix: http://terapix.iap.fr/.
[11] Observed, but currently still under processing.

**Table 9.** Merging decisions and definition of the *best band*. A source is defined as extended (E) in a given band if it satisfies extent > 5 arcsec and likelihood of extent > 15 in this band. Otherwise it is defined as point-like (P). For all sources, but the C1 clusters of galaxies, *flux* is computed from the energy conversion factors given in Table 7 using the point-source rates. For the C1 clusters, the fluxes are set to −1 as the reader is addressed to the *XMM*-LSS survey cluster data base (spatial and spectral fitting providing accurate flux and luminosity measurements). The numbers in [] indicate the number of spurious sources encountered in the interband associations (the counterpart in and B or CD is flagged as spurious, that is, has LH <15).

| B band | CD band | Best band | Coordinates from | Flux in B | Flux in CD | # of sources |
|---|---|---|---|---|---|---|
| E | Undetected | B | B E_fit | −1 for C1, *flux* for C2 | −1 | 59 |
| P | Undetected | B | B P_fit | *flux* | −1 | 1710 |
| Undetected | E | CD | CD E_fit | −1 | *flux* | 22 |
| Undetected | P | CD | CD P_ it | −1 | *flux* | 347 |
| E | E | Where detlik_ext is max | Best band E_fit | −1 for C1, *flux* for C2 | -1 for C1, *flux* for C2 | 2 |
| E | P | Where detlik_ext is max | Best band E_fit | −1 for C1, *flux* for C2 | −1 for C1, *flux* for C2 | 12 [4 CD] |
| P | E | Where detlik_pnt is max | Best band P_fit | *flux* | *flux* | 5 |
| P | P | Where detlik_pnt is max | Best band P_fit | *flux* | *flux* | 1228 [36 B, 358 CD] |

**Table 10.** List of additional optical information presented as columns in the XLSSOPT table. The latter table also includes the X-ray columns marked as such in Tables 4 and 5. Therefore columns with the X, B or CD prefixes refer to X-ray parameters, those with the O prefix to optical data, and those without prefix to combined properties. The arrangement of CFHTLS W1 fields is given at http://terapix.iap.fr/cplt/oldSite/Descart/cfhtls/cfhtlswidemosaictargetW1.html.

| Column name | Units | Meaning and usage | Terapix |
|---|---|---|---|
| Oseq | – | Internal sequence number (unique) | n/a |
| Oid | – | Original Terapix id in field | id |
| Ofield | – | CFHTS field identification in form $\pm x \pm y$ | n/a see the caption |
| Ora | ◦ | Right ascension of the optical candidate | RA |
| Odec | ◦ | Declination of the optical candidate | Dec. |
| Oflag | – | Binary flag combining 0/1 galaxy/star, 0/4 normal/masked, 0/8 normal/saturated | flag |
| Ou | mag | $u^*$ magnitude | u |
| Og | mag | $g'$ magnitude | g |
| Or_ | mag | $r'$ magnitude | r |
| Oi | mag | $i'$ magnitude | i |
| Oz | mag | $z'$ magnitude | z |
| Ou_e | mag | Error on $u^*$ magnitude | uerr |
| Og_e | mag | Error on $g'$ magnitude | gerr |
| Or_e | mag | Error on $r'$ magnitude | rerr |
| Oi_e | mag | Error on $i'$ magnitude | ierr |
| Oz_e | mag | Error on $z'$ magnitude | zerr |
| distance | arcsec | Distance from the X-ray corrected position to the optical position | n/a |
| prob | – | Chance probability of X-ray to optical association (see the text) | n/a |

**Table 11.** The *XMM*-LSS survey X-ray and optical data products. Individual parameter availability is given in Tables 4 and 5.

| Data sets | | Location | Address |
|---|---|---|---|
| Merged XLSS catalogue, main parameters | | CDS | http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/MNRAS/vol/pag |
| Single-band catalogues: | XLSSB, XLSSCD | Milan | http://cosmos.iasf-milano.inaf.it/∼lssadmin/Website/LSS/Query |
| Merged catalogue (all parameters): | XLSS | | |
| X-ray images | | | |
| Optical catalogue : | XLSSOPT | | |
| Optical postage stamps | | | |

the pointing where each object has the smallest off-axis angle are retained in the catalogue. In addition, fluxes assuming a power-law spectrum are provided for each point-like source.

(ii) The B–CD band merged catalogue assuming a correlation radius of 6 arcsec . This required the definition of the *best band* from which a number of parameters, such as the position, are taken and, hence, a selection of information from the single-band catalogues. Source counterparts in the other band are made available even if they have a detection likelihood below 15.

(iii) Optical: a panchromatic CFHTLS catalogue within 6 arcsec of each source of the merged catalogue as well as $i'$- and $g'$-band $40 \times 40$-arcsec$^2$ png images.

The *main* parameters (listed in Tables 4 and 5) of the merged X-ray catalogue are available in an electronic form at the Centre de Données de Strasbourg (CDS). The single-band and band-merged catalogues with *all* columns, as well as the associated data products (X-ray images and optical information) are accessible, with fully

interactive selection, through the *XMM*-LSS survey data base located in Milan and described by Chiappetti et al. (2005). User login details can be found in the entry web page.

## 5 FUTURE

From now on, the *XMM*-LSS survey catalogue and associated data sets will be regularly incremented following the receipt of new *XMM* and CFHTLS pointings. In parallel, XAMIN is being upgraded with the inclusion of information on the photon energy for the source characterization, in such a way as to better discriminate between AGN and cluster sources. When the new version is validated, we foresee reprocessing the entire data set and making it publicly available with the corresponding documentation.

## ACKNOWLEDGMENTS

The results presented here are based on observations obtained with *XMM–Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA. The optical data were obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada–France–Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at Terapix and the Canadian Astronomy Data Centre as part of the CFHT Legacy Survey, a collaborative project of NRC and CNRS. AG acknowledges support from Centre National d'Etudes Spatiales. The Italian members of the team acknowledge financial contribution from the contract ASI-INAF I/023/05/0. AD, OG, EG, P-GS and JS acknowledge support from the ESA PRODEX Programme 'XMM-LSS', and from the Belgian Federal Science Policy Office for their support. HQ acknowledges partial support from the FONDAP Centro de Astrofisica. PG is a Fellow of the Japan Society for the Promotion of Science.

## REFERENCES

Baldi A., Molendi S., Comastri A., Fiore F., Matt G., Vignali C., 2002, ApJ, 564, 190
Chiappetti L. et al., 2005, A&A, 439, 413
Gandhi P. et al., 2006, A&A, 457, 393
Giommi P., Perri M., Fiore F., 2000, A&A, 362, 799
Moretti A., Campana S., Lazzati D., Tagliaferri G., 2003, ApJ, 588, 696
Pacaud F. et al., 2006, MNRAS, 372, 578
Pacaud F. et al., 2007, MNRAS, in press (doi: 10.1111/j.1365-2966.2007.12468.x) (arXiv:0709.1950)
Pierre M. et al., 2004, J. Cosmol. Astropart. Phys., 9, 11
Pierre M. et al., 2006, MNRAS, 372, 591
Pratt G. W., Arnaud M., 2002, A&A, 394, 375
Starck J.-L., Pierre M., 1998, A&AS, 128, 397
Vikhlinin A., Forman W., Jones C., Murray S., 1995, ApJ, 451, 553

This paper has been typeset from a TEX/LATEX file prepared by the author.

## 7.5 L'échantillon complet C1 :

## "The XMM-LSS survey : the First Class cluster sample over the initial 5 deg$^2$ and its cosmological modelling"

### Contexte/motivation

Les amas C1 constituent l'échantillon le plus propre du XMM-LSS dans la mesure où la sélection a lieu uniquement en bande X. De plus, la possibilité de mesurer une température pour chacune des sources rend son exploitation beaucoup plus utile et informative. Le suivi spectroscopique des amas XMM-LSS a donc naturellement privilégié la complétion du catalogue C1 par rapport aux autres détections. Tous les amas C1 de Pierre et al. (2007) ont ainsi été confirmés et l'article suivant présente le plus grand echantillon complet d'amas XMM-LSS étudié à ce jour. Il s'agit également du premier catalogue controllé d'amas X pour lesquels la température a pu être mesurée directement à partir des données du relevé. Une attention particulière est donc portée sur l'impact des effets de sélection sur les propriétés des amas observées.

### Résumé

L'article présente un échantillons de 29 amas de galaxies provenant du relevé XMM-LSS sur une surface d'environ 5 deg$^2$, jusqu'à un redshift de 1.05. Ces amas respectent des critères de sélection X bien définis et sont tous confirmés spectroscopiquement. Nous fournissons, pour tous les amas, des luminosités X et des temperatures ainsi que des estimations de masses, obtenues par la modélisation spatial et spectral de leur émission. La distribution des amas pique vers z 0.3 et T=1.5 keV ; la moitié des systèmes étant des groupes avec des températures inférieures à 2 keV. La relation $L_X - T$ de l'échantillon pointe vers une évolution auto-similaire, mais n'exclue pas d'autres modèles physiquement plausibles. Sa Log(N)-Log(S) est compatible avec les résultats des précédents relevés ROSAT, bien que suggérant une normalisation plus élevée en deçà de $2\,10^{14}$ erg.s$^{-1}$.cm$^{-2}$. En supposant que les relations d'échelle des amas obéissent à une évolution auto-similaire, la densité numérique d'amas se révèle compatible avec les prévisions du modèle de concordance . Parmi les différents échantillons d'amas avec des températures, le XMM-LSS est le seul s'étendant à grand redshift associé à une fonction de sélection bien déterminée. Cela nous a permis de montrer que la prise en compte de la fonction de sélection est indispensable à une analyse correcte de l'évolution de la relation $L_X - T$. Cela pourrait expliquer les résultats contradictoires des études précédentes. Des simulations éxhaustives montrent qu'étendre la surface du relevé à 10 deg$^2$ permettra d'exclure l'hypothèse de non-évolution. Les contraintes sur des modèles plus raffinés de milieu intra-amas seront probablement limitées par la forte dispersion intrinseque de la relation $L_X - T$. Celle-ci néccessite donc de larges échantillons. Nous montrons de plus que le nombre d'amas détectable dépend de la dispersion suposée dans les relations d'échelle, révélant ainsi une dégénérescence supplémentaire (avec $\sigma_8$, $\Omega_m$, $L_X - T(z)$) dans l'interprétation cosmologique des

comptages d'amas. Nous présentons finalement des formules empiriques utiles pour les relations [masse - flux] et [masse - taux de comptage X] ainsi qu'une comparaison entre la sensibilité du XMM-LSS et celle des relevés S-Z à venir.

### Contribution personnelle

La majeur partie du contenu se fonde sur mon travail. Seul la mesure et la validation des redshifts, ainsi que la comparaison avec les relevés SZ proviennent d'autres co-auteurs.

Mon. Not. R. Astron. Soc. (2007) doi:10.1111/j.1365-2966.2007.12468.x

# The *XMM*-LSS survey: the Class 1 cluster sample over the initial 5 deg$^2$ and its cosmological modelling$^\star$

F. Pacaud,[1,2,3]† M. Pierre,[1,2] C. Adami,[4] B. Altieri,[5] S. Andreon,[6] L. Chiappetti,[7] A. Detal,[8] P.-A. Duc,[2] G. Galaz,[9] A. Gueguen,[2] J.-P. Le Fèvre,[10] G. Hertling,[9] C. Libbrecht,[8] J.-B. Melin,[11] T. J. Ponman,[12] H. Quintana,[9] A. Refregier,[1,2] P.-G. Sprimont,[8] J. Surdej,[8] I. Valtchanov,[5] J. P. Willis,[13] D. Alloin,[2] M. Birkinshaw,[14] M. N. Bremer,[14] O. Garcet,[8] C. Jean,[8] L. R. Jones,[12] O. Le Fèvre,[4] D. Maccagni,[7] A. Mazure,[4] D. Proust,[15] H. J. A. Röttgering[16] and G. Trinchieri[6]

[1]*DSM/DAPNIA/SAp, CEA Saclay, F-91191 Gif-sur-Yvette, France*
[2]*AIM – Unité Mixte de Recherche CEA – CNRS – Université Paris VII – UMR 715*
[3]*Argelander-Institut für Astronomie, University of Bonn, Auf dem Hügel 71, 53121 Bonn, Germany*
[4]*Laboratoire d'Astrophysique de Marseille, BP8, F-13376 Marseille Cedex 12, France*
[5]*ESA, Villafranca del Castillo, Spain*
[6]*INAF-Osservatorio Astronomico di Brera, via Brera 28, I-20121 Milano, Italy*
[7]*INAF-IASF Milano, via Bassini 15, I-20133 Milano, Italy*
[8]*Institut d'Astrophysique et de Géophysique, Université de Liège, Allée du 6 Août, 17, B5C, 4000 Sart Tilman, Belgium*
[9]*Departamento de Astronomía y Astrofísica, Pontificia Universidad Católica de Chile, Casilla 306, Santiago 22, Chile*
[10]*DSM/DAPNIA/SEDI, CEA Saclay, F-91191 Gif-sur-Yvette, France*
[11]*DSM/DAPNIA/SPP, CEA Saclay, F-91191 Gif-sur-Yvette, France*
[12]*School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT*
[13]*Department of Physics and Astronomy, University of Victoria, Elliot Building, 3800 Finnerty Road, Victoria, V8V 1A1, BC, Canada*
[14]*Department of Physics, University of Bristol, Tyndall Avenue, Bristol BS8 1TL*
[15]*GEPI, Observatoire de Paris-Meudon, F-92195 Meudon Cedex, France*
[16]*Leiden Observatory, PO Box 9513, 2300 RA Leiden, the Netherlands*

Accepted 2007 September 12. Received 2007 September 12; in original form 2007 March 7

**ABSTRACT**

We present a sample of 29 galaxy clusters from the *XMM*-LSS survey over an area of some 5 deg$^2$ out to a redshift of $z = 1.05$. The sample clusters, which represent about half of the X-ray clusters identified in the region, follow well-defined X-ray selection criteria and are all spectroscopically confirmed. For all clusters, we provide X-ray luminosities and temperatures as well as masses, obtained from dedicated spatial and spectral fitting. The cluster distribution peaks around $z = 0.3$ and $T = 1.5$ keV, half of the objects being groups with a temperature below 2 keV. Our $L_X$–$T(z)$ relation points towards self-similar evolution, but does not exclude other physically plausible models. Assuming that cluster scaling laws follow self-similar evolution, our number density estimates up to $z = 1$ are compatible with the predictions of the concordance cosmology and with the findings of previous *ROSAT* surveys. Our well-monitored selection function allowed us to demonstrate that the inclusion of selection effects is essential for the correct determination of the evolution of the $L_X$–$T$ relation, which may explain the contradictory results from previous studies. Extensive simulations show that extending the survey area to 10 deg$^2$ has the potential to exclude the non-evolution hypothesis, but those constraints on more refined intracluster medium models will probably be limited by the large intrinsic dispersion of the $L_X$–$T$ relation, whatever be the sample size. We further demonstrate that increasing the dispersion in the scaling laws increases the number of detectable clusters, hence generating

$^\star$Based on data collected with *XMM*, Very Large Telescope, Magellan, NTT and Canada–France–Hawaii Telescope; ESO programme numbers are: 070.A-0283, 070.A-907 (VVDS), 072.A-0104, 072.A-0312, 074.A-0360 and 074.A-0476.

†E-mail: pacaud@astro.uni-bonn.de

further degeneracy [in addition to $\sigma_8$, $\Omega_m$, $L_X$–$T(z)$] in the cosmological interpretation of the cluster number counts. We provide useful empirical formulae for the cluster mass–flux and mass–count rate relations as well as a comparison between the *XMM*-LSS mass sensitivity and that of forthcoming Sunyaev–Zel'dovich surveys.

**Key words:** surveys – cosmological parameters – large-scale structure of Universe – X-rays: galaxies: clusters.

## 1 INTRODUCTION

Along with cosmic microwave background (CMB) measurements and supernova observations, clusters of galaxies provide key cosmological information. It is especially instructive to cross-check the constraints from these three classes of data, since they originate from different physical processes. Moreover, since theory and numerical simulations allow us to follow cluster formation from the initial power spectrum, which is directly measured from the CMB, it is critical to test that the 'CMB *WMAP* concordance cosmology' is consistent with the observed properties of clusters in the low-$z$ Universe.

In the framework of hierarchical cosmic structure formation involving cold dark matter (CDM) type scenarios, where the smallest perturbations collapse first, clusters correspond to the mass scale that entered the non-linear regime between redshift three and the present epoch. In this sense, the most-massive galaxy clusters in the local Universe represent the largest virialized structures. This property of being both 'relaxed' and rare $3\sigma$ events' regime has been extensively exploited through formalisms like that of Press & Schechter (1974) in connection with the spherical collapse model, for the general case of Gaussian random field fluctuations. This connects, in an analytically tractable manner and for any redshift, both the cluster abundance as a function of mass and the cluster spatial distribution, to the properties of the initial fluctuation spectrum – in particular its normalization, $\sigma_8$, its shape, $\Gamma$, as a function of the density of the Universe, $\Omega_m$, and the equation of state of dark energy (e.g. Refregier, Valtchanov & Pierre 2002; Majumdar & Mohr 2003). This first-order approach is well supported by numerical CDM simulations: clusters lie at nodes of the cosmic network, have virialized cores, and are still growing by accretion along filaments at a rate that depends on the cosmology. However, at the same time as clusters started being used as cosmological tools, it was realized that the interpretation of their observed abundance as a function of time is actually very much dependent on the evolution of the observable cluster properties themselves. In order to break this latent degeneracy in such a way that clusters can effectively be used as cosmological candles, it is essential to understand how cluster properties impact on their delectability at any epoch.

While galaxies constitute only a few per cent of the total cluster mass, about 80 per cent of the baryonic mass resides in the tenuous X-ray-emitting intracluster medium (ICM), settled in the cluster gravitational potential. Because cluster X-ray emission is extended, clusters are readily identified among the high Galactic latitude X-ray population, which is dominated by point-like active galactic nuclei (AGN). However, whatever the detection method – optical or X-ray – the fundamental question of *how to relate observable quantities to cluster masses* remains. This is crucial because it is the cluster masses that enter the theory of structure formation as generated by theoretical cosmological calculations. This issue becomes particularly important outside the local Universe, since the evolution of cluster properties is not well established.

The Einstein Medium Sensitivity Survey (EMSS), followed by REFLEX based on the *ROSAT* All-Sky Survey, and a number of serendipitous clusters surveys (RDCS, SHARC, MACS, 160 deg$^2$, etc.) from *ROSAT* deep pointings, have provided the first 'cosmological cluster samples' (see the synoptic plot in Pierre et al. 2004 and references therein). The mass-observable relations used in the analysis of data from these surveys relied on the assumption of hydrostatic equilibrium and on the (mostly local) observed $L_X$–$T$ relation. The cluster selection function – a key ingredient – was modelled using a variable flux limit across the survey area. Under these hypotheses, the derived cosmological constraints appeared to be in agreement with the concordance model [see a review in Rosati et al. 2002].

Within the same period, deep *GINGA*, *ROSAT* and *ASCA* observations of nearby clusters revealed that the $L_X$–$T$ relation is significantly steeper than expected from purely gravitational heating (Arnaud & Evrard 1999), hence suggesting the presence of other heating/cooling sources such as feedback from star formation or AGN, in addition to the effects of cooling flows. This particularly affects the low-mass end of the cluster population – groups with temperatures [0.5–2] keV where the gravitational binding energies are low.

Nearby cluster observations at high spatial resolution by *Chandra* have also shown that the ICM is not the well-relaxed medium previously assumed: shocks, cold fronts, and bulk velocities are seen even in apparently relaxed clusters (e.g. Mazzotta, Fusco-Femiano & Vikhlinin 2002; Dupke & Bregman 2006). High spectral resolution *XMM* pointings have led to a totally new version of the putative 'cooling flow' scenario, where episodic heating/accretion by the central AGN could play a key role in the ICM, preventing any central 'cooling catastrophe'. Measurements of the $L_X$–$T$ relation for distant ($z > 0.5$) massive clusters are in progress; whether the evolution of scaling laws follows simple self-similar expectations is still hotly debated (see the review in Arnaud 2005a). All these results currently pertain to the upper end of the cluster mass function. They present a new challenge for high-resolution numerical simulations which, in turn, should quantify deviations from hydrostatic equilibrium.

In parallel, the building of large serendipitous *XMM* and *Chandra* cluster samples – from public archive data – has been initiated by a number of groups: the *XMM* Cluster Survey (XCS, Romer et al. 2001), SEXCLAS (Kolokotronis et al. 2006) and ChaMP (Barkhouse et al. 2006). Preliminary results over areas $<10$ deg$^2$ give cluster densities of the order of 5 deg$^{-2}$ for objects detected independently in both X-ray and optical wavebands. No cosmological analysis has been performed on these samples so far, but these searches have enabled the detection of the most-distant X-ray clusters to date at $z = 1.4$ (Mullis et al. 2005) and $z = 1.45$ (Stanford et al. 2006).

In addition to these large surveys, a number of contiguous surveys, such as COSMOS (Finoguenov et al. 2007) and XBootes (Kenter et al. 2005), are also being conducted. Following on from the REFLEX cluster survey (Böhringer et al. 2002), and exploiting the unrivalled sensitivity of the *XMM–Newton* X-ray observatory, the *XMM* wide-area survey (*XMM*-LSS, Pierre et al. 2004) is the largest contiguous X-ray cluster survey being undertaken at the present time. It has been designed to investigate the large-scale structure of the Universe as traced by galaxy clusters to redshifts $z = 1$ and beyond. The *XMM*-LSS sensitivity limit is $\sim$1000 times deeper than REFLEX, that is, $\sim 4 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ in the [0.5–2] keV band for point sources. Moreover, the *XMM*-LSS is able to make a systematic exploration for massive clusters out to at least $z \sim 1.5$ (Refregier et al. 2002; Bremer et al. 2006; Pacaud et al. 2006). Another key improvement is the spatial resolution of *XMM*: $\sim$6 arcsec full width at half-maximum (FWHM) on axis compared to $\sim$20 arcsec for *ROSAT*.

The two major requirements of the X-ray processing were to reach the sensitivity limit of the data in a statistically tractable manner in terms of cluster detection efficiency, and so to provide the selection function of the detected objects. To achieve these goals, it was necessary to design a new two-step X-ray pipeline, combining wavelet multiresolution analysis with maximum-likelihood fitting; Poisson statistics being used in both steps (Pacaud et al. 2006). We stress here that at our sensitivity and spatial resolution, the survey is primarily limited by surface brightness, rather than by flux as assumed by past generations of surveys. This led us to define cluster selection criteria in a two-dimensional parameter space, corresponding to specific levels of contamination and completeness, as discussed below.

Building on the preliminary results from the first square degree of the survey (Pierre et al. 2006), we have defined a well-controlled cluster sample covering all currently available *XMM*-LSS observations, that is, about 5 deg$^2$. The sample comprises some 30 clusters with fluxes in the [0.5–2] keV band ranging from 1 to 50 $\times$ $10^{-14}$ erg s$^{-1}$ cm$^{-2}$. They are referred to as 'Class 1', hereafter C1, clusters, because the criteria used to construct the sample from a two-dimensional X-ray parameter space guarantee no contamination by point-like sources. The observations were performed in a rather homogeneous way (10–20 ks exposures) and enable us, in addition, to estimate temperatures for all the sample clusters. This is a key datum which, along with careful modelling of the survey selection effects, allow us to tackle for the first time on real data the calibration of the mass-observable relations in a cluster survey – something which has been mostly addressed in a formal way thus far (e.g. Majumdar & Mohr 2003 and references therein). This led us to quantitatively investigate some of the degeneracies affecting the cosmological interpretation of the redshift distribution, d$n$/d$z$, of cluster number counts.

This paper is organized as follows. In Section 2, we summarize the principles of our X-ray analysis, provide a detailed calculation of the selection function, and subsequently present the cluster sample. The next section describes the spectroscopic confirmation of the clusters, and the determination of their X-ray temperature and luminosity. Section 4 provides a thorough discussion of the $L_X$–$T$ relation: special care is given to the modelling of the selection effects in the analysis of possible evolutionary trends. This is an especially informative exercise as it is the first time that the behaviour of the relation has been explored for a well-controlled sample. Section 5 is devoted to the modelling of the sample; starting from the linear spectrum of the initial density fluctuations, we compute the dark matter halo mass spectrum and infer the halo X-ray luminosity and temperature using empirical scaling laws compatible with our data. The cluster population thus obtained is then folded through the *XMM*-LSS selection function. The derived d$n$/d$z$ is compared with that observed, and the model provides us with mass-observable relations. The model also allows us to further explore (Section 6) the impact of several cluster parameters on classical cosmological tests; for instance, we examine the sensitivity of our observed d$n$/d$z$ to the shape of the $M$–$T$ relation in the group regime, to the amount of scatter in the scaling laws, and to changes in cosmological evolution resulting from different equations of state for dark energy.

Throughout this paper, data are analysed using the cosmological parameters estimated by Spergel et al. (2006), namely $H_0 = 73\,\mathrm{km\,s^{-1}\,Mpc^{-1}}, \Omega_\mathrm{m} = 0.24, \Omega_\Lambda = 0.76, \Omega_\mathrm{b} = 0.041, n_\mathrm{s} = 0.95$ and $\sigma_8 = 0.74$.

## 2 CONSTRUCTION OF THE SAMPLE

### 2.1 The data set

To date, the X-ray data received consist of a single region of roughly 6 deg$^2$ covered by 51 *XMM–Newton* pointings.[1] Most of the observations (32 pointings, hereafter B pointings) were obtained through *XMM* AO1/AO2, and have a nominal exposure time of $10^4$ s, whilst 19 (hereafter G pointings) are deeper $2 \times 10^4$ s guaranteed time observations. The latter constitute the *XMM* Medium Deep Survey (XMDS, Chiappetti et al. 2005).

Calibrated event lists were created using the XMM-SCIENCE ANALYSIS SYSTEM (XMM-SAS)[2] tasks EMCHAIN and EPCHAIN, and periods of intense proton flares were filtered out using the method suggested by Pratt & Arnaud (2002). Out of the 51 observations, six (including one G pointing) were too strongly contaminated to properly monitor and remove the high-flare periods. These pointings are scheduled for re-observation during *XMM* AO5 and, consequently, are not included in the analysis of this paper, which covers only 5 deg$^2$.

Details of the observations can be found in the publication of the full X-ray catalogue pertaining to this first 5 deg$^2$ (Pierre et al. 2007).

### 2.2 Data pre-analysis

For each pointing, images were generated in the [0.5–2] keV band from the filtered event lists using the XMM-SAS task EVSELECT.

These were subsequently analysed using the *XMM*-LSS source detection pipeline (described in detail by Pacaud et al. 2006): first, the images are filtered using a wavelet multiresolution algorithm (Starck & Pierre 1998) that was specifically designed to properly account for the Poisson noise in order to smooth the background and lower the noise level, while keeping unchanged the relevant information. SExtractor is then used on the filtered images to extract a very deep primary source catalogue. Finally, detailed properties of each detected source are assessed by XAMIN, a maximum-likelihood profile fitting algorithm that we have developed to characterize extended sources in *XMM–Newton* images.

[1] These are currently being complemented during *XMM–Newton* AO5 so as to cover the full 10 deg$^2$ of the *Spitzer*/SWIRE area surveyed in the *XMM*-LSS field.

[2] XMM-SAS, http://xmm.vilspa.esa.es/sas/

### 2.3 Source-selection process

#### *2.3.1 Definition of the sample*

As shown by Pacaud et al. (2006), the probability of detecting an extended source at the *XMM*-LSS sensitivity and resolution depends on both its flux and its angular extent. This led us to abandon the simple concept of a flux-limited cluster survey and to define a system of extended source classes: the final source selection is performed in the XAMIN output parameter space, and several samples are defined, allowing for various amounts of contamination from point sources and spurious detections. The subsamples, or classes, have been defined by means of extensive '*in situ*' simulations, involving a large range of cluster fluxes and apparent sizes, for the nominal exposure of the *XMM*-LSS, that is, $10^4$ s. In this paper, we focus on the C1 sample which was defined from our simulations to be the largest uncontaminated extended source sample available. It is obtained by selecting candidates with extension $>5$ arcsec, extension likelihood $>33$, and detection likelihood $>32$. Given that no spurious extended detection was found in our whole simulated data set, we are confident that the false detection rate is close to zero, and a visual inspection of the X-ray/optical overlays for the C1 sources showed the existence of a cluster of galaxies or of a nearby galaxy as an optical counterpart in every case. Furthermore, we find that the C1 sample coincides with clusters for which a reliable temperature can be obtained using X-ray spectroscopy from the survey data (Section 3.2).

Detailed information on the C1 selection process can be found in Pacaud et al. (2006).

#### *2.3.2 Selection function*

Given the varying effective exposure time per pointing, the survey coverage is not uniform ($7 < t_{\rm exp}/10^3$ s $< 20$, Fig. 1). Consequently, similar likelihood values do not exactly pertain to the same objects

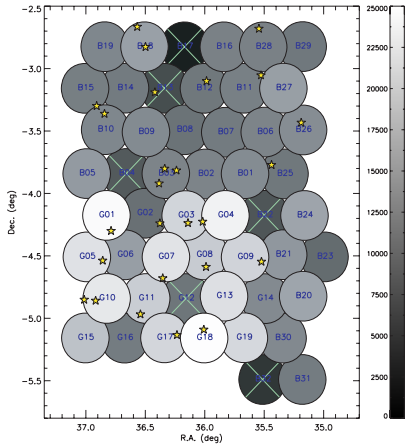


**Figure 1.** Sky distribution of the C1 clusters. The colour scale indicates the on-axis exposure time of each pointing in seconds (after particle flare filtering). Pointings marked by a cross are strongly affected by flares and will be re-observed during AO5.

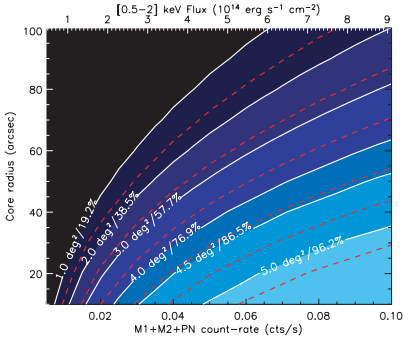


**Figure 2.** The survey coverage is displayed in a two-dimensional parameter space: the sky coverage is a function of both cluster flux and extent. The dashed lines are the result of extensive 10-ks simulations. The slightly shifted white lines are the analytical corrections accounting for exposure variations across the surveyed area, hence indicating the effective selection function of the current data set. The sample completeness corresponding to each line is indicated in terms of both the sky coverage and the percentage of detected sources. Extent values correspond to the core radius of a $\beta$-model with $\beta = 2/3$. The count rate to flux conversion assumes a 2-keV spectrum at $z = 0$.

across the survey. Rather than lowering all exposures to $7 \times 10^3$ s, we apply the C1 criteria to every pointing in order to maximize the size of the cluster sample. The probability of detecting a cluster of a given flux and extent for any pointing is subsequently derived by applying an analytic correction to the $10^4$-s simulations, scaling as a function of exposure time the signal-to-noise ratio (S/N) produced by such an object.[3] This is justified by the fact that the range spanned by the exposure times is modest ($1/2 < t_{\rm exp}/10^4$ s $< 2$), implying also that source confusion does not change significantly. The survey selection function, that is, the probability to detect a source of a given flux and apparent size as a C1 cluster, is obtained by integrating the contributions from all pointings. It is this function that is used in the following for cosmological applications. It is displayed in the form of a sky coverage in Fig. 2. The maximum area covered reaches 5.2 deg$^2$. This happens, for example, for sources with 20 arcsec core radius, and fluxes above $5 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$. Our sensitivity drops by roughly a factor of 2 for sources with flux around $1.5 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ and 20 arcsec core radius, or $5 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ flux and 75 arcsec core radius.

### 2.4 The sample

Applying the C1 selection criteria to the inner 11.5 arcmin of the 45 valid observations yields 37 detections, among which four are duplicated sources detected on adjacent pointings. Optical/X-ray overlays reveal that five out of the remaining 33 sources are actually nearby X-ray-emitting galaxies whose properties are presented in Appendix C. Hence, we have a total of 28 galaxy clusters. An additional unambiguous C1 source is detected in pointing B13 (highly contaminated by flares); we provide the basic properties of this source in this paper, but it is not used in the scientific analysis as we

[3] The mean background values measured by Read & Ponman (2003) are assumed.

cannot accurately compute the selection function for this pointing. The cluster sample is listed on Table 1. We note that a complementary cluster sample of about the same size has also been identified at lower significance in the same area; these clusters belong to our C2 and C3 classes, for which the selection criteria are much less well defined. They will be published in a separate paper (Adami et al., in preparation).

## 3 ASSESSING INDIVIDUAL SOURCE PROPERTIES

### 3.1 Spectroscopic validation

Each C1 X-ray source was confirmed as a bona fide galaxy cluster at specific redshift via optical photometry and spectroscopy.

The overall procedure used to confirm individual clusters is very similar to the approach taken in previously published *XMM*-LSS papers (Valtchanov et al. 2004; Willis et al. 2005a; Pierre et al. 2006) and is summarized below.

A combination of either CTIO/MOSAIC II $Rz\prime$ (Andreon et al. 2004) or Canada–France–Hawaii Telescope (CFHT)/MEGACAM[4] *ugriz* imaging was used to associate the location of each X-ray source with the spatial barycentre of a significant overdensity of galaxies displaying characteristically red colours. Based on the CTIO data, galaxies lying within a given colour tolerance of this 'red sequence' were then flagged as candidate cluster members and given a high priority in subsequent multi-object spectroscopic observations. A small number of low X-ray temperature ($\lesssim$1 keV) groups at moderate redshift ($z > 0.2$) and local ($z < 0.2$) compact groups did not display a statistically significant red galaxy overdensity. These systems were inspected visually and spectroscopic targets were assigned manually.

Due to the moderately large size and the extended redshift range covered by the sample, cluster targets were observed using a number of facilities over several observing semesters. Details of the observing configurations can be found in Table 2.

Spectroscopic data were reduced using standard procedures described in detail in previously published *XMM*-LSS publications (Valtchanov et al. 2004; Willis et al. 2005a; Pierre et al. 2006). Where possible, redshift values and associated uncertainties of individual galaxy targets were computed via cross-correlation with galaxy reference templates. In the remaining cases, redshifts were assigned manually, from emission features.

The nominal condition adopted by *XMM*-LSS to confirm a cluster redshift is to observe three concordant redshifts (typically $\Delta z \lesssim 0.01$) within a projected scale of about 500 kpc of the cluster X-ray centre. Only cluster XLSSC-052 does not fulfill this condition: the X-ray emission is associated with a pair of $z = 0.06$ galaxies very close to a bright star. Except for this specific system, the number of galaxy members confirmed per cluster ranges between 3 and 15, with a typical cross-correlation velocity error of the order of 50–150 km s$^{-1}$ per galaxy. Despite this velocity accuracy, we are limited for a large fraction of our sources by small-number statistics and consequently quote redshifts only to $\pm 0.01$. In each case, the final cluster redshift is computed from the unweighted mean of all galaxies lying within $\pm 3000$ km s$^{-1}$ of the visually assigned redshift peak.

[4] Data are taken from the CFHT Wide Synoptic Legacy Survey. See the URL www.cfht.hawaii.edu/Science/CFHTLS/ for further details.

### 3.2 Spectral analysis

In order to measure the temperature of the intracluster gas, X-ray spectra were extracted in a circular aperture around each source. The corresponding background emission was estimated within a surrounding annulus having inner radius large enough for the cluster contribution to be considered negligible. Preliminary modelling of the cluster surface brightness profile allowed the determination of the optimal extraction radii in terms of the S/N.

The resulting spectra were fitted using XSPEC[5] to a thermal plasma model (APEC) assuming a fixed hydrogen column density set to the Galactic value as derived from H I observations by Dickey & Lockman (1990). The metal abundance of the gas was held fixed during the fitting process at 0.3 times the solar abundance, as estimated by Grevesse & Sauval (1998). As explained in Willis et al. (2005a), the cluster spectra were constructed imposing a minimum requirement of five background photons per bin in order to avoid the apparent bias we identified in XSPEC temperature estimates when using the Cash statistic on very sparse spectra. Our simulations (Willis et al. 2005a) showed that this procedure provides quite reliable temperature measurements ($\pm$10–20 per cent) for $\sim$1–3 keV clusters having only a few hundred counts. We further investigated the impact of fixing the metal abundance at $0.3\,Z_{\odot}$, by computing best-fitting temperatures obtained using extreme mean abundances of 0.1 and $0.6\,Z_{\odot}$. In most cases, the temperatures fell within the $1\sigma$ error bars from our initial fit. For five systems (namely XLSSC-008, XLSSC-028, XLSSC-041, XLSSC-044 and XLSSC-051), one of these two extreme measurements just fell a few per cent outside our error range. The measured temperatures are presented in Table 1.

### 3.3 Spatial analysis

#### *3.3.1 Surface brightness modelling*

To accurately determine the cluster fluxes and luminosities, we modelled the observed photon spatial distribution in the [0.5–2] keV band with a radial $\beta$-profile:

$$S(r) = \frac{S_0}{[1 + (r/R_c)^2]^{3\beta - 1/2}}, \quad (1)$$

using a refined version of the method described in Pierre et al. (2006).

For each source, we started by fitting the mean background levels (vignetted and particle components) over the inner 13 arcmin of the pointing, excluding all sources detected by the pipeline. For each EPIC camera, the photons were then binned within 3 arcsec annuli centred on the cluster peak. The resulting profiles were subsequently rebinned imposing a minimum S/N of 3 with respect to the estimated background level, weighted by the annular exposure times, and finally co-added to build a MOS1 + MOS2 + PN count-rate profile. The fitted model is constructed by convolving the circular $\beta$-profile with an analytical parametrization of the point spread function (PSF) (Ghizzardi et al. 2001), as implemented by Arnaud et al. (2002). The $\chi^2$ statistic is computed over a discrete grid of $\beta$ and $R_c$ values, with the value of the normalization coefficient, $S_0$, optimized analytically.

As already discussed in Pierre et al. (2006), the majority of the clusters are faint, and the detected photon distribution in many cases represents only a fraction of the extended X-ray surface brightness distribution. Under such conditions, the parameters $\beta$ and $R_c$ are degenerate when fitted simultaneously, limiting the extent to which

[5] http://heasarc.gsfc.nasa.gov/docs/xanadu/xspec/

**Table 1.** List of the C1 galaxy clusters and their basic optical/X-ray properties. More information on peculiar individual clusters is provided in Appendix B.

| Source name | XLSSC number | Pointing[a] | RA (J2000) | Dec. (J2000) | Redshift | $N_{gal}$[b] | $R_{spec}$ (arcsec) | Counts ($R_{spec}$) | $T$ (keV) | $R_{fit}$ (arcsec Mpc$^{-1}$) | Counts ($R_{fit}$) | $F_X$[c] (0.5 Mpc) | $r_{500}$ (Mpc) | $L_{500}$[d] ($10^{43}$ erg s$^{-1}$) | $M_{500}$[e] ($10^{13}\,h^{-1}\,M_\odot$) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| XLSS J022023.5−025027 | 039 | B29 | 35.0983 | −2.8409 | 0.23 | 3 | 66.5 | 144 | $1.3^{+0.3}_{-0.1}$ | 54/0.198 | 112 | $2.5^{+0.9}_{-0.9}$ | 0.417 | $0.9^{+0.2}_{-0.3}$ | 2.4 |
| XLSS J022045.4−032558 | 023 | B26 | 35.1894 | −3.4328 | 0.33 | 7 | 55.0 | 338 | $1.7^{+0.3}_{-0.2}$ | 84/0.399 | 384 | $4.4^{+0.4}_{-0.4}$ | 0.457 | $3.8^{+0.3}_{-0.3}$ | 3.1 |
| XLSS J022145.2−034617 | 006 | B25 | 35.4385 | −3.7715 | 0.43[f] | 26 | 80.0 | 1304 | $4.8^{+0.6}_{-0.5}$ | 606/3.399 | 2099 | $25.0^{+0.8}_{-0.7}$ | 0.838 | $60.3^{+1.8}_{-1.8}$ | 19 |
| XLSS J022205.5−043247 | 040 | G09 | 35.5232 | −4.5464 | 0.32 | 3 | 41.5 | 116 | $1.6^{+1.1}_{-0.3}$ | 117/0.544 | 273 | $2.1^{+0.3}_{-0.2}$ | 0.442 | $1.6^{+0.2}_{-0.2}$ | 3.4 |
| XLSS J022206.7−030314 | 036 | B11 | 35.5280 | −3.0539 | 0.49 | 3 | 46.5 | 507 | $3.6^{+0.6}_{-0.4}$ | 213/1.286 | 659 | $10.7^{+0.6}_{-0.5}$ | 0.676 | $28.9^{+1.5}_{-1.5}$ | 11 |
| XLSS J022210.7−024048 | 047 | B28 | 35.5447 | −2.6801 | 0.79 | 12 | 60.0 | 114 | $3.9^{+2.8}_{-1.4}$ | 75/0.561 | 89 | $1.6^{+0.3}_{-0.2}$ | 0.592 | $13.2^{+2.2}_{-2.3}$ | 12 |
| XLSS J022253.6−032828 | 048 | B07 | 35.7234 | −3.4745 | 1.00 | 3 | 55.0 | 137 | $1.8^{+0.7}_{-0.4}$ | 66/0.529 | 94 | $1.5^{+0.2}_{-0.2}$ | 0.327 | $17.4^{+3.1}_{-3.1}$ | 2.9 |
| XLSS J022348.1−025131 | 035 | B16 | 35.9507 | −2.8588 | 0.17 | 5 | 60.0 | 145 | $1.2^{+0.1}_{-0.1}$ | 162/0.469 | 334 | $4.8^{+0.5}_{-0.5}$ | 0.394 | $0.77^{+0.07}_{-0.07}$ | 1.5 |
| XLSS J022356.5−030558 | 028 | B12 | 35.9857 | −3.0997 | 0.30[g] | 8 | 60.0 | 144 | $1.3^{+0.2}_{-0.2}$ | 60/0.267 | 118 | $2.7^{+0.4}_{-0.4}$ | 0.399 | $1.5^{+0.2}_{-0.2}$ | 1.5 |
| XLSS J022357.4−043517 | 049 | G08 | 35.9892 | −4.5883 | 0.49 | 4 | 25.0 | 76 | $2.2^{+0.9}_{-0.5}$ | 456/2.753 | 722 | $1.9^{+0.2}_{-0.2}$ | 0.493 | $4.3^{+0.4}_{-0.4}$ | 3.4 |
| XLSS J022402.0−050525 | 018 | G18 | 36.0087 | −5.0904 | 0.32[f] | 14 | 35.0 | 203 | $2.0^{+0.7}_{-0.4}$ | 75/0.349 | 245 | $1.5^{+0.2}_{-0.2}$ | 0.521 | $1.3^{+0.2}_{-0.2}$ | 4.7 |
| XLSS J022404.1−041330 | 029 | G04 | 36.0172 | −4.2251 | 1.05[h] | 5 | 33.5 | 310 | $4.1^{+0.9}_{-0.7}$ | 60/0.486 | 355 | $3.1^{+0.2}_{-0.2}$ | 0.524 | $48.3^{+3.7}_{-3.4}$ | 8.9 |
| XLSS J022433.8−041405 | 044 | G03 | 36.1411 | −4.2347 | 0.26[h] | 10 | 54.5 | 319 | $1.3^{+0.2}_{-0.1}$ | 294/1.183 | 544 | $2.8^{+0.2}_{-0.2}$ | 0.399 | $1.2^{+0.1}_{-0.1}$ | 2.0 |
| XLSS J022456.2−050802 | 021 | G17 | 36.2345 | −5.1339 | 0.08 | 7 | 28.5 | 265 | $0.68^{+0.04}_{-0.02}$ | 75/0.113 | 372 | $3.3^{+0.6}_{-0.5}$ | 0.297 | $0.11^{+0.01}_{-0.01}$ | 0.9 |
| XLSS J022457.1−034856 | 001 | B03 | 36.2381 | −3.8157 | 0.61[i] | 23 | 60.0 | 730 | $3.2^{+0.4}_{-0.3}$ | 129/0.869 | 770 | $7.6^{+0.3}_{-0.4}$ | 0.584 | $33.2^{+1.5}_{-1.5}$ | 8.3 |
| XLSS J022520.8−034805 | 008 | B03 | 36.3370 | −3.8015 | 0.30[f] | 11 | 45.0 | 99 | $1.3^{+0.7}_{-0.2}$ | 57/0.254 | 104 | $2.1^{+0.3}_{-0.3}$ | 0.396 | $1.2^{+0.2}_{-0.2}$ | 1.4 |
| XLSS J022524.7−044039 | 025 | G07 | 36.3531 | −4.6776 | 0.26[h] | 15 | 35.0 | 661 | $2.0^{+0.2}_{-0.2}$ | 153/0.615 | 925 | $8.7^{+0.4}_{-0.4}$ | 0.533 | $4.6^{+0.2}_{-0.2}$ | 4.1 |
| XLSS J022530.6−041420 | 041 | G02 | 36.3777 | −4.2391 | 0.14[h] | 15 | 45.0 | 523 | $1.34^{+0.14}_{-0.06}$ | 159/0.392 | 785 | $20.2^{+1.0}_{-1.1}$ | 0.440 | $2.4^{+0.1}_{-0.1}$ | 2.2 |
| XLSS J022532.2−035511 | 002 | B03 | 36.3844 | −3.9200 | 0.77[i] | 11 | 37.5 | 225 | $2.8^{+0.8}_{-0.5}$ | 66/0.489 | 200 | $2.8^{+0.3}_{-0.3}$ | 0.493 | $19.6^{+1.8}_{-1.8}$ | 5.8 |
| XLSS J022540.6−031121 | 050 | B13 | 36.4195 | −3.1894 | 0.14 | 13 | 86.0 | 1386 | $3.5^{+0.6}_{-0.5}$ | 177/0.436 | 1509 | $48.2^{+2.2}_{-2.4}$ | 0.804 | $9.3^{+0.5}_{-0.6}$ | 16 |
| XLSS J022559.5−024935 | 051 | B18 | 36.4982 | −2.8265 | 0.28 | 11 | 60.0 | 224 | $1.2^{+0.1}_{-0.1}$ | 57/0.242 | 169 | $1.5^{+0.6}_{-0.3}$ | 0.384 | $0.9^{+0.2}_{-0.1}$ | 4.4 |
| XLSS J022609.9−045805 | 011 | G11 | 36.5413 | −4.9682 | 0.05[h] | 9 | 67.5 | 424 | $0.64^{+0.06}_{-0.04}$ | 465/0.454 | 1045 | $11.6^{+1.1}_{-1.1}$ | 0.290 | $0.11^{+0.01}_{-0.01}$ | 0.6 |
| XLSS J022616.3−023957 | 052 | B18 | 36.5681 | −2.6660 | 0.06 | 2 | 46.5 | 529 | $0.63^{+0.03}_{-0.03}$ | 102/0.118 | 802 | $14.0^{+1.7}_{-1.4}$ | 0.285 | $0.25^{+0.02}_{-0.02}$ | 0.9 |
| XLSS J022709.2−041800 | 005 | G01 | 36.7885 | −4.3000 | 1.05[i,h] | 9 | 35.0 | 164 | $3.7^{+1.5}_{-1.0}$ | 96/0.777 | 179 | $1.1^{+0.1}_{-0.1}$ | 0.489 | $17.1^{+2.0}_{-2.0}$ | 8.9 |
| XLSS J022722.4−032144 | 010 | B10 | 36.8435 | −3.3623 | 0.33[f] | 9 | 60.0 | 452 | $2.4^{+0.5}_{-0.4}$ | 96/0.456 | 440 | $6.3^{+0.4}_{-0.4}$ | 0.574 | $6.1^{+0.5}_{-0.5}$ | 6.1 |
| XLSS J022726.0−043216 | 013 | G05 | 36.8586 | −4.5380 | 0.31[f,h] | 18 | 30.0 | 160 | $1.0^{+0.1}_{-0.1}$ | 417/1.899 | 536 | $2.1^{+0.2}_{-0.2}$ | 0.340 | $1.3^{+0.1}_{-0.1}$ | 1.3 |
| XLSS J022738.3−031758 | 003 | B15 | 36.9098 | −3.2996 | 0.84[i] | 13 | 48.0 | 231 | $3.3^{+1.1}_{-0.7}$ | 393/2.998 | 523 | $4.2^{+0.4}_{-0.4}$ | 0.518 | $37.8^{+3.5}_{-3.3}$ | 6.3 |
| XLSS J022739.9−045127 | 022 | G10 | 36.9165 | −4.8576 | 0.29[h] | 9 | 39.5 | 1305 | $1.7^{+0.1}_{-0.1}$ | 171/0.744 | 1791 | $9.8^{+0.3}_{-0.3}$ | 0.471 | $6.2^{+0.2}_{-0.2}$ | 3.2 |
| XLSS J022803.4−045103 | 027 | G10 | 37.0143 | −4.8510 | 0.29 | 7 | 60.0 | 438 | $2.8^{+0.6}_{-0.5}$ | 123/0.535 | 577 | $6.1^{+0.4}_{-0.4}$ | 0.653 | $4.8^{+0.4}_{-0.3}$ | 9.1 |

[a] *XMM* pointing identifiers from the *XMM*-LSS internal labelling as defined in table 1 of Pierre et al. (submitted), and displayed in Fig. 1.
[b] Number of spectroscopically confirmed cluster members.
[c] Flux in units of $10^{-14}$ erg s$^{-1}$ cm$^{-2}$ in the [0.5–2] keV band.
[d] Bolometric luminosity in units of $10^{43}$ erg s$^{-1}$.
[e] In units of $10^{13}\,h^{-1}\,M_\odot$. Note that these are just rough estimates based on the isothermal assumption.
[f] Already published in Willis et al. (2005a).
[g] Two foreground galaxies in the centre of the X-ray emission have a redshift of 0.08 – see Appendix B.
[h] Already published in Pierre et al. (2006).
[i] Already published in Valtchanov et al. (2004).

**Table 2.** Instrumental characteristics for each spectrograph configuration employed during the observations. All spectral observations were performed with a slit width between 1.0 and 1.4 arcsec.

| Telescope | Instrument | Grism + Filter | Wavelength interval (Å) | Pixel sampling (Å pixel$^{-1}$) | Spectral resolution[a] (Å) | Identifier |
|---|---|---|---|---|---|---|
| VLT | FORS2 | 300V + GG435 | 4000–9000 | 3.2 | 14 | 1 |
| VLT | FORS2 | 600RI + GG435 | 5000–8500 | 1.6 | 7 | 2 |
| VLT | FORS2 | 600z + OG590 | 7500–10 000 | 1.6 | 7 | 3 |
| Magellan | LDSS2 | Medium red | 4000–9000 | 5.1 | 14 | 4 |
| NTT | EMMI | Grism #3 | 4000–9000 | 3.0 | 8 | 5 |

[a]Estimated via the mean FWHM of the bright, isolated arc emission lines.

'best-fitting' parameters can be viewed as a physically realistic measure of the cluster properties, although they provide a useful ad hoc parametrization (Appendix A). For this reason, we do not quote here best-fitting values of $\beta$ and $R_c$ derived for each confirmed cluster. A dedicated analysis of the mean cluster profiles, obtained by stacking the data as a function of redshift and temperature, is underway (Alshino et al., in preparation).

The photon count rate within a specified radius is obtained by integrating the best-fitting spatial profile. The conversion into [0.5–2] keV flux and unabsorbed rest-frame bolometric ([0.001–50] keV) luminosity is performed via XSPEC using the cluster temperatures previously derived.

### 3.3.2 *Luminosity and flux determination*

Luminosities are integrated within $r_{500}$, that is, the radius at which the cluster mass density reaches 500 times the critical density of the Universe at the cluster redshift. As in Willis et al. (2005a) and Pierre et al. (2006), this radius is estimated from the cluster mean temperature using the $M$–$T$ relation of Finoguenov, Reiprich & Böhringer (2001), converted to $\Lambda$ CDM cosmology, which gives

$$r_{500} = 0.375\ T^{0.63}\ h_{73}(z)^{-1}\ \mathrm{Mpc}, \qquad (2)$$

where $T$ is expressed in keV and $h_{73}$ is the hubble constant in units of 73 km s$^{-1}$ Mpc$^{-1}$. Although 'total' fluxes are often quoted in the cluster literature, our simulations show that the present data do not allow us to reliably perform such measurements (see Appendix A). To limit extrapolation uncertainties, aperture flux values are computed by integrating within a fixed radius of 0.5 Mpc. As conspicuous in Table 1, 0.5 Mpc is generally smaller than $2R_{\rm fit}$ except for the nearby low-temperature groups XLSSC-039, XLSSC-021 and XLSSC-052. We chose to compute the flux within 0.5 Mpc (rather than $R_{500}$, which is generally smaller), since this is a similar approach to aperture photometry. For all clusters, it corresponds to about two-third of the total flux, that is, integrated to infinity, assuming a profile defined by $\beta = 2/3$ and $R_c = 180$ kpc. We emphasize that these fluxes are only used in the comparative analysis of the log($N$)–log($S$) relation, while $L_{500}$ and $M_{500}$ (along with $T$) are the actual physical quantities used in our cosmological modelling and in the subsequent discussion.

Finally, the $1\sigma$ errors on the extrapolated fluxes and luminosities are computed by identifying the region of the $(S_0, R_c, \beta)$ parameter space where $\Delta\chi^2 \leqslant 1$, and computing the extreme values of the extrapolated count rates allowed by these models (see Appendix A for further discussion on the flux measurement accuracy).

### 3.3.3 *Mass determination*

Based on our best spatial fit profile, we estimate cluster masses, assuming that the gas is isothermal (the limited number of photons does not allow us to derive temperature profiles). Under these assumptions, for $M_{500}$ in $M_\odot$, the hydrostatic equilibrium assumption yields (see e.g. Ettori 2000):

$$M_{500} = (1.11 \times 10^{14})\, \beta\, R_c\, T \left( \frac{x_{500}^3}{1 + x_{500}^2} \right), \qquad (3)$$

where $x_{500} = r_{500}/R_c$ (with $r_{500}$ given by equation 2), $R_c$ is expressed in Mpc, and $T$ in keV.

Values of $r_{500}$, flux and luminosity for the C1 clusters are listed in Table 1, together with $M_{500}$.

## 4 RESULTS

### 4.1 Global properties of the sample

The detected clusters span the [0.05–1.05] redshift range with a pronounced peak around $z \sim 0.3$ (Fig. 3). Roughly half of the objects have a temperature lower than 2 keV, pointing to a large fraction of groups in our sample. As a natural consequence of its sensitivity, good PSF and dedicated source-detection procedure, the *XMM*-LSS systematically unveils for the first time the $z \sim 0.3, T \lesssim 2$ keV cluster population on large scales. We note also that none of the objects detected so far exhibits strong lensing features. In the following, we will generally use the word 'cluster' regardless of the temperature, while 'group' specifically refers to objects below 2 keV. A

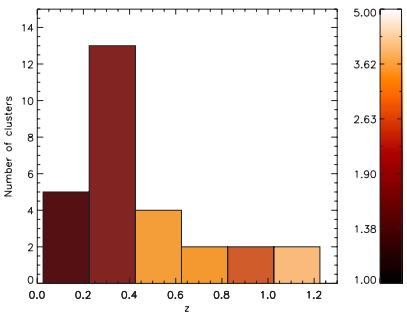


**Figure 3.** Redshift distribution of the C1 sample. The colour scale indicates the cluster mean temperature for each bin (unweighed mean of the individual cluster temperatures in keV) .

notable feature of the sample, illustrated by Fig. 3, is the increasing mean cluster temperature with redshift. Given the strong correlation between X-ray luminosity and temperature (Section 4.3), this can readily be understood as the unavoidable Malmquist bias. Moreover, the current sky coverage of the survey makes the sample likely to be affected by cosmic variance. Both issues will be given special attention in the following cosmological analysis.

Additional information on individual sources, including possible contamination by AGN emission, is discussed in Appendix B. We also display the optical images of each cluster with confirmed member galaxies and X-ray contours overlaid.

### 4.2 The $M_{500}$–$L_X$ relation

The correlation between cluster luminosities and masses derived from the spatial analysis is shown in Fig. 4. Once self-similar evolution is assumed, our data points appear to be continuous with the massive cluster samples of Zhang et al. (2006) and Zhang et al. (2007) (respectively, at $z \sim 0.3$ and $z \sim 0.2$). Moreover, the scatter in the relation is relatively low (i.e. almost comparable to that of the high-mass sample). This overall consistency is remarkable in the sense that, for half of the clusters, less than 500/300 photons were available for the spatial/spectral fitting. This provides an important indicator of the reliability of our measurements, and could be further interpreted as adding weight to the self-similar evolution hypothesis.

The comparison between our data points and the indicative $M$–$L_X$ relation derived from Arnaud, Pointecouteau & Pratt (2005) and Arnaud & Evrard (1999) also suggests some flattening of the correlation at the low-mass end. Our luminosity sampling, however, suffers significant Malmquist bias, and masses derived from a single isothermal $\beta$-model fitting are known to be underestimated (Vikhlinin et al 2006). Both effects tend to contribute to the apparent flattening in the $M$–$L_X$ relation.

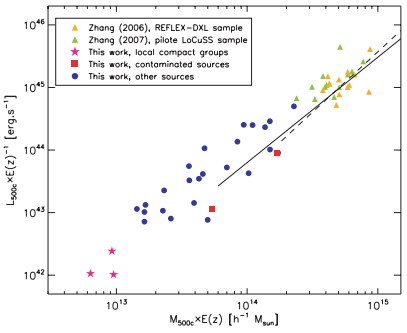


**Figure 4.** The mass–luminosity relation for the sample clusters. Because of the large redshift range spanned by the data, the mass and luminosity parameters are scaled assuming self-similar evolution [factor $E(z)$, see Section 4.3]. For comparison, the massive cluster samples from Zhang et al. (2006) and Zhang et al. (2007) are also shown. The dashed and solid lines also show the mean $M_{500}$–$L_X$ relation inferred from the $M_{500}$–$T$ and $L_X$–$T$ measured in Arnaud et al. (2005) and Arnaud & Evrard (1999), respectively, for $T > 3.5$ keV and $T > 2$ keV.

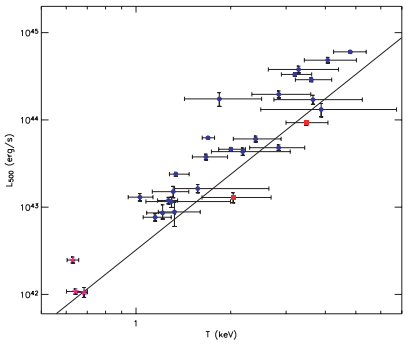


**Figure 5.** The $L_{500}$–$T$ relation. Same symbols as in Fig. 4. The sources that are used for the analysis of the $L_X$–$T$ evolution (circles) have redshifts in the range [0.14–1.05] and fall on average above the local $L_X$–$T$ relation of Arnaud & Evrard (1999) (shown by the solid line).

### 4.3 The $L_X$–$T$ relation

The correlation between the observed luminosity and the temperature of our clusters is shown in Fig. 5. The comparison with the local relation measured by Arnaud & Evrard (1999) suggests a positive redshift evolution of the luminosity at a given temperature. To test this, we performed an analysis of the enhancement factor, $F(z, T) = L_X/L_X(T, z = 0)$, with respect to this local reference. In most of the intracluster gas models, $F$ is a simple function of redshift: the self-similar assumption yields $F(z) = E(z)$, where $E(z) = H(z)/H_0$ is the evolution of the Hubble constant. More elaborate models that include non-gravitational physics (i.e. cooling and heating) generally propose an evolution of the form $F(z) = (1 + z)^\alpha$ (e.g. Voit 2005). All these models assume that the entire cluster population follows a unique evolutionary track. In this study, we adopted this hypothesis, since the Malmquist bias prevents us from investigating the extent to which evolution could be a function of cluster temperature.

To date, various attempts to constrain $F(z)$ as a power law in $(1 + z)$ have yielded discordant results (see Table 3). Several explanations for this discrepancy have been invoked: poor statistics, deviations between the several available local reference relations, disparity in mass and redshift of the present high-$z$ cluster samples (if $F$ is not a simple function of $z$), as well as biased samples. Although limited by photon statistics, the *XMM*-LSS cluster sample, with its well-controlled selection function, provides an

**Table 3.** Trends in the literature for the modelling of $F(z)$ as a power law of $(1 + z)$. For comparison, constraints from the C1 sample were added, ignoring the selection effects.

| Reference | Power of $(1 + z)$ |
|---|---|
| Vikhlinin et al. (2002) | $1.5 \pm 0.3$ |
| Novicki, Sornig & Henry (2002) | $2.1 \pm 1.1$ |
| Ettori et al. (2004) | $0.6 \pm 0.3$ |
| Lumb et al. (2004) | $1.5 \pm 0.3$ |
| Kotov & Vikhlinin (2005) | $1.8 \pm 0.3$ |
| Maughan et al. (2006) | $0.8 \pm 0.4$ |
| C1 clusters with $0.1 < z < 0.4$ (14 sources) | $2.3 \pm 0.8$ |
| C1 clusters with $0.4 < z < 1.1$ (10 sources) | $1.3 \pm 0.5$ |
| All C1 clusters above $z = 0.1$ (24 sources) | $1.5 \pm 0.4$ |

important opportunity to test the impact of the selection process in such studies.

In a preliminary analysis, Willis et al. (2005b) measured a mean value $F = 1.46$ from an initial *XMM*-LSS sample covering the [0.30–0.43] redshift range. Another estimate using the D1 subsample (Pierre et al. 2006, section 3.4) also found a luminosity enhancement for the $z \sim 0.3$ domain, and showed that it could be statistically significant only above $T \sim 1.5$ keV. However, none of these studies fully addressed the impact of selection biases on the derived results.

In this paper, involving a larger sample, we perform a self-consistent likelihood analysis of our data parameter space, for a given $F(z)$. We excluded from the analysis XLSSC-050 (contaminated by flares), XLSSC-018 (contaminated by an AGN), and the three local groups ($z \lesssim 0.1$ and $T \lesssim 1$ keV) for which our measurements could be affected by a large radial extrapolation and by X-ray emission from member galaxies. The available sample for this analysis is thus restricted to 24 objects.

#### *4.3.1 The likelihood model*

Folding our selection function into the analysis of the $L_X$–$T$ relation cannot be achieved by the usual $\chi^2$ fitting. Instead, we have to estimate the likelihood of our data, given a $F(z)$ model. We thus start by assuming that cluster luminosities are distributed by a log-normal probability density function (PDF) around the mean value $L_{\rm mean}(T, z) = L(T, z = 0) \times F(z)$, which we will call the true distribution:

$$\frac{{\rm d}P_{\rm true}(L_{\rm X} \mid T, z)}{{\rm d}\ln L_{\rm X}} = \frac{1}{\sqrt{2\pi}\sigma_{\ln L_{\rm X} \mid T}} \exp\left[-\frac{\ln (L_{\rm X}/L_{\rm mean})^2}{2\sigma^2_{\ln L_{\rm X} \mid T}}\right]. \quad (4)$$

An estimate of the scatter in the local $L_X$–$T$ relation is provided by Stanek et al. (2006). The authors fitted a single dispersion value to the Reiprich & Böhringer (2002) sample (thus assuming that it is the same from the group regime to large clusters) and obtained $\sigma_{\ln T \mid L_{\rm X}} = 0.25$. We further assume that this scatter does not evolve with redshift. Combining the lognormal distribution with the power-law shape of the mean $L_X$–$T$ relation, we get $\sigma_{\ln L_{\rm X} \mid T} \sim 2.7 \times 0.25 \sim 0.7$, where 2.7 is the approximate $L_X$–$T$ slope in the Reiprich & Böhringer (2002) sample.

The normalized PDF for detection of a cluster with such $L_X$ and $T$ is then the product of the 'true' PDF with our selection function:

$$\frac{{\rm d}P_{\rm det}(L_{\rm X} \mid T, z)}{{\rm d}\ln L_{\rm X}} = P_0(T, z) \times \frac{{\rm d}P_{\rm true}(L_{\rm X} \mid T, z)}{{\rm d}\ln L_{\rm X}} \times f(L_{\rm X}, T, z), \quad (5)$$

where

$$P_0(T, z) = \left(\int \left[\frac{{\rm d}P_{\rm true}(L_{\rm X} \mid T, z)}{{\rm d}\ln L_{\rm X}} \times f(L_{\rm X}, T, z)\right] {\rm d}(\ln L_{\rm X})\right)^{-1}. \quad (6)$$

We compute $f(L_X, T, z)$ (from Fig. 2), assuming a canonical value of 180 kpc for the cluster core radii. The resulting modification of the PDF is illustrated in Fig. 6 at different redshifts assuming self-similar evolution: including selection effects in the likelihood model renders underluminous cool clusters undetectable. Note that for clusters with temperatures close to the detection threshold at any given redshift, the bias in mean $L_X$ at a given $T$ due to selection effects can be a factor of 2 or more.

Practically, errors on the luminosity can be considered as negligible, since the fractional errors on temperature are usually much larger and the $L_X$–$T$ relation goes like $T^3$. The likelihood of an observed cluster C being drawn from a given $F(z)$ is thus obtained by

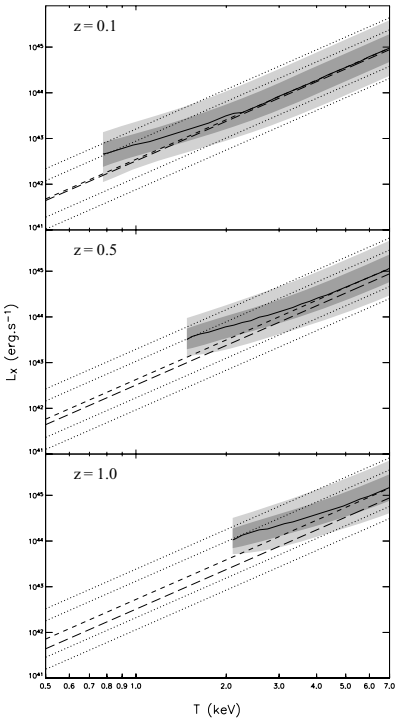


**Figure 6.** Comparison of the modelled $L_X$–$T$ distribution with and without selection effects for several redshifts, assuming self-similar evolution. The light and dark-shaded regions show, for each temperature, the luminosity interval than contains, respectively, 95.5 and 68.3 per cent of the expected detections. The solid line gives the maximum of the luminosity PDF for each temperature; it is cut at the temperature for which the detection probability (integrated over $L_X$) falls to 5 per cent. For comparison, the local Arnaud & Evrard (1999) relation, evolved according to the self-similar model, is indicated by the short-dashed line, and the 1 and 2$\sigma$ bounds arising from our scatter model are shown by the dotted lines. The long-dashed line is the local relation.

marginalizing over the possible cluster temperatures, that is,

$$L(F \mid {\rm C}) = \int_0^{+\infty} \left[\frac{{\rm d}P_{\rm det}(L_{\rm X} \mid T, z)}{{\rm d}\ln L_{\rm X}}\right] P_{\rm C}(T)\, {\rm d}T, \quad (7)$$

where $P_C(T)$, the temperature PDF of the cluster, is derived from the C-statistic distribution provided by the XSPEC's STEPPAR command. Finally, the likelihood of the enhancement factor is computed as the product of the detection probabilities, for the given $F(z)$, of all clusters pertaining to our sample:

$$L(F) = \prod_{i=0}^{N} L(F \mid {\rm C}_i). \quad (8)$$

To estimate errors on fitted parameters, we identify $-2 \ln L(F)$ with a $\chi^2$ distribution, as is asymptotically valid in the limit of

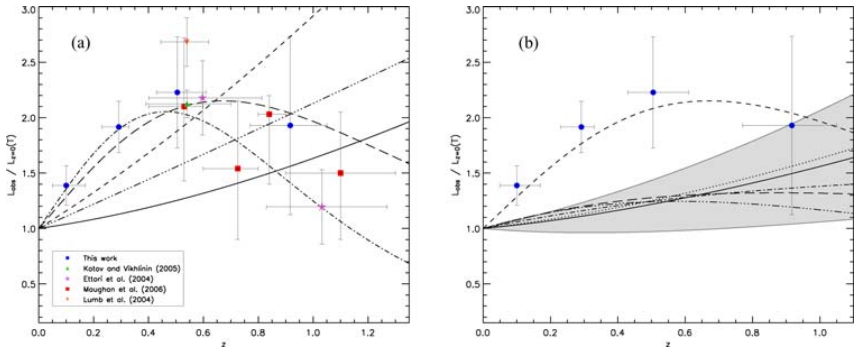


**Figure 7.** Evolution of the $L_X$–$T$ relation. (a): *Raw analysis*. The data points from several studies are plotted (stacked into redshift bins for clarity). Whenever possible, we converted luminosity estimates from the other samples to $L_{500}$ using the information provided by the authors. The following differences remain: the luminosities of the Lumb et al. (2004) clusters are estimated within the virial radius; the data points of Maughan et al. (2006) also include the clusters from Vikhlinin et al. (2002) and have luminosities estimated within $R_{200}$. Overlaid are several enhancement factor fits from our baseline analysis: the $(1 + z)$ power law and ad hoc two-parameter fits to our data alone (respectively short- and long-dashed lines), and the $(1 + z)$ power law and ad hoc two-parameter fits (respectively, dot–dashed and three-dot–dashed lines) fitted to the C1 clusters combined with the samples of Kotov & Vikhlinin (2005) and Ettori et al. (2004). For comparison, the self-similar prediction is indicated by the solid line. (b): *Taking into account selection effects*. The filled circles and short-dashed line recall the location of our raw data points and best-fitting model from the preliminary analysis. The corrected enhancement factor fit for the two-parameter model is shown as the long-dashed line. The final one-parameter fit and its $1\sigma$ confidence interval are displayed by the solid line and the shaded region. Expectations from several intracluster gas models are plotted for comparison: the self-similar predictions (dotted line) and two models by Voit (2005) including non-gravitational physics (dot–dashed and three-dot–dashed lines).

large samples according to Wilks' theorem, and quote 68 per cent confidence intervals. The same identification is used to assess how much a given model deviates from the best-fitting $F(z)$ by correcting the measured $\Delta\chi^2$ to the equivalent value that yields the same probability for one free parameter.

#### 4.3.2 Raw fit

In a first step, we neglect the selection function in our formalism (i.e. effectively imposing $f(L_X, T, z) = 1$ in the likelihood computation). This will both provide us with a reference point to assess the impact of the selection process, and allow for a comparison with previous work, where the issue of sample selection has been ignored.

As shown in Table 3, fitting the usual $F(z) = (1 + z)^\alpha$ yields $\alpha = 1.5 \pm 0.4$, which is consistent with Vikhlinin et al. (2002), Lumb et al. (2004) and Kotov & Vikhlinin (2005). Nevertheless, this simple model does not seem to correctly reproduce the observations over the full redshift range: the values of $\alpha$ obtained over $[z = 0.1$–$0.4]$ and $[z = 0.4$–$1.1]$ are incompatible at the $1\sigma$ level (see Table 3). Interestingly, Ettori et al. (2004) noted a similar discrepancy within their own sample, but did not consider its implications for their subsequent analysis.

In Fig. 7(a), we display simultaneously our data points along with those from previous studies. Error bars are large, but the points suggest a non-monotonic evolution, with a maximum around $z \sim 0.5$, thus excluding an enhancement factor of the form $F(z) = (1 + z)^\alpha$. This may explain the diverging results of Table 3. Consequently, we fitted an ad hoc two-parameter model of the form $F(z) = (1 + z)^\alpha \times E(z)^\beta$. This model is intrinsically degenerate and would need very accurate data to individually constrain $\alpha$ and $\beta$, but has the ability to reproduce the apparent non-monotonic evolution. Since cases $\alpha = 0$, $\beta = 1$ and $\beta = 0$ correspond to the self-similar and power-law models, respectively, we may evaluate through this parametrization how much the observations deviate from any of them.

A fit over the C1 clusters yields $\alpha = 4.7$ and $\beta = -5.4$ and excludes the self-similar evolution at the $3.5\sigma$ level. Our best $(1 + z)^\alpha$ model is, however, less than $0.4\sigma$ away from the best fit and the evidence of a non-monotonic evolution is thus rather weak from our data alone (the probability of such a deviation being real is roughly 45 per cent). This is not the case any longer when including also in the fit the data from Kotov & Vikhlinin (2005) and Ettori et al. (2004),[6] for which we could convert the luminosity to $L_{500}$. A simple power-law model is then ruled out at the $8.4\sigma$ level based on 59 sources.

#### 4.3.3 Taking into account selection effects

With this comparison baseline, we can now turn to investigate the impact of our selection process on the observed $L_X$ enhancement (restoring the correct value of $f(L_X, T, z)$ in the likelihood). This task is tractable here because the *XMM*-LSS C1 selection function has been thoroughly assessed.

Maximizing the likelihood over our ad hoc two-parameter model now yields $\alpha = 1.3$ and $\beta = -1.3$. As can be seen in Fig. 7(b), the inferred evolution rate is considerably lower than that with the raw fit. As a result, the self-similar expectation is now less than $7 \times 10^{-2}\sigma$ away from the best-fitting model, and the statistical evidence for a non-monotonic evolution is virtually null. To evaluate the remaining deviation from self-similarity, we switch back to a one-parameter model by fixing $\beta = 1$. The best-fitting value for $\alpha$ is then $-0.07^{+0.41}_{-0.55}$, and the corresponding $F(z)$ is shown in Fig. 7(b).

[6] We assumed a Gaussian PDF of the cluster temperature for these sources in our likelihood model.

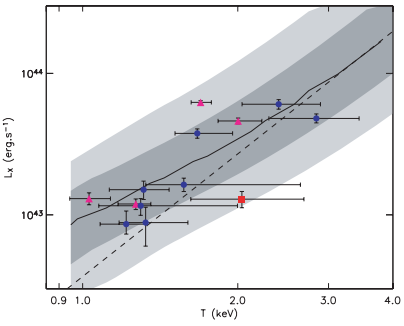


**Figure 8.** Similar to figure 6 for $z = 0.3$, with the $0.2 < z < 0.4$, C1 clusters overlaid. Again, the dashed line is the evolved model of Arnaud & Evrard (1999). The triangles indicate the D1 clusters presented by Pierre et al. (2006). The only cluster (XLSSC-018, square) that falls well below the Arnaud & Evrard (1999) $L_X$–$T$ relation, and at the limit of our $2\sigma$ contour, is likely to be contaminated by an AGN (see Appendix B).

The latter analysis points to a negative evolution of the $L_X$–$T$ relation (i.e. $L(z \mid T)$ increases with $z$) over the $0 < z < 1$ range. These results quantitatively favour the self-similar hypothesis, although our current data set is still marginally consistent with no evolution at all as is clear from Fig. 7(b).

This is illustrated in more detail in Fig. 8 for the $z \sim 0.3$ intermediate redshift, where the 1–2 keV groups make up the peak of the *XMM*-LSS sensitivity (cf. Section 5).

These results are still preliminary because of the poor statistics (in both the number of sources and the temperature estimates). Moreover, they are likely to be very sensitive to the choice of the local $L_X$–$T$ reference (including scatter) due to the redshift–temperature correlation within the sample. Nevertheless, our findings seriously question any attempt to assess the evolution of cluster scaling laws without modelling the full source-selection process.

In the rest of this paper, we assume that the cluster scaling laws evolve following the self-similar model, as suggested by the above analysis.

## 5 COSMOLOGICAL MODELLING

In this section, we compare the observed properties of the C1 sample with cosmological expectations based on the latest *WMAP* constraints (Spergel et al. 2006).

### 5.1 The model

We model the cluster population and its evolution following Pacaud et al. (2006), but with slightly modified assumptions as to the cluster scaling laws.

The linear power spectrum $P(k)$ of dark matter fluctuations is computed using an initial power law of index $k^{n_s}$ and the transfer function from Bardeen et al. (1986), taking into account the shape parameter:

$$\Gamma = \Omega_m h \exp\left[-\Omega_b\left(1 + \frac{\sqrt{2h}}{\Omega_m}\right)\right] \quad (9)$$

introduced by Sugiyama (1995). The overall power spectrum is normalized at $z = 0$ to $\sigma_8$, and the redshift evolution is obtained from the linear growth factor (Carroll, Press & Turner 1992). Based on this power spectrum, we use the Jenkins et al. (2001) formalism to derive the dark matter halo mass function $\mathrm{d}\,n/\mathrm{d}\,m(z)$. This procedure is an accurate fitting of the mass function obtained from numerical simulations, provided one defines the mass of the halo to be that included inside $r_{200b}$, the radius enclosing an overdensity of 200 with respect to the mean background density.

In order to reproduce the C1 selection function, we need to translate the mass parameter into observable parameters, which is practically achieved by means of scaling laws. Usually, halo masses in such relations are defined with respect to the critical density of the Universe (and not to the background density), hence the need to assume a model for the halo profiles to connect the two mass definitions. For this purpose, we used NFW profiles (Navarro et al. 1997) with a scaling radius $r_s$ provided by the model of Bullock et al. (2001) which relates $r_s$ to the virial mass of the halo through the concentration parameter $c = r_{vir}/r_s$. The $M_{500}$–$M_{vir}$ conversion was obtained from the formulae provided in appendix C of Hu & Kravtsov (2003), with the estimate of the overdensity within the virial radius from Kitayama & Suto (1996).

The gas temperature within the dark matter haloes is then computed using the $M_{500}$–$T$ relation of Arnaud et al. (2005), that is, a power low of slope $\alpha = 1.49$, valid for clusters with $T > 4$ keV. To account for the steepening of the relation at lower temperatures suggested by their data, we used a higher slope of $\alpha = 1.9$ below 3 keV. The lower halo mass limit is then fixed by imposing $T > 0.5$ keV. Bolometric luminosities are derived using the $L_X$–$T$ relation of Arnaud & Evrard (1999). Self-similar evolution is assumed for both these scaling laws.

To account for the scatter observed in cluster properties, we encapsulate the dispersion of the $M$–$T$ and $L$–$T$ relations in the $M$–$L$ relation, for the sake of simplicity. We assign the X-ray luminosity assuming a lognormal distribution. Following Stanek et al. (2006), who measured $\sigma_{\ln M \mid L} = 0.37$, we use $\sigma_{\ln L \mid M} \sim 0.37 \times 1.59 \sim 0.6$, where 1.59 is the slope of their $M$–$L_X$ relation. The total *XMM*–*Newton* EPIC count rate is estimated using the same spectral model as described in Section 3.2, with a fixed hydrogen column density of $2.6 \times 10^{20}$ cm$^{-2}$ (representative of our field), folded through the EPIC response matrices for the THIN filter in accordance with our observing mode. The selection function (Fig. 2) is finally applied assuming, as in the previous section, a constant physical core radius of 180 kpc.

In what follows, given the still modest size of the present cluster sample (some 30 objects over 5 deg$^2$), we restrict most of the comparisons to qualitative visual ones.

### 5.2 The redshift distribution

Our model (using $\sigma_8 = 0.74$, from *WMAP*3) predicts 6.2 C1 clusters per square degree which, assuming Poisson noise alone (i.e. $\pm 1.1$ for the current 5.2 deg$^2$), is compatible with our observed density of 5.4 deg$^{-2}$ as given by the objects listed in Table 1.

The observed redshift distribution of the C1 sources, shown in Fig. 9, shows good overall agreement with the model expectations, and suggests that we are crossing a void-like region within $0.4 < z < 0.6$. As is evident from the errors in Fig. 9, this apparent underdensity can be fully accounted for by statistical plus cosmic variance. We estimated the cosmic variance using the formalism from Hu & Kravtsov (2003), under the assumption that the surveyed volume is enclosed within a top-hat sphere; in the sensitivity regime of the *XMM*-LSS survey, the total uncertainty on the $N(z)$ bins scales roughly as $2\sqrt{N}$.

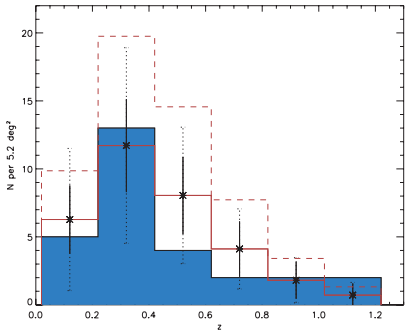


**Figure 9.** Redshift distribution of the C1 sample. The filled histogram shows the redshift distribution of our sample, while the solid line shows the expectations of our cosmological model (*WMAP3*: $\sigma_8 = 0.74$ and self-similar evolution for the $L_X$–$T$ relation). Expected fluctuations around the mean model density due to shot noise and (shot noise + cosmic variance) are, respectively, displayed with the solid and dotted error bars. The dashed line shows the expectations for a model with *WMAP* first-year cosmological parameters ($\sigma_8 = 0.85$) and a non-evolving $L_X$–$T$ relation; combining the latter $\sigma_8$ value with self-similar evolution would produce a normalization even higher, by about 35 per cent.

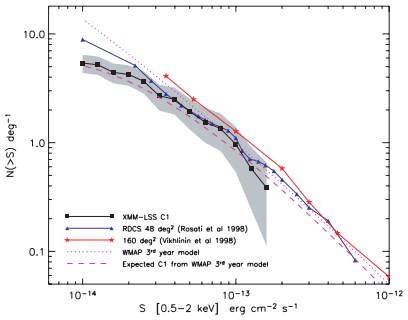


**Figure 10.** The C1 log($N$)–log($S$). The squares show the observed C1 number counts and the shaded region indicates the $1\sigma$ error region. Overlaid are the predictions from our model: the dashed line for C1 clusters, and the dotted line for the whole cluster population (i.e. all clusters above a given flux). Since only a fraction of cluster fluxes are detected in any survey, there is an important caveat in this comparison: our model delivers total cluster fluxes (i.e integrated to infinity) while our measurements are performed within 0.5 Mpc (which includes approximately two-third of the total flux for a cluster with $R_c = 180$ kpc and $\beta = 2/3$). Previously, Vikhlinin et al. (1998) claimed to measure total fluxes in an unbiased manner, while Rosati et al. (1998) estimated that they had recovered at least 80–90 per cent of the total flux.

Interestingly, the C1 cluster density predicted assuming $\sigma_8 = 0.85$ (as inferred by Spergel et al. 2003 from the first-year *WMAP* data) is 14.6 deg$^{-2}$ and appears to be much higher than required to match our observations. In this case, even using a non-evolving $L_X$–$T$, which is roughly our $1\sigma$ lower bound from the previous section, our model predicts 10.7 C1 clusters per square degree. Our data are $3.7\sigma$ (considering Poisson fluctuations) below the latter model. Even including cosmic variance, the difference is approximately $2\sigma$. This suggests that increasing the area by a factor of 2 would already permit a stringent test of the value of $\sigma_8$ indicated by *WMAP*.

In this context, we examine in detail in Section 6.3 some parameters which play a role in the interpretation of the cluster number counts.

### 5.3 The flux and luminosity distribution

As our selection function does not depend on flux alone, as assumed in previous generations of surveys, we have no direct estimate of the log($N$)–log($S$) relation. In this section, we provide the necessary information to enable the intercomparison between our statistical distributions and those of past as well as future surveys.

Our cosmological model can be both compared to the local log($N$)–log($S$) and used to predict an expected distribution for our C1 sample . This is shown in Fig. 10. The raw log($N$)–log($S$) from our model is first compared with the estimates from Rosati et al. (1998) and Vikhlinin et al. (1998) based on flux-limited samples, and said to be corrected for selection effects as well as for incompleteness. We find overall consistency, although our model suggests a somewhat higher total cluster density than the RDCS measurements above a flux of $\sim 10^{-14}$ erg s$^{-1}$ cm$^{-2}$. Furthermore, the predicted flux distribution, once folded with the C1 selection function, is consistent within the errors with the C1 log($N$)–log($S$) relation derived from the present sample. This adds credibility to our full selection process and log($N$)–log($S$) modelling.

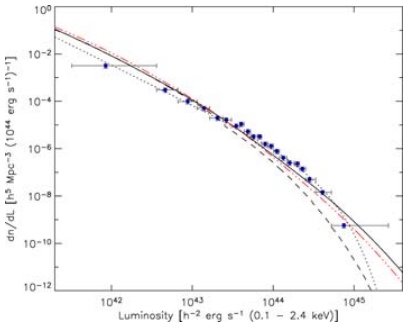


**Figure 11.** The local luminosity function predicted by our model, compared with the REFLEX data points (Böhringer et al. 2002). The dotted line is the best fit to the REFLEX data from the original paper. The dashed line is the expectation from our model assuming no scatter in the $M$–$L_X$ relation. The solid line shows our model (including scatter) yielding a much better agreement. The three-dot–dashed line is the prediction of our model for a redshift of 0.8.

Similarly, our sample is currently too small to derive the cluster luminosity function at any redshift, but this again can be predicted by our model. We first compute as a cross-check the local luminosity function and compare it with the estimates from the REFLEX survey (Böhringer et al. 2002), which constitutes the largest complete cluster sample currently available at low redshift (Fig. 11). The model shows, on average, good agreement with the measured REFLEX luminosity function (which accounts for the survey sky coverage). It lies slightly below the REFLEX measurements for luminosities in

the range $5 \times 10^{43}$–$5 \times 10^{44}$ erg s$^{-1}$, behaviour that is also present in the *ROSAT* Bright Cluster Sample (Ebeling et al 2000) as can be appreciated in the comparative plots of Mullis et al. (2004). We further emphasize that, as is conspicuous in Fig. 11, the inclusion in the model of scatter in the $M$–$L_X$ relation is necessary in order to obtain predictions that are compatible with the REFLEX constraints at the high-$L_X$ end. Finally, we note that the prediction of our model at $z = 0.8$ (also displayed in Fig. 11) is very similar to the local one out to $L_X = 10^{44}$ erg s$^{-1}$, and about a factor of 2 lower at $L_X = 10^{45}$ erg s$^{-1}$ (although still within the REFLEX error bars). Since EMSS (Gioia et al. 1990), the question of the evolution of the cluster luminosity function has been controversial, especially for luminosities above $5 \times 10^{44}$ erg s$^{-1}$ (see a review by Mullis et al. 2004 including several *ROSAT* samples). The most-recent determination of the cluster luminosity function for $L_X < 10^{44}$ erg s$^{-1}$ from the COSMOS–*XMM* data shows no evolution (Finoguenov et al. 2007). Our calculations suggest that, interestingly, for a concordance cosmology, the observed luminosity function does not significantly evolve out to $z \sim 0.8$ for $L_X < 10^{44}$ erg s$^{-1}$, while the cluster mass function and scaling laws do. This can be regarded as a 'conspiracy' between cosmology and cluster physics.

### 5.4 Mass-observable relations

Last but not the least, our model provides us with a tool for deriving heuristic mass-observable relations. For this purpose, we adopt the very general parametrization introduced by Hu (2003):

$$\mathrm{OBS} = \mathrm{OBS}_0 \left(\frac{M_\mathrm{d}}{M_0}\right)^{p(z)} \mathrm{e}^{A(z)}, \qquad (10)$$

where $M_\mathrm{d}$ is the mass defined within some overdensity $d$, OBS is the observable of interest (flux, count rate) and $M_0$ and $\mathrm{OBS}_0$ are reference values for both quantities. In practice, we used $M_0 = 10^{12}\,h^{-1}\,\mathrm{M}_\odot$ and $M_\mathrm{d} = M_{200\mathrm{b}}$, that is, the mass enclosed within the radius delineating an overdensity of 200 with respect to the mean background density. Model data points for a large range of masses, redshifts and fluxes (count rates) were fitted with the above formula as illustrated in Fig. 12. We found that the functions $p(z)$ and $A(z)$

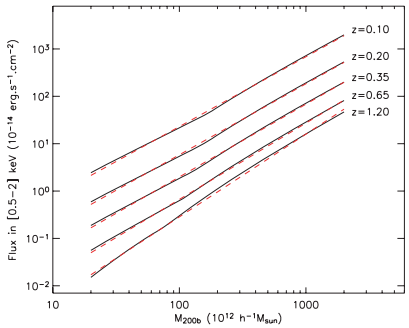


**Figure 12.** Calibration of the [0.5–2] keV band total flux versus $M_{200\mathrm{b}}$ relation. The exact correlations derived from our model are shown for several redshifts with plain lines. The overplotted dashed lines show the recovered values using the fit presented in Table 4.

**Table 4.** Parameters of the mass-observable calibration from our cosmological model, as defined by equations (10) and (11).

| Parameter | Mass–flux relation[a] | Mass–count rate relation[b] |
|---|---|---|
| $\mathrm{OBS}_0$ | 0.0329 | $1.23 \times 10^{10}$ |
| $a_0$ | −37.68 | −36.50 |
| $a_1$ | 0.0645 | 0.0680 |
| $p_0$ | 1.45 | 1.41 |
| $p_1$ | 0.248 | 0.269 |

[a]For a total source flux in the [0.5–2] keV band in erg s$^{-1}$ cm$^{-2}$.
[b]For a MOS1 + MOS2 + PN count rate in the [0.5–2] keV band in counts s$^{-1}$.

are sufficiently well described by functionals of the form

$$\begin{cases} p(z) = p_0 + p_1\, z \\ A(z) = a_0\, z^{a_1}. \end{cases} \qquad (11)$$

The best-fitting values for the mass versus count rate and mass versus flux relations are provided in Table 4. We do not give errors on the parameters as the overall accuracy of both fits with respect to the model data is better than 15 per cent over the $0.05 < z < 1.2$ and $2 \times 10^{13} < M < 2 \times 10^{15}\ h^{-1}\,\mathrm{M}_\odot$ ranges. In practice, the intrinsic dispersion of the $M$–$L$ relation also needs to be taken into account. Assuming $\sigma_{\ln L|M} \sim 0.6$ (Section 5.1), this translates into a dispersion of −45 to +82 per cent in the flux–mass or count rate–mass relations. Thus, 'perfect' flux measurements may yield, via this formula, mass accuracies of the order of −60 to +100 per cent. Such an empirical relation has obvious useful practical usages. It should be noted, however, that extrapolating the formulae in the present form above $z > 1.2$ is not straightforward as a number of prominent low-temperature lines (O VII, Fe VII, Fe VIII) are redshifted below the 0.5-keV boundary, creating discontinuities that cannot be accounted for with the above simple functionals.

## 6 DISCUSSION

We have shown that an accurate understanding of the sample-selection effects is essential for a proper study of the evolution of the $L$–$T$ relation of clusters of galaxies. Once these are allowed for, our data are compatible with self-similar evolution but cannot exclude no evolution at all. Subsequently, we have defined a cosmological cluster evolutionary model to which we applied the survey-selection function. It matches well our observations in terms of the cluster number density, redshift and flux distribution, hence providing a theoretical basis to further explore the ability of large *XMM* cluster surveys to constrain cosmology.

In this respect, one of the critical issues is the role and the number of cluster parameters involved in the cluster scaling law evolution. As a first step, we have shown that for our best-fitting model, it is possible to derive good mass-observable relations for the entire population of interest. We realistically explore below the impact of various parameters and hypotheses on the cluster-scaling laws, on the selection function and, subsequently, on cluster number counts. In this way, we hope to pave the way for future investigations.

### 6.1 Cluster shape parameters

Our selection function, as applied to our model, assumes a fixed core radius of $R_\mathrm{c} = 180$ kpc. Since $R_\mathrm{c}$ is related to the overall physical size of the system, one naturally expects this value to depend on the cluster mass (or on any related parameter). This should ideally be

included in the cosmological analysis, but we are currently lacking a well-established $R_c$–$T$ relation. The slope reported by Ota & Mitsuda (2004) is, for instance, much stronger than the one observed in the Sanderson & Ponman (2003) sample. Ota & Mitsuda (2002) also revealed that, due to cool cores, the core radius distribution inferred by performing a single $\beta$-model analysis over the whole cluster population is actually double peaked and thus ill-defined.

Furthermore, another known observational feature of X-ray groups and clusters is the tendency of the $\beta$ parameter to drop with decreasing system temperature (see e.g. Jones & Forman 1999, or more recently Osmond & Ponman 2004), but this correlation remains poorly defined, especially in the group regime. As our selection function is mainly surface brightness limited, we expect the detection probability to depend on $\beta$ as well as $R_c$ and so, two additional scaling relations are needed.

One could, in principle, derive them from the observed sample but, given the generally small number of collected photons, it is not possible to resolve precisely the $\beta$–$R_c$ degeneracy for most of our sources. Furthermore, such a study would be complicated by the fact that our selection process is dependent on these parameters, thus requiring a sophisticated self-consistent modelling similar to what was done in Section 4.3 for the $L_X$–$T$ relation.

Given the small number of systems in the present sample, we postpone such a study to a future paper investigating the evolution of mean cluster/group profiles by stacking images (Alshino et al., in preparation)

### 6.2 The evolution of the $L_X$–$T$ relation

The constraints we were able to put on the evolution of the $L$–$T$ relation ($\alpha = -0.07^{+0.41}_{-0.55}$ for the $F(z) = E(z)(1+z)^{\alpha}$ model) are not only limited by the relatively modest size of the sample and the uncertainties on the temperature measurements, but also depend on the assumed scatter of the relation (this parameter was fixed in the fit).

#### *6.2.1 Impact of the assumed scatter value*

In order to quantify the impact of the scatter in the determination of the evolution of the $L$–$T$ relation, we ran a series of fits assigning different values to $\sigma_{\ln L|T}$. We find that the $\alpha$ index describing the deviation from self-similarity can be modelled by the following empirical formula:

$$\alpha = 1.78 - \left(\frac{\sigma_{\ln L|T}}{0.45}\right)^{1.41}. \quad (12)$$

The uncertainty on $\sigma_{\ln T|L}$ of $\sim$25 per cent quoted by Stanek et al. (2006) along with the $L_X$–$T$ slope uncertainty (say between 2.5 and 3) yield an uncertainty of $\sim$20 per cent for $\sigma_{\ln L|T}$ (same reasoning as in Section 4.3.1). This translates to a conservative confidence range of [0.5–0.9] for $\sigma_{\ln L|T}$, which in turn gives fitted values of $\alpha$ in the range [0.6, $-$0.87]. The corresponding deviation is thus larger than the statistical errors from our fit, and precise constraints on $\sigma_{\ln L|T}$ are definitely required in order to place firm and rigorous constraints on $F(z)$. Given that the scatter in the $L_X$–$T$ relation could well be a function of redshift and mass, one will have to wait for large samples of high-$z$ clusters with well-monitored selection effects in order to undertake a fully self-consistent determination of the evolution of the slope and scatter of the $L_X$–$T$ relation.

**Table 5.** Expected accuracy in the determination of the evolution of the $L_X$–$T$ ($z$) $\propto E(z)(1+z)^{\alpha}$ relation for various sizes of *XMM*-LSS type surveys, estimated from simulations. The temperature measurements come either directly from the 10-ks *XMM* exposures, with corresponding accuracy, or, part of them are improved by subsequent deep *XMM* observations, assuming an uncertainty of 10 per cent; the last column gives the number of such clusters to undergo deep *XMM* pointings. $\sigma_\alpha$ is the mean $1\sigma$ error on $\alpha$ for a survey realization.

| Area | Temperature accuracy | $\sigma_\alpha$ | $N_{\mathrm{re-obs}}$ |
|---|---|---|---|
| (deg$^2$) | | | |
| 5 | From the survey | 0.59 | – |
| 20 | " " | 0.28 | – |
| 64 | " " | 0.15 | – |
| 5 | 10 per cent for the $z > 0.8$ clusters | 0.58 | 2 |
| 20 | " " | 0.25 | 14 |
| 64 | " " | 0.14 | 41 |
| 5 | 10 per cent for the $z > 0.5$ clusters | 0.50 | 9 |
| 20 | " " | 0.22 | 45 |
| 64 | " " | 0.13 | 139 |

#### *6.2.2 Impact of the sample size and of the uncertainties on the temperature*

We have further investigated the extent to which the determination of the slope of the evolution of the $L_X$–$T$ relation is conditioned by the size of the cluster sample and by the magnitude of the error bars on the temperature (in comparison, errors on the luminosity are assumed to be negligible). We have thus simulated a large number of random cluster catalogues corresponding to 5, 20 and 64 deg$^2$ *XMM*-LSS type surveys. The realizations are drawn from our cosmological model, reproducing our observed $n(z, T)$ distribution along with selection effects; self-similarity is assumed for the $L_X$–$T$ evolution. Each cluster temperature is assigned a mean error estimated from the spectral fitting simulations presented in the appendix of Willis et al. (2005a), for 10-ks *XMM* exposures. In addition, we have considered the possibility of dedicated deep *XMM* follow-up observations providing a temperature accuracy of 10 per cent for cluster subsamples of various sizes. For each sample, the slope was fitted following the method described in Sections 4.3.1 and 4.3.3. Results are summarized in Table 5.

The simulations show that a significant improvement can be reached by increasing the sample size as the accuracy on $\alpha$ scales roughly as the square root of the surveyed area. This is a very notable result, given that two-third of our clusters have no more than 500 counts available for the spectral fit. The spectral accuracy is of the order of 20 per cent below 2 keV, and 50 per cent around 5 keV (Willis et al. 2005a), the latter concerning mostly distant clusters. The simulations further show that increasing the accuracy of the temperature measurements does not yield a very significant improvement in the slope of the $L_X$–$T$ relation – compared to the amount of time that would be necessary to obtain accurate temperatures for the $0.5 < z < 1$ hot clusters. This is mostly due to the fact that the dispersion in the $L_X$–$T$ relation itself is large. Quantitatively, applying the $\sqrt{\mathrm{area}}$ rule of thumb in Fig. 7(b) shows that a 10-deg$^2$ area (or any sample with a similar size and controlled selection effects) has the potential to exclude the non-evolution hypothesis. However, even with 600 clusters (i.e. a 64-deg$^2$ area, and 140 very well measured clusters above $z > 0.5$), it seems difficult to, for example, discriminate between the two modified entropy models of cluster evolution proposed by Voit (2005).

*6.2.3 Working on samples close to the detection limit*

Fig. 8 shows that for the $0.2 < z < 0.4$ range, most of our clusters are detected in a $L_X$–$T$ region where the selection effects are significant, as is to be expected, given that lower mass haloes are much more numerous than massive ones. As a result, overluminous clusters will be over-represented in our sample. For the present data set, this is also true at any redshift, and this situation is easily understandable as a result of the combined effects of (1) the high efficiency of XAMIN close to the detection limit (it was designed to enable the construction of the largest possible controlled cluster sample), and (2) the steepness of the cluster mass function. Generally speaking, given the contradictory former results obtained from data based on heterogeneous *ROSAT*-selected samples, our findings suggest that graphs such as Fig. 8 should be first carefully constructed, when studying the cluster $L_X$–$T$ relation at any redshift.

### 6.3 Modelling the cluster number density

While most of the ingredients linking cosmology to the expected cluster distribution are rather well established (mass functions, mean scaling relations for massive clusters), many details of the mass-observable issues are still pending. Among these, we would cite: the evolution of cluster scaling laws, the scaling laws in the group regime, and the role of the scatter in all relations. Our analysis has shown that our current data are compatible with self-similar evolution for the $L$–$T$ relation.[7] In this section, we thus focus on two other specific aspects, namely the slope of the $M$–$T$ relation for the groups and the scatter in the $M$–$L$ relation. We illustrate below the sensitivity of the cluster number counts to these quantities by means of our model.

*6.3.1 The M–T relation in the group regime*

Our model assumes that the $M$–$T$ relation for both massive clusters and groups can be modelled by a broken power law. Recent measurements (Arnaud et al. 2005; Vikhlinin et al. 2006) agree very well in the high-mass regime, but yield contradictory results for groups. Vikhlinin et al (2006) claimed that the previously reported steepening of the relation (by e.g. Finoguenov et al. 2001) resulted from a bias towards low mass, due to an incorrect modelling of the density profile at large radius. However, their assertion relied on only two low-temperature systems. On the other hand, Arnaud et al. (2005) found a steepening at a low temperature based on a larger sample of groups, but their analysis assumes an NFW profile at a large radius. As this point is presently still controversial, we adopted an intermediate solution: a steepening to 1.9 of the relation below 3 keV (note that a change in the $L_X$–$T$ relation for groups would have similar impact). Fig. 13 illustrates how the various hypotheses on the shape of the $M$–$T$ relation in the low-mass regime impact on the cluster number counts. Our assumed broken power-law model provides the best fit on the basis of a visual inspection, but the difference to the Arnaud et al. (2005) model is not statistically significant with the present sample size.

[7] In this respect, we stress that combining a fully self-similar $M$–$T$ relation with a $L_X$–$T$ relation whose slope is not the one predicted by the self-similar model (only gravitational physics) implies that the $L_X$–$T$ and $M$–$L_X$ relations cannot simultaneously evolve in a self-similar way. The point seems to have been overlooked in the cluster detailed studies so far.

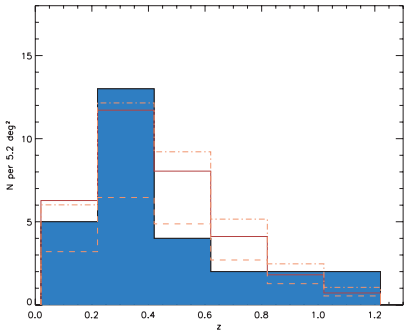


**Figure 13.** Impact of the steepening of the $M$–$T$ relation in the group regime on the modelling of the cluster number density. The full line corresponds to our fiducial model (broken power law, Section 5.1), giving 6.2 clusters per square degree. If we assume the Vikhlinin et al (2006) result (single slope of 1.5, dashed line), the cluster density drops to 3.6 deg$^{-2}$. The relation of Arnaud et al. (2005), motivated by a possible steepening of the mass function below 3.5 keV (single slope of 1.7, dash–dotted line), seems to be closer to our observations and yields a cluster density of 6.9 deg$^{-2}$.

*6.3.2 Impact of the scatter*

We discussed above how the scatter in the $L_X$–$T$ relation influences the determination of the evolution of the relation from the observations (Section 6.2.1). In our model, all sources of dispersion are encapsulated in the $M$–$L_X$ relation, since the determination and understanding of dispersion in $L_X$–$T$ and $M$–$T$, and thus $M$–$L_X$, and of their possible evolution, is still in its infancy. In this respect, we note that measurements of scatter other than those of Stanek et al. (2006) are available: lower values for $\sigma_{\ln L|M}$ have been found by Zhang et al. (2006) and Zhang et al. (2007) (respectively, 0.41 and 0.33), but they are dominated by higher temperature systems which are known to have more regular properties. In terms of the accuracy with which the scatter is determined, the typical errors estimated by Stanek et al. (2006) lead to an uncertainty in $\sigma_{\ln L|M}$ of the order of 20 per cent.

In Fig. 14, we show the impact on the C1 redshift distribution of neglecting the scatter in $M$–$L_X$, as well as the influence of plausible values of scatter. If the luminosities are symmetrically scattered around the expected value for a given mass, the net effect on our cluster counts (approximately the number of clusters detected above a given luminosity at a given redshift) is an increase in the detected cluster density since, due to the steepness of the mass function, substantially more low-mass clusters become detectable than high-mass cluster undetectable.

### 6.4 Cosmology and cluster survey self-calibration

We have demonstrated that taking into account survey selection effects is not only critical for the proper modelling of the cluster number counts (hence cosmology) but also for the determination of the cluster evolutionary scaling laws which, in turn, also impact on $dn/dz$ (Fig. 9). Further, we explored the impact on $dn/dz$ of the slope of the $M$–$T$ relation (Fig. 13) and of the dispersion in the scaling laws (Fig. 14). This illuminated, by means of real data, the magnitude of some of the degeneracies between cosmology and cluster evolution.

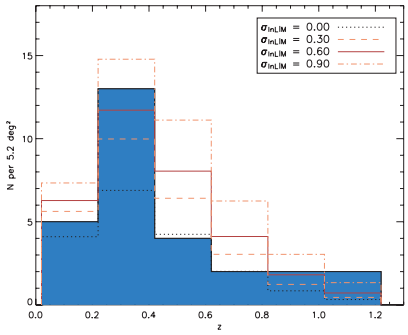


**Figure 14.** Impact of scatter in the $M$–$L_X$ relation on the predicted C1 redshift distribution. For given scatter values, the predicted total number of clusters per square degree are: 3.6 (no scatter), 5.1 (0.3) and 8.4 (0.9); our model gives 6.2 (0.6).

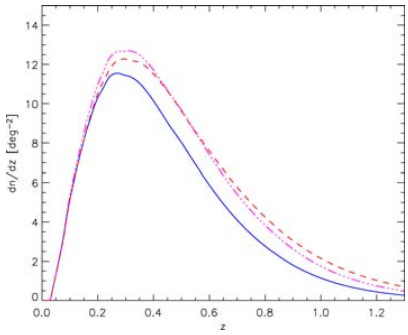


**Figure 15.** Dark energy versus scaling laws for our modelled cluster reshift distribution. The solid line shows the prediction for the Spergel et al. (2006) cosmology and self-similar evolution of the scaling laws. Assuming instead $F(z) = E(z) \times (1+z)^{0.34}$ for the $L_X$–$T$ evolution, which is our $1\sigma$ higher bound from Section 4.3, yields the dot–dashed line. As a comparison, the dashed line shows the distribution obtained by keeping self-similar evolution but changing the dark energy equation of state from $-1$ to $-0.6$.

However, the number of cosmological free parameters that we have explored in this study is very limited (basically, only $\sigma_8$, Fig. 9) and a more general analysis should, among others, include constraints on the dark energy parameter $w$. For instance, in Fig. 15 we exhibit the predictions of our model for different values of $w$ and $\alpha$ (evolution parameter of $L_X$–$T$). A clear degeneracy is apparent, adding to those already reviewed.

In order to cope with this critical issue, it has been suggested that X-ray [or Sunyaev–Zel'dovich (SZ)] cluster number counts, in principle, have the ability to *self-calibrate* even an evolving mass-observable relation, because the number density of clusters as a function of redshift has a fixed functional form given by cosmological simulations (Levine, Schulz & White 2002; Hu 2003; Majumdar & Mohr 2003). This suggests the possibility of simultaneously solving for the cosmology and fitting an ad hoc parametrized mass-observable relation, provided that a sample of several thousands of clusters is available.

While our – still somewhat phenomenological – best-fitting model, allowed us to derive flux–mass and count rate–mass relations, several caveats still have to be pointed out. We have noted that the flux–mass relation is very sensitive to the details of the X-ray-emitting plasma (e.g. emission lines) and also, obviously, to the way fluxes are estimated in shallow *XMM* exposures; in earlier studies, even the temperature dependence of the bremsstrahlung continuum was ignored (e.g. Majumdar & Mohr 2003). Even neglecting measurement uncertainties, our mass-observable mapping exhibits $1\sigma$ uncertainties in the range [$-60$ per cent, $+100$ per cent] coming solely from the dispersion of the $M$–$L_X$ relation.

In this respect, while vastly increasing the number of clusters may be useful, additional independent observations on cluster evolutionary physics will constitute a crucial input. This will improve the modelling of the mass-observable relations, and thus the constraints on the mass function. In particular, deeper *XMM* pointings will provide accurate temperatures and, consequently, a better estimate of the dispersion in the $L$–$T$ relation. Weak lensing, using the CFHTLS images, will provide independent mass measurements (Berge et al., in preparation).

In the future, the *XMM*-LSS field will be covered by SZ observations [One Centimeter Receiver Array (OCRA), Array for MIcrowave Background Anisotropy (AMiBA) and Atacama Pathfinder EXperiment (APEX)]. Since the SZ decrement is an integral of the cluster pressure ($n_e T_e$) along the line of sight, and is independent of $z$, it provides an especially interesting complement to the X-ray emissivity (scaling as $n_e^2 T_e^{1/2}$ for moderately hot clusters). Comparing our limiting mass as a function of redshift with those of these SZ surveys is all the more informative as we introduced a new class-oriented selection to define the *XMM*-LSS cluster sample. This is done, as a test case, in Fig. 16 for the *XMM*-LSS and the APEX–SZ surveys [Schwan et al. (2003), Dobbs et al. (2006)]. The SZ limit has been obtained under the assumption that APEX will observe at two frequencies (150 and 220 GHz) with a 1 arcmin beam and 10 $\mu$K beam$^{-1}$ white noise in each channel. We have included the effect of contamination by primary CMB anisotropies. Given the spatial spectrum of the CMB, the instrumental noise and cluster

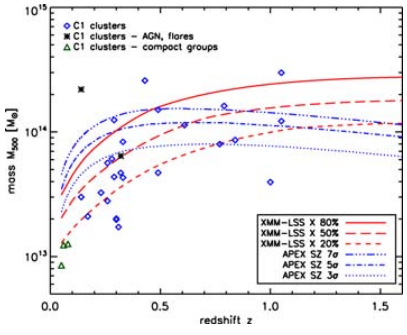


**Figure 16.** Comparison of the *XMM*-LSS and APEX–SZ sensitivity in terms of limiting mass. The red lines show various probability detection thresholds for the C1 clusters. The blue lines are the predictions for the 10-$\mu$K APEX survey (see the text). The X-ray and SZ curves make use of the best-fitting model to our current data set (Section 5.1).

profiles, we compute the expected SZ flux limit as a function of cluster size using matched filters. We then convert from SZ fluxes and sizes to masses and redshifts, using the same cosmology and scaling laws as in our X-ray modelling. Details of the computation are given in Melin, Bartlett & Delabrouille (2005).

The *XMM*-LSS mass limit of $10^{14}$ $M_{\odot}$ around $z = 1$ for the C1 selection, appears comparable to the prediction for the deepest SZ surveys to date. In addition, the *XMM*-LSS uncovers more low-mass systems below $z < 0.5$. According to our current model, the APEX–SZ cluster density detected at $3\sigma$ is 3.8 deg$^{-2}$ out to $z = 1$, to be compared with 5.4 deg$^{-2}$ for the observed C1 X-ray selection. In a forthcoming paper (Melin et al., in preparation), we explore the cosmological constraints expected from a few hundred clusters, whose masses are determined by a joint X-ray/SZ/weak-lensing analysis.

## 7 SUMMARY AND CONCLUSION

(i) This work presents a well-controlled *XMM* cluster sample over 5 deg$^2$ with a density of 5.4 deg$^{-2}$ at medium X-ray sensitivity (down to $\sim 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ for the extended sources in question). A complementary sample of about the same size, but with less well defined selection criteria is in preparation (Adami et al., in preparation). We provide positions, redshifts, fluxes, luminosities, temperatures and masses along with X-ray and optical images.

(ii) The selection is based on well-defined criteria pertaining to spatial extent properties, and are similar to a surface brightness (rather than flux) limit.

(iii) The resulting cluster redshift distribution extends out to $z = 1.2$ and peaks at $z \sim 0.3$. Half of the clusters have temperatures in the range 1–2 keV, occupying the $0.2 < z < 0.4$ interval. This intermediate population, the building blocks of the present-day clusters, is systematically unveiled by the *XMM*-LSS survey. It is also the first time that a wide-area blind X-ray survey has provided reliable cluster temperature measurements.

(iv) We demonstrate that taking into account the sample selection effects is vital for a proper determination of the evolution of the $L_X$–$T$ relation. This is due in large part to the steepness of the cluster mass function, which results in sources clustering close to the detection limit, and of these the overluminous systems are preferentially detected. This may explain the often contradictory results obtained with heterogeneous samples selected from *ROSAT* surveys.

(v) Modelling the selection effects carefully, our data appear to point to approximately self-similar evolution. A sample twice as large should allow exclusion of the no-evolution hypothesis, at better than the $1\sigma$ level. Using extensive simulations, we find that increasing the sample size is more efficient than increasing the accuracy of the temperature measurements for constraining the evolution; this is due to the large intrinsic dispersion of the $L_X$–$T$ relation.

(vi) Our results suggest a higher normalization of the cluster log($N$)–log($S$) at faint fluxes than previously obtained by deep *ROSAT* surveys.

(vii) Cluster masses were estimated from surface brightness profile fitting, assuming hydrostatic equilibrium. Our sample follows surprisingly well (given the very modest *XMM* exposures of $\sim 10$ ks and our rudimentary modelling of the mass profile) the local $L_X$–$M$ relation, when evolved self-similarly.

(viii) Self-consistent cosmological modelling of the cluster population, convolved with the accurately determined survey selection function, confirms that the properties of our current data set are compatible with the concordance cosmology along with cluster self-similar evolution.

(ix) This model then allowed us to investigate several degeneracies arising from cluster physics, regarding the cosmological interpretation of the number counts. In particular, we stress the influence of the, still poorly determined, scatter in the cluster scaling laws.

(x) This study led us to investigate many of the issues raised when attempting to perform precision cosmology using real cluster data, including the need for precise selection effects and a realistic error budget. Thanks to the temperature information obtained from the detected photons for all C1 clusters, this work constitutes one of the first attempts to break, in a self-consistent way, the degeneracy between cosmology and cluster evolution, in the analysis of the cluster number counts. The next step will be to apply cosmological modelling to the full *XMM*-LSS area (10 deg$^2$ – to be completed in 2007) adding input from the combined weak-lensing and SZ surveys as well as from the three-dimensional cluster distribution.

(xi) All data presented in this paper – cluster images taken at X-ray and optical wavebands in addition to detailed results for the spectral and spatial analyses – are available in electronic form at the *XMM*-LSS cluster online data base: http://l3sdb.in2p3.fr:8080/l3sdb/.

## ACKNOWLEDGMENTS

The authors would like to thank Monique Arnaud for sharing her PSF redistribution procedure. We are also grateful to the referee C. Collins for his constructive remarks on this manuscript. The results presented here are based on observations obtained with *XMM–Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA. The cluster optical images were obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at CFHT which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the CFHT Legacy Survey, a collaborative project of NRC and CNRS. AG acknowledges support from Centre National d'Etudes Spatiales. The Italian members of the team acknowledge financial contribution from contract ASI-INAF I/023/05/0. AD, OG, EG, P-GS and JS acknowledge support from the ESA PRODEX Programme 'XMM-LSS', and from the Belgian Federal Science Policy Office for their support. HQ acknowledges partial support from the FONDAP Centro de Astrofisica. We thank the IN2P3/DAPNIA computer centre in Lyon.

## REFERENCES

Andreon S., Willis J., Quintana H., Valtchanov I., Pierre M., Pacaud F., 2004, MNRAS, 353, 353
Arnaud M., 2005, in Melchiorri F., Rephaeli Y., eds, Proc. Enrico Fermi International School of Physics, Course CLIX. ESA Publications, Noordwijk
Arnaud M., Evrard A. E., 1999, MNRAS, 305, 631
Arnaud M., Majerowicz S., Lumb D. et al., 2002, A&A, 390, 27
Arnaud M., Pointecouteau E., Pratt G. W., 2005, A&A, 441, 893
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Barkhouse W. A. et al., 2006, ApJ, 645, 955
Böhringer H. et al., 2002, ApJ, 566, 93
Bremer M. et al., 2006, MNRAS, 371, 1427
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Carroll S. M., Press W. H., Turner E. L., 1992, ARA&A, 30, 499
Chiappetti L. et al., 2005, A&A, 439, 413

Dickey J. M., Lockman F. J., 1990, ARA&A, 28, 215
Dobbs M. et al., 2006, New Astron. Rev., 50, 960
Dupke R. A., Bregman J. N., 2006, ApJ, 639, 781
Ebeling H., Edge A. C., Allen S. W., Crawford C. S., Fabian A. C., Huchra J. P., 2000, MNRAS, 318, 333
Ettori S., 2000, MNRAS, 311, 313
Ettori S., Tozzi P., Borgani S., Rosati P., 2004, A&A, 417, 13
Finoguenov A., Reiprich T. H., Böhringer H., 2001, A&A, 368, 749
Finoguenov A. et al., 2007, ApJS, 172, 182
Ghizzardi S., 2001, EPIC-MCT-TN-011 (*XMM*-SOC-CAL-TN-0022), http://xmm.vilspa.esa.es/docs/documents/CAL-T-0022-1-0.ps.gz
Gioia I. M., Henry J. P., Maccacaro T., Morris S. L., Stocke J. T., Wolter A., 1990, ApJ, 356, L35
Grevesse N., Sauval A. J., 1998, Space Sci. Rev., 85, 161
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Jones C., Forman W., 1999, ApJ, 511, 65
Kenter A. et al., 2005, ApJS, 161, 9
Kitayama T., Suto Y., 1996, ApJ, 469, 480
Kolokotronis V., Georgakakis A., Basilakos S., Kitsionas S., Plionis M., Georgantopoulos I., Gaga T., 2006, MNRAS, 366, 163
Kotov O., Vikhlinin A., 2005, ApJ, 633, 781
Hu W., 2003, Phys. Rev. D, 67, 081304
Hu W., Kravtsov A. V., 2003, ApJ, 584, 702
Levine E. S., Schulz A. E., White M., 2002, ApJ, 577, 569
Lumb D. H. et al., 2004, A&A, 420, 853
Majumdar S., Mohr J., 2003, ApJ, 585, 603
Maughan B. J., Jones L. R., Ebeling H., Scharf C., 2006, MNRAS, 365, 509
Mazzotta P., Fusco-Femiano R., Vikhlinin A., 2002, ApJ, 569, L31
Melin J.-B., Bartlett J. G., Delabrouille J., 2005, A&A, 429, 417
Moretti A., Campana S., Lazzati D., Tagliaferri G., 2003, ApJ, 588, 696
Mullis C. R. et al., 2004, ApJ, 607, 175
Mullis C. R., Rosati P., Lamer G., Böhringer H., Schwope A., Schuecker P., Fassbender R., 2005, ApJ, 623, L85
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Novicki M. C., Sornig M., Henry J. P., 2002, AJ, 124, 2413
Osmond J. P. F., Ponman T. J., 2004, MNRAS, 350, 1511
Ota N., Mitsuda K., 2002, ApJ, 567, L23
Ota N., Mitsuda K., 2004, A&A, 428, 757
Pacaud F. et al., 2006, MNRAS, 372, 578
Pierre M. et al., 2004, JCAP, 9, 11
Pierre M. et al., 2005, MNRAS, 372, 591
Pierre M. et al., MNRAS, 382, 279
Pratt G., Arnaud M., 2002, A&A, 394, 375
Press W., Schechter P., 1974, ApJ, 187, 425
Read A. M., Ponman T. J., 2003, A&A, 409, 395
Refregier A., Valtchanov I., Pierre M., 2002, A&A, 390, 1
Reiprich T. H., Böhringer H., 2002, ApJ, 567, 716
Romer A. K., Viana P. T. P., Liddle A. R., Mann R. G., 2001, ApJ, 547, 594
Rosati P., della Ceca R., Norman C., Giacconi R., 1998, ApJ, 492, L21
Rosati P., Borgani S., Norman C., 2002, ARA&A, 40, 539
Sanderson A. J. R., Ponman T. J., 2003, MNRAS, 345, 1241
Schwan D. et al., 2003, New Astron. Rev., 47, 933
Spergel D. N. et al., 2003, ApJS, 148, 175
Spergel D. N. et al., 2007, ApJS, 170, 377
Stanek R., Evrard A. E., Böhringer H., Schuecker P., Nord B., 2006, ApJ, 648, 956
Stanford S. A. et al., 2006, ApJ, 646, L13
Starck J.-L., Pierre M., 1998, A&AS, 128, 397
Sugiyama N., 1995, ApJS, 100, 281
Tasse C. et al., 2006, A&A, 456, 791
Valtchanov I. et al., 2004, A&A, 423, 75
Vikhlinin A., McNamara B. R., Forman W., Jones C., Quintana H., Hornstrup A., 1998, ApJ, 502, 558
Vikhlinin A., VanSpeybroeck L., Markevitch M., Forman W. R., Grego L., 2002, ApJ, 578, L107
Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691
Voit G. M., 2005, Rev. Mod. Phys. 77, 207
Willis J. P. et al., 2005a, MNRAS, 363, 675
Willis J. P. et al., 2005b, MNRAS, 364, 751
Yuan F., Narayan R., 2004, ApJ, 612, 724
Zhang Y.-Y., Böhringer H., Finoguenov A., Ikebe Y., Matsushita K., Schuecker P., Guzzo L., Collins C. A., 2006, A&A, 456, 55
Zhang Y.-Y., Finoguenov A., Böhringer H., Kneib J.-P., Smith G. P., Czoske O., Soucail G., 2007, A&A, 467, 437

## APPENDIX A: NOTES ON SPATIAL FITTING AND PHOTOMETRIC ACCURACY

In this appendix, we test the accuracy of our luminosity-measurement procedure, running the spatial fitting algorithm on simulated images.

For this purpose, we created $10^4$-s *XMM* pointing simulations with vignetted and particle background levels set to the mean observed values of Read & Ponman (2003), and point sources distributed according to the log($N$)–log($S$) relation of Moretti et al. (2003). We then randomly included in these simulations extended sources with $\beta = 2/3$ and several $R_c$ and count-rate values (respectively, from 10 to 50 arcsec and 0.005 to 0.1 counts s$^{-1}$). A $\beta$-profile was fitted to all the simulated sources that were detected by the first pass of our pipeline (see Pacaud et al. 2006, for details) and for which the spatial fitting procedure detects at least three radial bins above $3\sigma$.

The simulations showed that below 400 detected counts in the spatial fitting radius, the degeneracy between $\beta$ and $R_C$ is large, but that the integrated flux (within the fitting radius) is well modelled by the best-fitting ($\beta$, $R_C$) combination. Results of the simulations are presented in Fig. A1.

The measurements are almost unbiased: less than a few per cent in general, and up to ∼10 per cent for very faint sources (100–200 counts). This small offset can be interpreted as a weak Eddington bias, as it is of the order of Poisson fluctuations, and probably results from our requirement of recovering at least three significant bins (which means retaining only the clusters that appear brighter

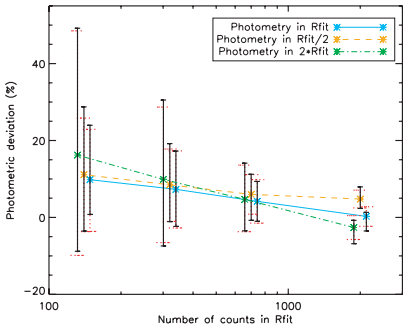


**Figure A1.** Photometric accuracy of our spatial fitting procedure using simulations. The data points show the mean deviation from the true value in our simulation as a function of available signal (number of counts within spectral fitting radius $R_{\rm fit}$), and the extrapolation factor (flux measurement within $R_{\rm fit}$, $R_{\rm fit}/2$ and $2R_{\rm fit}$). The solid error bars with short hats show the dispersion among measured values, while the dotted error bars with large hats give the mean value of the $1\sigma$ bounds estimated by the procedure.

among the faintest ones). This should not be an issue for the C1 clusters as they are generally brighter (see Table 1). The scatter around the mean value is very low for bright sources and mildly increases (up to ∼15 per cent) for the faintest ones. Estimating the luminosity within a different radius from the fitting one does not result in significant differences (although the scatter increases) as long as the profile is not extrapolated too far out, which is the case for our sample. Finally, the mean estimated errors are of the same order as the dispersion of the best-fitting value, assessing our confidence intervals.

Other sources of error have already been considered in the appendix A of Pierre et al. (2006): the impact of neglecting the errors on the temperature (which affects our estimate of $R_{500}$) was shown to be smaller than a few per cent, while the contribution from undetected weak AGN is probably lower than the per cent level (which is comforted by the present simulations in which AGN are included).

We thus conclude that the possible systematics generated with our fitting procedure are rather low, especially compared to the statistical errors.

## APPENDIX B: INDIVIDUAL CHARACTERISTICS OF THE C1 GALAXY CLUSTERS

Notes on individual sources.

(i) *XLSSC−028*. The velocity information obtained to date on this object does not allow us to firmly conclude on its redshift. The two bright galaxies on which the X-ray emission is centred have a redshift of 0.08; their spectra are typical of elliptical galaxies, without emission lines. No other galaxies with this redshift have been measured in the field, but a number of $z = 0.3$ objects are found within 500 kpc of the X-ray centroid. The measured X-ray temperature is 1.5 keV for $z = 0.3$, and the luminosity is found to be $1.6 \times 10^{43}$ erg s$^{-1}$, which puts the object close to the observed $L_X$–$T$ relation. At $z = 0.08$, the source is about a factor of 17 fainter, and corresponds to a temperature of 0.75 keV which put it also exactly on the $L_X$–$T$ relation.

(ii) *XLSSC−018*. The first measurement of the temperature of this system ($\sim 2.7^{+2.5}_{-0.9}$ keV, published in Willis et al. 2005a) was one of the hottest in our sample at a redshift around 0.3. Yet, the cluster does not seem massive, neither in the X-ray nor in the optical. The emission in the [2–10] keV appears to be significantly displaced from the low-energy centroid (Fig. B1), and is possibly associated with an optical counterpart. This suggests that a fraction of the emission could be due to an AGN. Excluding the probable contaminated region from the spectral fit lowers by a factor of 2 the number of available photons, but surprisingly leads to much tighter error bars around a temperature of 2.0 keV (as reported in Table 1). We regard this as strong evidence that the cluster's X-ray emission suffers AGN contamination, and consequently discard this source from the $L_X$–$T$ analysis.

(iii) *XLSSC−006*. This source was already studied in Willis et al. (2005a), and assigned a luminosity of $(4.5 \pm 0.3) \times 10^{44}$ erg s$^{-1}$ which is incompatible with our estimated $(6.0 \pm 0.2) \times 10^{44}$ erg s$^{-1}$. In that paper, the luminosity was derived from the XSPEC spectral fitting measurement, to which a correction factor was applied in order to extrapolate to $r_{500}$. We interpret the discrepancy as resulting from the combination of two effects. First, XSPEC uses a background estimate coming from an annulus surrounding the source, while in our spatial fitting procedures, we model it over the whole pointing. As a consequence, the mean background level used within XSPEC actually happened to be higher than our fitted value due to residual contamination of the bright XLSSC-006 cluster in this local annulus. Additionally, the observed source count rate in the [0.5–2] keV band (on which our spatial fitting luminosity estimate is based), is found to be higher within XSPEC than the one inferred from the best spectral fitting model.

Information on other C1 sample clusters is already published by Valtchanov et al. (2004), Willis et al. (2005a) and Pierre et al. (2006).

## APPENDIX C: THE C1 NEARBY GALAXIES

These sources were discarded from the C1 galaxy cluster sample based on an obvious link between the main X-ray emission and the

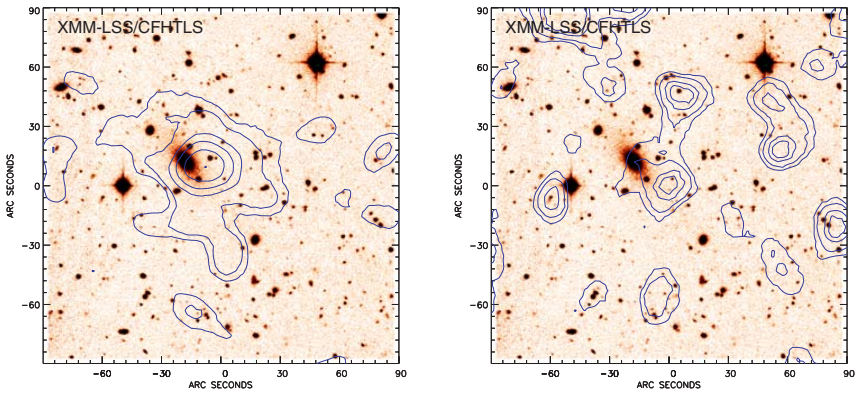


**Figure B1.** Gaussian-smoothed X-ray emission from the source XLSSC-018 at $z = 0.32$ in bands [0.5–2] keV and [2–10] keV. The emission in the soft band seems slightly offset from the cluster cD galaxy. In the hard band, the offset becomes much larger.

**Table C1.** The nearby galaxy sample. *XMM* pointing identifiers refer to the *XMM*-LSS internal naming convention as described in Pierre et al. (submitted); the location of each pointing on the sky is shown in Fig. 1.

| Source name | Pointing | RA (J2000) | Dec. (J2000) |
|---|---|---|---|
| XLSS J022528.7−040041 | B03 | 36.3699 | −4.0115 |
| XLSS J022251.4−031151 | B11 | 35.7146 | −3.1975 |
| XLSS J022659.2−043529 | G06 | 36.7469 | −4.5916 |
| XLSS J022430.4−043617 | G08 | 36.1268 | −4.6048 |
| XLSS J022617.6−050443 | G16 | 36.5735 | −5.0788 |

presence of a nearby galaxy on the same line of sight. A significant fraction of the total X-ray emission can, however, originate from another object in the field. The source list is given in Table C1, and details regarding the origin of the X-ray emission are given below:

(i) *XLSS J022528.7−040041*. The nearby galaxy lying at the centre of the X-ray isophotes is alreday known as APMUKS(BJ) B022258.83−041412.5. The X-ray emission from this galaxy certainly originates from its interaction with two satellite companions, as conspicuous in the images of Fig. C1 .

(ii) *XLSS J022251.4−031151*. This source is likely to result from the confusion of a point source with the nearby elliptical galaxy APMUKS(BJ) B022018.98−032531.1. It is associated with a radio source detected in Tasse et al. (2006).

(iii) *XLSS J022659.2−043529*. The nearby galaxy lying at the centre of the X-ray isophotes is already known as MCG -01-07-011. We cannot exclude it to be the dominant galaxy of a poor nearby group, as its aspect is reminiscent of the compact group XSLC-011 at $z = 0.05$.

(iv) *XLSS J022430.4−043617*. This source is likely to result from the confusion of a point source with a nearby spiral galaxy. The galaxy is already known as 6dF J0224300−043614, and has a redshift of $z = 0.069\,16$.

(v) *XLSS J022617.6−050443*. The nearby galaxy lying at the centre of the X-ray isophotes [APMUKS(BJ) B022347.26−051811.5] is detected as a weak radio source in Tasse et al. (2006). As for XLSS J022659.2−043529, we cannot exclude the possibility that it is the dominant galaxy of a small group. Another option would be that it is part of the recently identified XBONG class (see e.g. Yuan & Narayan 2004).

## SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article.

**Figure B2.** Images of the C1 clusters (sorted by right ascension). Left: X-ray/*Z*-band overlay of the central 7 arcmin. Right: true colour (*g*, *r*, *i*) image with X-ray contours and measured redshift in the central 1.5 Mpc.
**Figure C1.** Images of the C1 nearby galaxies. Left: 7 arcmin wide *L*-band image overlayed with X-ray contours. Right: 5 arcmin wide true colour (*g*, *r*, *i*) image overlayed with X-ray contours.

This material is available as part of the online paper from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-2966.2007.12468.x
(this link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.

## 7.6 La fonction de corrélation des AGNs :

## "The XMM-LSS survey : properties and two-point angular correlations of point-like sources"

### Contexte/motivation

Il s'agit de la première exploitation du catalogue complet d'AGN. Le XMM-LSS fournit le plus grand échantillon à cette profondeur sur une large surface contigüe. L'étude de leur distribution spatiale constitue donc une des contributions les plus originales du relevé à l'étude des sources ponctuelles (avec les analyses multi-longueurs d'onde).

### Résumé

L'article étudie les sources X détectées sur 4.2 $\deg^2$ pseudo-contigus dans les bandes 0.5-2 keV et 2-10 keV jusqu'à des flux resectifs de $2 \times 10^{-15}$ et $8 \times 10^{-15}$ erg.s$^{-1}$.cm$^{-2}$, dans le cadre du relevé XMM-LSS. La log(N)-log(S) dans chaque bande présente une forte pente aux flux élevés, mais est en bon accord avec d'autres détermination en deçà de $2 \times 10^{-14}$ erg.s$^{-1}$.cm$^{-2}$. Les sources détectées résolvent près de 30 pour cent du fond diffus X dans la bande 2-10 keV. Nous étudions la corrélation angulaire à deux points des sources ponctuelles en utilisant la statistique du plus proche voisin et la fonction de corrélation. Nous détectons un faible signal positif avec $\sim$1130 sources dans la bande 0.5-2 keV, mais pas de corrélation pour les $\sim$400 sources de la bande 2-10 keV en dessous d'une échelle de 100 arcsec. Un sous-échantillon de 200 sources avec des rapports de dureté élevés, probablement dominé par des AGNs obscurcis, affiche un signal positif compatible avec une grande longueur de corrélation, mais seulement à un niveau de $\sim 2\,(3)\sigma$, en se basant sur des méthodes de rééchantillonage statistique (poissonien). Nous discutons les implications probables et soulignons l'importance de relevés plus larges et complets en vue de mieux comprendre les structures à grande échelle du ciel en rayons X.

### Contribution personnelle

J'ai constitué et mis à disposition le catalogue utilisé. J'ai ensuite participé aux études visant à déterminer la meilleure manière de modéliser les effets de sélection.

A&A 457, 393–404 (2006)
DOI: 10.1051/0004-6361:20065284


Astronomy & Astrophysics

# The XMM large scale structure survey: properties and two-point angular correlations of point-like sources⋆

P. Gandhi[1,2], O. Garcet[3], L. Disseau[4,2], F. Pacaud[5], M. Pierre[5], A. Gueguen[5], D. Alloin[5,2], L. Chiappetti[6], E. Gosset[3], D. Maccagni[6], J. Surdej[3], and I. Valtchanov[7]

[1] Institute of Astronomy, Madingley Road, University of Cambridge, CB3 0HA, UK
e-mail: pg@ast.cam.ac.uk
[2] European Southern Observatory, Casilla 19001, Santiago, Chile
[3] Institut d'Astrophysique et de Géophysique, Université de Liège, Belgium
[4] École Normale Supérieure, Paris, France
[5] CEA/DSM/DAPNIA, Service d'Astrophysique, Saclay, 91191 Gif-sur-Yvette, France
[6] INAF IASF Milano, via Bassini 15, 20133 Milano, Italy
[7] Astrophysics group, Blackett Laboratory, Imperial College, London, UK

Received 27 March 2006 / Accepted 6 July 2006

**ABSTRACT**

We analyze X-ray sources detected over 4.2 pseudo-contiguous sq. deg. in the 0.5–2 keV and 2–10 keV bands down to fluxes of $2 \times 10^{-15}$ and $8 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ respectively, as part of the *XMM-Newton* Large Scale Structure Survey. The log $N$–log $S$ in both bands shows a steep slope at bright fluxes, but agrees well with other determinations below $\sim 2 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$. The detected sources resolve close to 30 per cent of the X-ray background in the 2–10 keV band. We study the two-point angular clustering of point sources using nearest neighbours and correlation function statistics and find a weak, positive signal for ∼1130 sources in the 0.5–2 keV band, but no correlation for ∼400 sources in the 2–10 keV band below scales of 100 arcsec. A sub-sample of ∼200 faint sources with hard X-ray count ratios, that is likely to be dominated by obscured AGN, does show a positive signal with the data allowing for a large angular correlation length, but only at the ∼2 (3)$\sigma$ level, based on re-sampling (Poisson) statistics. We discuss possible implications and emphasize the importance of wider, complete surveys in order to fully understand the large scale structure of the X-ray sky.

**Key words.** X-rays: diffuse background – large-scale structure of Universe – galaxies: Seyfert

## 1. Introduction

With the latest generation of X-ray satellites, *XMM-Newton* and *Chandra*, it has become possible to easily identify active galactic nuclei (AGN) and galaxy clusters, and map out their distribution to high redshifts (Brandt & Hasinger 2005; Rosati et al. 2002). We are finally in a position to answer questions such as: In what environments do AGN preferentially form? Are AGN formation and fuelling influenced by large-scale structure, or are their properties decided by factors local to the AGN and its hosting bulge alone? Is there a dependence of AGN obscuring column density on their larger-scale environment?

Several previous AGN spatial and angular clustering measurements have been carried out in X-rays and provide a mixed picture. The *Einstein* and *ROSAT* missions sampled bright and typically unobscured AGN populations, resulting in detections of moderate clustering signatures (e.g., Vikhlinin & Forman 1995; Carrera et al. 1998; Fabian & Barcons 1992, and references therein). In the optical, Wake et al. (2004) have found that Seyferts in the Sloan Digital Sky Survey at $z < 0.2$ selected on the basis of their [OIII] or [NII] emission line strengths are unbiased tracers of mass, with neither their auto-correlation properties, nor cross-correlation with galaxies showing significant excess above the field. How these results extend to high redshift and connect with AGN selected at other wavelengths is a subject of intense study.

More recent measurements between 2 and 10 keV are capable of probing through increasing columns of absorbing material associated with the tori of obscured AGN. Since obscured AGN outnumber their unobscured counterparts by a factor of anywhere between 3 and 10 (Maiolino & Rieke 1995; Matt et al. 2000) and since X-ray surveys select AGN much more efficiently than at other wavelengths (Brandt et al. 2004), hard-band studies (above ∼2 keV) are essential to draw conclusions from a representative AGN census. Such work has been carried out by Yang et al. (2003) with *Chandra* and Basilakos et al. (2004) with *XMM-Newton* over areas covering 0.4 deg$^2$ and 2 deg$^2$ respectively, and both find a significant auto-correlation signal, possibly associated with the distribution of obscured AGN.

In this paper, we describe initial results on the properties and distribution of X-ray-detected AGN in a large survey: the *XMM-Newton* Large Scale Structure survey[1] (hereafter XMM-LSS, Pierre et al. 2004). This is a contiguous, wide-area (currently ∼6 deg$^2$) survey with the primary goals of studying the physical

⋆ Based on observations obtained with *XMM-Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA.

[1] http://vela.astro.ulg.ac.be/themes/spatial/xmm/LSS/

properties of cluster/group populations; the impact of environment on star, AGN and galaxy formation; and, reciprocally, the effect of star formation activity on cluster properties. This is currently the widest, medium-deep survey of a contiguous patch of the X-ray sky with spatial resolution better than 10 arcsec above 2 keV. Full characterization of the nature of the detected X-ray population will be possible with extensive multi-wavelength follow-up currently under way. In addition, the XMM-LSS area and sub-regions have already been (or will be) observed as part of numerous large and "legacy" surveys at other wavelengths, including the radio (VLA, Cohen et al. 2003), optical (CFHT, VLT, Le Fèvre et al. 2004), near-infrared (UKIDSS[2]) and mid- to far-infrared wavelengths (Spitzer; Lonsdale et al. 2003). Initial follow-up of cluster candidates has proven highly successful. Results include the confirmation of several high-redshift clusters at $z > 0.6$ (Andreon et al. 2005; Valtchanov et al. 2004), extension of the lower-redshift ($0.3 < z < 0.6$) sample to the luminosity regime of poor groups and clusters (Willis et al. 2005a,b) as well as compilation of the highest sky density cluster sample to date (Pierre et al. 2006).

While the full XMM-LSS cluster survey is expected to provide sensitive measurements and consistency checks of cosmological parameters (Refregier et al. 2002), more than ∼80 per cent of the X-ray sources detected to the flux limits of the survey are point sources, predominantly AGN. We present the basic properties of the detected point-like sources in the XMM-LSS field, including distributions of source flux above specific sensitivity limits (Sect. 2.3) and a measurement of the resolved fraction of the X-ray background (Sect. 2.4). This is followed by an analysis of the projected two-point correlations on the sky of various sub-samples of point-like sources (Sect. 3), including a sample with hard count ratios most likely dominated by obscured AGN. Lastly, we compare these results with previous work and discuss possible implications (Sect. 4). While some source classification and separation is discussed in Sect. 2.2, detailed identification and follow-up (still in progress) will be presented in future works.

Analysis of these sources over a sub-region of guaranteed-time pointings (hereafter referred to as the "G pointings") covering ∼3sq degs has been described by Chiappetti et al. (2005). The present paper is an extension of these results to cover the point-sources detected in the full area of the XMM-LSS pointings observed so far, and to study their two-point angular correlations. We include an additional 30 pointings (hereafter referred to as the "B pointings") observed as part of guest observer time. Additionally, the source detection pipeline used for our work is different from that of Chiappetti et al. While these last authors use well-tested, standard XMM-SAS point-source detection algorithms (`eboxdetect, emldetect`, etc.), the main driver of the XMM-LSS survey to detect faint groups and clusters has motivated the development of a custom-built wavelet technique with a full profile fitting algorithm in order to best distinguish between point-like and extended sources (Pacaud et al. 2006), large parts of which we utilize in the present work.

Throughout this paper, the hard band refers to the X-ray 2–10 keV band; the soft band to 0.5–2 keV; and the term "hard-spectrum" sources refers to sources with a hardness ratio of X-ray counts (*HR*; the relative excess of hard band counts over the soft band) greater than –0.2. Where required, we use the concordance cosmology (Spergel et al. 2003), unless otherwise stated.

[2] `http://www.ukidss.org/`

**Table 1.** The list of *XMM-Newton* pointings (field label and Observation ID) considered for the present analysis. The nominal durations of the G (guaranteed time) and B (guest observer) pointing exposures are 20 and 10 ks respectively. Additional details at `http://cosmos.iasf-milano.inaf.it/~lssadmin/Website/LSS/Anc/`

| Field – ObsID | Field – ObsID | Field – ObsID |
|---|---|---|
| G01 – 0112680101 | G18 – 0111110401 | B15 – 0037981501 |
| G02 – 0112680201 | G19 – 0111110501 | B16 – 0037981601 |
| G03 – 0112680301 | B01 – 0037980101 | B18 – 0037981801 |
| G04 – 0109520101 | B02 – 0037980201 | B19 – 0037981901 |
| G05 – 0112680401 | B03 – 0037980301 | B20 – 0037982001 |
| G06 – 0112681301 | B04 – 0037980401 | B21 – 0037982101 |
| G07 – 0112681001 | B05 – 0037980501 | B22 – 0037982201 |
| G08 – 0112680501 | B06 – 0037980601 | B23 – 0037982301 |
| G09 – 0109520601 | B07 – 0037980701 | B24 – 0037982401 |
| G10 – 0109520201 | B08 – 0037980801 | B25 – 0037982501 |
| G11 – 0109520301 | B09 – 0037980901 | B26 – 0037982601 |
| G13 – 0109520501 | B10 – 0037981001 | B27 – 0037982701 |
| G14 – 0112680801 | B11 – 0037981101 | B28 – 0147110101 |
| G15 – 0111110101 | B12 – 0037981201 | B29 – 0147110201 |
| G16 – 0111110701 | B13 – 0037981301 | B30 – 0147111301 |
| G17 – 0111110301 | B14 – 0037981401 | B31 – 0147111401 |

## 2. The sample of point-like sources in the XMM-LSS

The XMM-LSS observations consist of 19 guaranteed-time (G) and 32 guest-observer time (B) overlapping pointings covering a total area of ∼6 deg$^2$. The nominal exposure times are 20 ks and 10 ks for the G and B pointings, respectively. One G pointing and two B pointings[3] with high flaring background were not analysed for this work; the remaining 48 pointings (18 G + 30 B) are listed in Table 1, and the layout of the fields on the sky is shown in Fig. 1. Details of the X-ray observations are described in Pierre et al. (2004), and complete details of the detection pipeline and source classification will be presented in Pacaud et al. (2006). The unique feature of the pipeline is a custom-built maximum-likelihood profile fitting algorithm (Xamin) that runs on a list of initial detections found by first using SExtractor (Bertin & Arnouts 1996) on wavelet-filtered, reduced images in each pointing and energy band. This increases sensitivity towards faint, extended sources while properly accounting for Point-Spread-Function (PSF) variation with all instrumental and position–dependent effects (e.g., energy, off-axis position, bad pixels and CCD gaps). Herein, we present results obtained from pipeline runs on the 0.5–2 keV (soft) and 2–10 keV (hard) band photon images, and refer the reader to Pacaud et al. (2006) for the full XMM-LSS pipeline algorithm.

### 2.1. Source detection and photometry

For each source, the maximum-likelihood normalization of the PSF profile over the local background determines the photometry in counts, after appropriately masking out any neighbouring sources by the use of segmentation maps and accounting for chip gaps. The fit is carried out over an aperture large enough to encompass the bulk of the counts (typically a 70 arcsec box for point-sources, and a larger aperture for extended ones, depending on source extent). Conversion from count-rate (*CR*) to flux (*F*) is computed from a combination of the inverse

[3] These pointings have the following *XMM-Newton* Observation IDs respectively: 0109520401, 0037981701 and 0147111501.

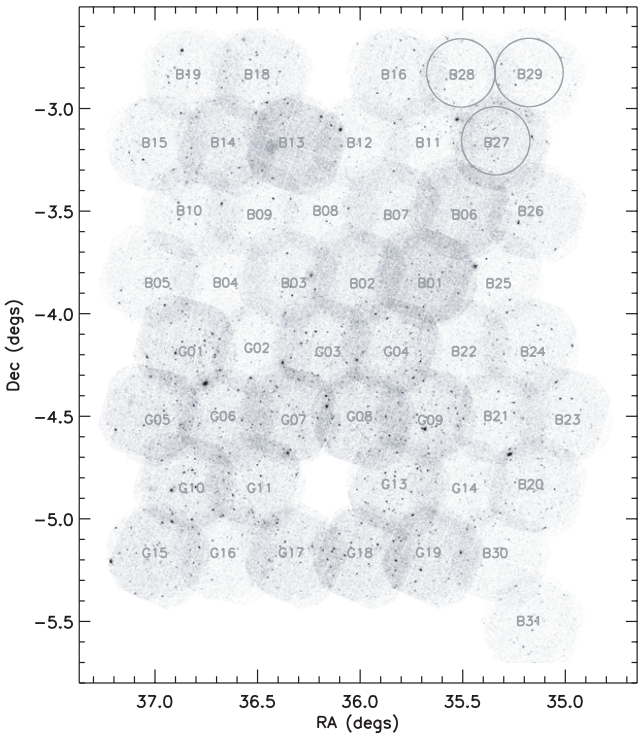


**Fig. 1.** The layout of the 48 pointings of the XMM-LSS survey. Smoothed 0.5–10 keV photon images for all cameras have been coadded with the same scaling in counts for display (without exposure map correction). Each pointing has a full field-of-view diameter of just under 30 arcmin. The placement is such that most adjacent pointings overlap beyond ∼10 arcmin from the respective optical axis centres (the pointings are labelled with their field names (compare with Table 1) at approximately these centres). In the present work, we restricted the analysis to the respective central 10′ regions, outlined for three pointings only (for clarity) at the top right as the 10′-radius large circles. The limiting sensitivity varies as a function of the exposure times, background level and off-axis angles. Many of the brighter X-ray sources are easily discerned in black. The longer-exposed guaranteed time pointings (G) are in the southern part of the survey; the rest (B pointings) being obtained during guest observer time.

**Table 2.** The count-rate-to-flux conversion factors for the individual EPIC cameras and energy bands, stated in units of erg s$^{-1}$ cm$^{-2}$ for a rate of 1 ct s$^{-1}$. A photon-index power-law of $\Gamma = 1.7$ affected by Galactic absorption of $2.6 \times 10^{20}$ cm$^{-2}$ was assumed. Both MOS cameras were assumed to be identical.

| EPIC Camera | 0.5–2 keV | 2–10 keV |
|---|---|---|
| MOS | $4.990 \times 10^{-12}$ | $2.296 \times 10^{-11}$ |
| pn | $1.460 \times 10^{-12}$ | $7.912 \times 10^{-12}$ |

conversion factors ($CF$) for each of the cameras, scaled by the exposure times as in Baldi et al. (2002). The conversion factors for individual cameras were calculated using XSPEC (Arnaud 1996) and the latest available EPIC response matrices. The thin filter response was considered, as for the actual observations. The model used was an intrinsic power-law $\Gamma = 1.7$ affected by Galactic absorption of $N_{\rm H} = 2.6 \times 10^{20}$ cm$^{-2}$, appropriate to the XMM-LSS sight-line. The individual $CF$s are listed in Table 2.

### 2.2. Source selection and classification

Intensive follow-up of *Chandra* and *XMM-Newton* surveys has shown that the vast majority of X-ray detections at high galactic latitudes are associated with AGN, in deep as well as in shallow surveys (Brandt et al. 2004; Szokoly et al. 2004; Nandra et al. 2004). The optical identifications of the associated counterparts are varied, and include Seyferts as well as late and early-type galaxies. Despite the fact that these do not always show obvious optical signs of AGN activity, nevertheless the presence of a powerful accretion source is indisputable, based on the inferred power of the X-ray source, and, in many cases, the detection of weak, high-ionization emission lines. In the XMM-LSS survey as well, the bulk (more than ∼80 per cent) of X-ray detections will be due to AGN (compare with, e.g., Crawford et al. 2002). Stars and bremsstrahlung emission from galaxy clusters will constitute the remaining sample.

Obvious stars were removed by cross-correlating our X-ray detections with the USNO survey (Zacharias et al. 2004) and identifying all optical point-like sources with a $B$-band

magnitude brighter than 13. This cross-correlation resulted in 17 sources being removed from the 0.5–2 keV sample, and these are not included in any of the following analysis. No stars coincided with a hard-band X-ray source above the nominal significance threshold (see below). Though there will undoubtedly remain some contamination due to fainter non-AGN point-like sources, the *ROSAT* deep surveys have found that stars represent only a minority of all X-ray detections (Schmidt et al. 1998) and typically display soft X-ray spectral slopes. Thus, our results should not be affected significantly by these contaminants, especially in the hard-band.

In the process of identifying clusters and groups (hereafter, jointly referred to as clusters), extensive simulations are being carried out to determine the best X-ray classification parameters, based on already published clusters (Valtchanov et al. 2004; Willis et al. 2005a,b, amongst others). For the purpose of the present work, we take as "extended", only those soft-band (0.5–2 keV) sources that have already been followed-up at other wavelengths and confirmed to be clusters, or found to have a high probability of being associated with such systems, based on i) optical-spectroscopy; ii) X-ray source profile extent; and, iii) the presence of a red sequence of galaxies (see Pacaud et al. 2006, for full details). At the time of writing, this consists of 28 confirmed clusters, and 28 clusters with provisional spectroscopic redshifts over the full 5.7 deg$^2$ of Fig. 1. The rest of the sources in the soft band are classified as "point-like". In the 2–10 keV (hard) band, we treat all sources as being point-like.

For the analysis presented in this paper, only those sources that lay within 10 arcmin of the optical axis centres of each pointing were retained. This was done in order to minimize biases due to PSF distortion at large off-axis angles and possible confusion due to the same source being detected on adjacent pointings. This results in pseudo-contiguous coverage of the field, with holes in between neighbouring pointings (that are fully accounted for in our analysis; see below).

Finally, the significance of source detection was estimated by the value of signal:noise ($S/N$). The "signal" used is that corresponding to 68 per cent of the full background-subtracted counts for each source, and the "noise" assumes Poisson errors on the total counts (source+background) within a fixed aperture of radius of 17 arcsec. Strictly speaking, this is consistent only for on-axis point-sources[4], though the effect of the off-axis PSF degradation is mitigated somewhat by our restriction to the central 10 arcmin regions of each pointing. To compute the Poisson error, we specifically use Eq. (7) of Gehrels (1986) for the $1\sigma$ upper-limit on the noise. Most results in the following sections are presented for the sample with $S/N > 3$, but we also consider an effectively-fainter sample with $S/N > 2$ in Sect. 3.4, and briefly mention results relevant to sub-samples with higher $S/N$ thresholds of 4 and 5.

### 2.3. Results: sky coverage and log N–log S

Sensitivity maps for each pointing were constructed by using the background measured by the source detection pipeline. The minimum number of counts necessary at any spatial pixel on a pointing (above the local background) in order for a source to have a S/N matching the adopted threshold is computed, and converted to a limiting-flux using the exposure maps and conversion-factors. The sky coverage is computed by summing the area covered by these sensitivity maps as a function of limiting-flux, and is shown for the hard and soft bands in Fig. 2, for the nominal S/N threshold of 3. The longer-exposed guaranteed time (G) pointings have a deeper sensitivity, but smaller overall coverage than the guest observer (B) pointings. The maximum coverage at the brightest fluxes is close to 4.2 deg$^2$; overlap between adjacent pointings in this selected area of the central 10 arcmin regions is negligible (only 0.3 per cent of the total).

[4] http://xmm.vilspa.esa.es/external/xmm_user_support/document-ation/uhb/node17.html

**Table 3.** Sizes of various sub-samples of X-ray sources analyzed in this paper. For the S/N threshold of 2, only hard-band sources are analyzed in Sect. 3.4. The parentheses in the first column specify whether the sample was selected over the whole area (B+G pointings), or only over the deeper G pointings.

| Selection criterion | Classification | 0.5–2 keV | 2–10 keV |
|---|---|---|---|
| $S/N > 3$ (B+G) | Point sources | 1134 | 413 |
| " | Extended sources | 36 | 0[a] |
| " | Stars | 17[b] | 0 |
| $S/N > 2$ (B+G) | Point sources | – | 912 |
| $S/N > 2$ (G) | Point sources | – | 473 |
| " | $1 \geq HR > -0.2$ | – | 209 |
| " | $1 > HR > -0.2$ | – | 140 |

Notes: [a] All sources were considered as point-like in the 2–10 keV band. If we relax this condition, only 8 sources are affected. [b] Only obvious stars with $B < 13$ were identified (see text), and these are counted separately from "point sources".

Above a $S/N$ threshold of 3, we find a total of 1134 and 413 point sources in the soft and hard bands respectively (Table 3 lists the numbers of sources in various bands and with different selection criteria as discussed in the text). The $\log N$–$\log S$ in each band (also shown in Fig. 2) was computed by identifying all sources with a detected flux above any given value $S$ and summing the inverse areas over which these sources could have been detected, as measured from the sky coverage. The flux distributions of the detected sources peak at $\sim 5 \times 10^{-15}$ and $\sim 1.5 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ in the soft and hard bands respectively, and at these flux levels, the $\log N$–$\log S$ is in excellent agreement with other published surveys (e.g., we show the fit of Baldi et al. 2002 in Fig. 2). The faintest fluxes detected (typically near the optical axes) are $\sim 2$ and $8 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ in the two bands, respectively. At the lowest fluxes in the soft band, our $\log N$–$\log S$ begins to drop off due to incompleteness, while at bright fluxes in both bands, our $\log N$–$\log S$ is lower than (but consistent with the $1\sigma$ lower limit of) other surveys. This difference was also observed by Chiappetti et al. (2005) over the area of the G pointings, based on a completely different source detection procedure (their $\log N$–$\log S$ is also shown in Fig. 2). We find excellent agreement with Chiappetti et al. (2005, especially in the soft band) and refer to their paper for power-law fits and discussion of the $\log N$–$\log S$. The agreement gives us confidence that the noted deficit is intrinsic and not due to pipeline systematics.

### 2.4. Results: resolved fraction of the X-ray background

Only with the new generation of X-ray satellites has the hard cosmic X-ray background radiation been resolved substantially into discrete sources (a combination of obscured and unobscured AGN), supporting the basic tenet of AGN Unification (Setti & Woltjer 1989; Mushotzky et al. 2000). While the soft X-ray background below 2 keV can be accounted for almost completely as a combination of emission from Galactic emission, AGN and clusters (e.g., Fabian & Barcons 1992), the exact fraction resolved out in the hard band remains a contentious issue.

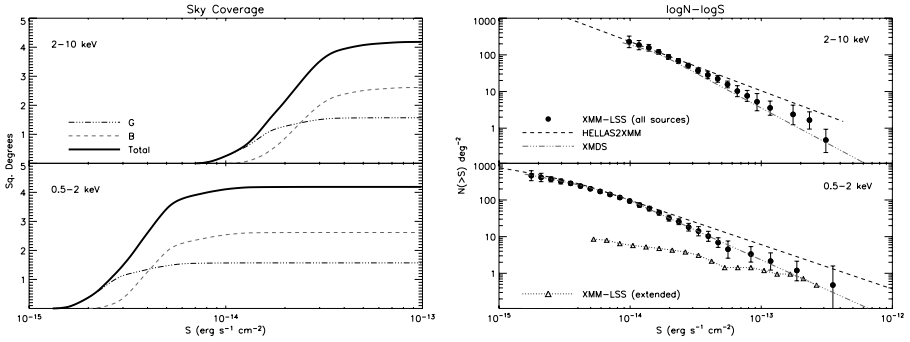


**Fig. 2.** The sky coverage (*left*) and $\log N$–$\log S$ (*right*) of the XMM-LSS sample within the central 10 arcmin-radius pointing regions, for the 2–10 keV (*top*) and 0.5–2 keV (*bottom*) bands, for a threshold $S/N > 3$ in both bands. The sky coverage is shown separately for the guaranteed time (deeper; dots-dashed; marked "G") and guest observer time (shallower; dashed; marked "B") pointings. The $\log N$–$\log S$ is shown for all sources (clusters have a minor contribution, except at bright fluxes in the soft band; their contribution is shown as the triangles; marked "extended"). The errors on the $\log N$–$\log S$ denote $1\sigma$ Poisson uncertainties on the independent differential count bins, subsequently scaled to the integral counts. The results of the HELLAS2XMM survey (Baldi et al. 2002) and *XMM-Newton* Medium Deep Survey (XMDS Chiappetti et al. 2005) are also overplotted for comparison.

There is even intriguing evidence that a new population of obscured AGN remains undiscovered in even the deepest surveys (De Luca & Molendi 2004; Worsley et al. 2004).

We compute the resolved intensity of detected sources by summing over the flux of each source dividing by the inverse area over which the source would have been detected, similarly to the computation of the $\log N$–$\log S$. The result is shown in Fig. 3 and tabulated in Table 4. At the nominal flux limit of the XMM-LSS survey (with $S/N > 3$), we resolve close to 30 per cent of some latest measurements of the 2–10 X-ray background (De Luca & Molendi 2004; Hickox & Markevitch 2006). This matches well with the level resolved in other surveys over our flux regime (e.g., Manners et al. 2003). Note that hard-spectrum sources (discussed further in Sect. 3.4) contribute a larger fraction at fainter fluxes due to their increasing dominance, relative to those with softer X-ray spectra.

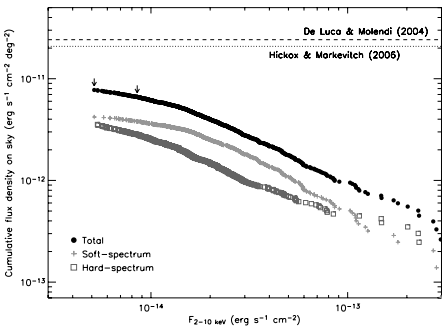


**Fig. 3.** Cumulative X-ray background intensity in the 2–10 keV band for the detected point sources in the full field of the XMM-LSS survey. Arrows denote the faintest flux limits of sub-samples compiled with S/N thresholds of 3 (*right arrow*) and 2 (*left arrow*), respectively. The dashed line is the total X-ray background measurement reported by De Luca & Molendi (2004); the dotted line is the measurement of Hickox & Markevitch (2006) converted to 2–10 keV assuming a power-law with photon index $\Gamma = 1.4$. The resolved source contribution is sub-divided into soft-spectrum and hard-spectrum sources (discussed in Sect. 3.4).

## 3. Angular correlation statistics

Full clustering analysis requires *spectroscopic* redshifts, since even optimistic photometric redshift errors of $\Delta z \sim 0.1$ translate into typical physical separations of hundreds of Mpc, washing out any intrinsic clustering signal. Until the on-going multi-wavelength follow-up of X-ray sources produces a useful number of accurate redshifts, we restrict our study to the areal distribution of sources only.

Angular clustering is related to the excess probability of finding source pairs at any given angular separation $\theta$, relative to a sample distributed with uniform probability. A variety of methods have been used for studying clustering properties of astronomical surveys, including power spectrum (Webster 1976), counts-in-cells (e.g., Carrera et al. 1998) and fractal analyses (Joyce et al. 1999), each of which can be implemented either in real, or in projected space. Here, we describe results obtained from nearest-neighbours and correlation function statistics for the point-like sources in the XMM-LSS.

### 3.1. Generation of random (uncorrelated) catalogues

The most straightforward way to account for instrument and pipeline selection effects in the survey (e.g., off-axis vignetting, holes between adjacent pointings, chip gaps) is to simulate an ensemble of catalogues over randomly-chosen sky positions that also correctly accounts for these effects. Source fluxes were randomized to have an overall distribution similar to the data. A random flux ($S_r$) can be sampled from the data $\log N$–$\log S$ by choosing a random number ($p$) uniformly distributed between 0

**Table 4.** Resolved fraction of the X-ray background over the B+G pointings for sub-samples with different S/N thresholds. The resolved fraction is stated as a percentage of the total intensity of $2.4 \times 10^{-11}$ erg s$^{-1}$ cm$^{-2}$ deg$^{-2}$ found by De Luca & Molendi (2004). The resolved fraction is further sub-divided into the contribution of hard-spectrum and soft-spectrum sources (see Sect. 3.4).

| $S/N$ threshold | Faintest flux | Median flux | XRB sky intensity | Resolved fraction (per cent) | | |
|---|---|---|---|---|---|---|
| | erg s$^{-1}$ cm$^{-2}$ | erg s$^{-1}$ cm$^{-2}$ | erg s$^{-1}$ cm$^{-2}$ deg$^{-2}$ | All sources | Hard-spectrum sources | Soft-spectrum sources |
| $S/N > 3$ | $8.5 \times 10^{-15}$ | $2.6 \times 10^{-14}$ | $6.6 \times 10^{-12}$ | 28 | 12 | 16 |
| $S/N > 2$ | $5.1 \times 10^{-15}$ | $1.8 \times 10^{-14}$ | $7.7 \times 10^{-12}$ | 33 | 16 | 17 |

and 1, and searching for the flux $S_r$ that satisfies the following transformation:

$$N(> S_r) = N(> S_{lim}) \times [1 - p] \quad (1)$$

where $N(>S_r)$ is the value of the $\log N$–$\log S$ at $S_r$ and $S_{lim}$ is the limiting flux above which the $\log N$–$\log S$ is defined (see also Mullis et al. 2004). Sky positions for these random catalogues were assigned with equal probability within the central 10′ regions of the pointings. If the limiting-flux at the assigned coordinates was larger (worse) than the randomly-chosen flux for the source, the source was discarded, and another one randomized in its place. We verified that the $\log N$–$\log S$ distributions of the random catalogues were similar to those of the parent catalogues used to generate them in each band and with each selection criterion.

This procedure assumes that a source with a randomly generated flux $S_r$ would have been detected with exactly that flux at the assigned sky position. In reality, any detection procedure will introduce shot noise, which results in some faint sources below the nominal threshold still being detected due to Poisson uncertainties "boosting" their counts. On the other hand, some bright sources may have their counts "depressed" and thus lie below the detection threshold. We simulated this by converting the flux $S_r$ to counts (with the use of local exposure maps and conversion-factors) and drawing a random number from the Poisson distribution of the total counts (source+local background). The source is then retained in the final random catalogue with a probability that the Poisson distribution exceeds the total limiting counts required by the S/N threshold. This simulation of shot noise was used for all random catalogues in each of the bands and sub-samples tested.

The above manual introduction of shot noise was necessary because full simulations of the entire dataset proved to be unrealistic. Still, as a cross-check, we did carry out full simulations of three pointing exposures by randomly generating sources with appropriate off-axis and energy-dependent PSFs for each camera, along with simulated backgrounds in accordance with Read & Ponman (2003), all scaled for the respective pointing exposure times. Our full detection pipeline was then run over these. We verified that the statistical distribution of sources over these simulated pointings compares well with the random catalogues above.

With an *XMM-Newton* Mirror Module PSF characteristic size of ∼6 arcsec, there is a small, but finite probability that blending of two close, neighbouring sources could result in them being classified as a single, extended source. The XMM-LSS pipeline has been designed to optimally detect and probe the evolution of extended sources, and we thus make use of its strength of separating point-like sources from extended clusters. We studied the effect of blending through several sets of simulations and confirmed that our pipeline is able to resolve close-by pairs successfully, though the efficiency of doing so depends on source flux as well as on pair separation. While no sources will be resolved at separations of less than ∼6 arcsec, the pipeline resolves effectively all source pairs at separations of ≥30 arcsec. Based on the simulations and observed source counts, the efficiency of pair resolution at a separation of ∼20 (10) arcsec was determined to be ∼70 (30) per cent. Further details will be presented in Pacaud et al. (2006), but for our purposes, we simply remove a corresponding fraction of such close pairs from the random catalogues in order to simulate this blending.

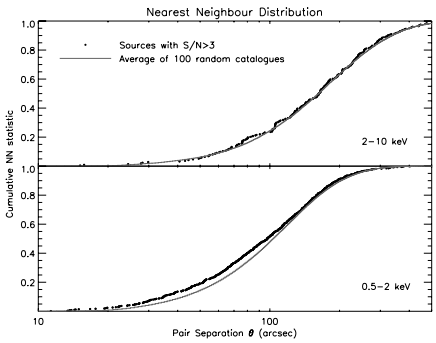


**Fig. 4.** Cumulative nearest-neighbour distribution function (statistic) for the soft (*bottom*) and hard (*top*) bands for the point-sources with $S/N > 3$ (filled circles) compared to the average statistic of 100 random catalogues. The plot shows the fraction of sources with an angular separation less than any given $\theta$ to their respective nearest neighbours. The soft band distribution has an excess of pairs compared to random, as opposed to the hard band sample.

A hundred simulated catalogues are generated for each band and sub-sample studied, each with the same number of sources as the parent data catalogue. The average number of data and random pairs in the ensemble are then counted for the computation of correlation statistics.

### 3.2. Results: angular nearest neighbours statistic

First, we show the distribution of projected separations of point-like sources from their first nearest neighbours (NN). This is shown as the cumulative (normalized) distribution in Fig. 4, compared to the average distribution of nearest neighbours for 100 random catalogues, in both the soft and the hard bands. An excess of nearest neighbours is observed in the soft band (*bottom plot*) below ∼100 arcsec. A Kolmogorov-Smirnov (K-S) test returns a small probability of $\approx 10^{-3}$ for the null hypothesis that the two distributions are identical, implying possible clustering in this band. In the hard band, no such excess is observed, and the K-S probability is consistent with the data and random distributions being drawn from the same population.

### 3.3. Results: angular correlation function

Optimal estimators are widely used to quantify the overall excess of data-data pairs over random-random ones at different scales: we chose to use the Hamilton estimator (Hamilton 1993) for the angular correlation function (ACF), but found very similar results using others, e.g., that of Efstathiou et al. (1991). Excess clustering compared to a uniform distribution is parametrized in terms of $\omega(\theta)$, defined as

$$\omega(\theta) = f\frac{DD(\theta)RR(\theta)}{DR(\theta)DR(\theta)} - 1 \quad (2)$$

where $DD$, $RR$ and $DR$ are the number of data-data, random-random and data-random pairs at separation $\theta$, all subjected to the survey selection effects. The normalizing factor $f$ is $4N_{\rm D}N_{\rm R}/(N_{\rm D}-1)(N_{\rm R}-1)$, where $N_{\rm D}$ and $N_{\rm R}$ are the number of sources in the data and random catalogues, respectively. Source pairs were binned in equal logarithmic intervals of $\theta$: the bin sizes being chosen to include at least ∼20 pairs in each bin in order to minimize the effect of small-number statistics. This restriction also defined the minimum pair-separation bin over which $\omega$ is plotted and fitted: this typically lies between 20–50 arcsec. In any case, the overall results presented below are not sensitive to reasonable binning choices.

We show the ACF results for point-like sources in Fig. 5. The plotted error-bars are Poisson 1-$\sigma$ uncertainties, calculated for each bin as $(1+\omega)/\sqrt{DD}$. In the soft band, there is an overall positive auto-correlation signal at most scales smaller than ∼1000 arcsec. The amplitude of this correlation is small: if we characterize $\omega$ as a power-law of the form

$$\omega(\theta) = (\theta_0/\theta)^{\gamma-1} \quad (3)$$

we find $\theta_0 = 6.3 \pm 3$ arcsec, with a slope of $\gamma = 2.2 \pm 0.2$, with the quoted errors being appropriate to 68 per cent confidence intervals for one interesting parameter. By simply counting the excess number of data-data pairs compared to random-random ones with all separations less than, say, 200 arcsec, we find an excess at the 2.3$\sigma$ level. This result was measured for the sample of 1134 point-like sources alone. The auto-correlation signal is almost identical for the combined catalogue of 1170 soft-band detections, including the 36 extended sources, suggesting that there is no obvious and strong cross-correlation signal between the extended and point-like samples; however, this will be studied in detail once a larger, complete sample of extended sources is compiled.

In terms of numbers of pairs, we detect 58 (251) independent data pairs with separations of <100 (200) arcsec, while the random catalogues contain 58 (230) pairs over the same scales, on average. This result is consistent with the null hypothesis implied by the nearest neighbour distribution. Though statistics are small, we find similar null results for sub-samples of hard-band sources selected with $S/N > 4$ (209 sources) or with $S/N > 5$ (123 sources).

We note that bias (underestimation of $\omega$) related to the integral constraint (the fact that a finite sky area with an unknown source density is used to estimate the correlation signal) is negligible for our field. Assuming the power-law form of Eq. (3) for the intrinsic correlation function, the bias can be estimated by numerical integration of $\theta_0^{\gamma-1}/\Omega^2 \int\int \theta^{1-\gamma} {\rm d}\Omega_1 {\rm d}\Omega_2$ over the whole area ($\Omega$) of the survey. For a range of relevant power-law parameters, we find this bias to be ∼0.01, a quantity sufficiently small to ignore during fitting.

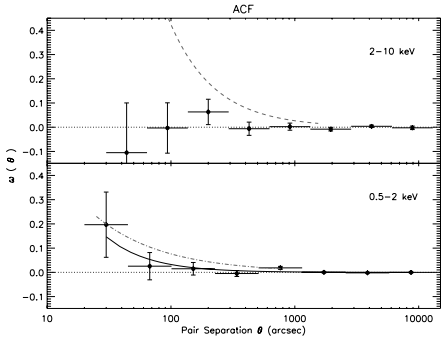


**Fig. 5.** The ACF, as defined by Hamilton (1993) and measured for the XMM-LSS survey in the soft (*bottom*) and hard (*top*) bands for the samples with $S/N > 3$. The solid curve is the best-fit power-law model (shown for the soft band only), while the dotted line marks $\omega = 0$. The y-axes are plotted on a linear-scale to aid visualization of the significance of any correlation and the axes ranges are kept the same in both bands for comparison. Previous power-law ACFs of Basilakos et al. (2004, for the hard band) and of Vikhlinin & Forman (1995, for the soft band) are shown as the dashed and dot-dashed lines respectively.

### 3.4. Results: sources with hard X-ray spectral indices

The above result on the non-detection of clustering in the 2–10 keV band seems to be at odds with the findings of Yang et al. (2003) and Basilakos et al. (2004). In both these works, excess clustering was detected in the hard band, and associated with the clustering of obscured AGN, which should constitute a larger fraction of the hard-band detections as compared to the soft band. We have made predictions for the fraction of AGN detected in the XMM-LSS survey that are expected to be obscured (see Sect. 4.2 below) and find that, at the flux limits probed by our sample, obscured AGN will represent approximately 30–40 per cent of the hard band detections, and about 10 per cent of the soft band sample; this fraction increases with decreasing flux levels. Obscured AGN can be efficiently (but not uniquely) selected by computing the hardness ratio ($HR$) of source counts in the hard ($H$) and the soft ($S$) bands. If we define

$$HR = \frac{H-S}{H+S}, \quad (4)$$

then a large fraction of sources with $HR > -0.2$ (hereafter, "hard-spectrum sources") are likely to be obscured AGN. This limit corresponds to an obscuring column density of approximately $10^{22}$ cm$^{-2}$ due to cold gas local to a source with an intrinsic power-law photon index of 1.7 at a redshift $z = 0.7$, and has been often used in the literature to separate intrinsically obscured AGN from unobscured ones (e.g., Gandhi et al. 2004; Padovani et al. 2004).

Applying the above $HR$ criterion to the hard band sample of 413 sources results in only 133 hard-spectrum sources over the B+G pointings, which is not sufficient for a proper correlation analysis. Since obscured AGN begin to emerge at the faintest fluxes, it is possible to probe them in larger numbers by decreasing our significance threshold for source detection. We therefore searched for detections with $S/N > 2$, and found a total of 912 sources in the 2–10 keV band, of which 409 have

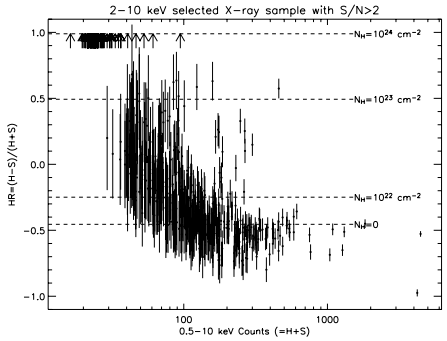


**Fig. 6.** Hardness ratio (*HR*) vs counts (Hard+Soft; i.e. $H + S$) of the G pointings sample of sources with $S/N > 2$ in the 2–10 keV band with either a unique counterpart (400 black dots with error bars) or no counterpart (69 arrows) in the soft band. The horizontal dashed lines mark the hardness ratios of a power-law source at $z = 0.7$ with photon-index $\Gamma = 1.7$ observed on the pn CCD, and obscured by various columns of gas (labelled as $N_{\rm H}$) local to the source, in addition to the constant Galactic column. We take all sources with $HR > -0.2$ to be "hard-spectrum" sources.

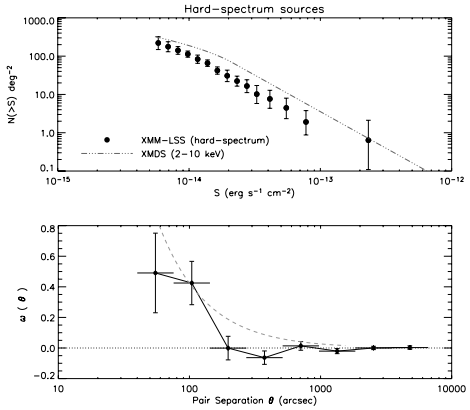


**Fig. 7.** $\log N$–$\log S$ (*top*) and ACF (*bottom*) for the 2–10 keV sample of 209 hard-spectrum ($HR > -0.2$) sources with $S/N > 2$ in the G pointings. In the $\log N$–$\log S$ plot, the dot-dashed line shown for comparison is the best-fit to the full 2–10 keV XMDS $\log N$–$\log S$, which also traces the slope of our full sample of hard-band sources (see Fig. 2). In the ACF plot, notice the larger range of the $y$-axis compared to Fig. 5. The dashed line is the best-fit ACF of Basilakos et al. (2004).

$HR > -0.2$. But only a marginal correlation signal was detected for these.

We note, however, that there is a large variation in the exposure-times and limiting fluxes of the individual pointings, especially between the B and G pointings, which could bias results due to varying efficiency of selecting obscured AGN across the field. In order to have a more uniform coverage, we then restricted our source selection to the deeper G-pointings only. The total number of 2–10 keV detections with $S/N > 2$ over the G pointings is 473. Of these, 400 have unique counterparts in the 0.5–2 keV band within a threshold inter-band distance of 10 arcsec, and 140 of these are hard-spectrum sources with $HR > -0.2$ (Fig. 6). In order to include the hardest sources, no S/N selection was imposed on the soft-band sample for this cross-correlation. Additionally, 69 sources (of 473) have no counterpart in the 0.5–2 keV band, and are likely to be very highly obscured ($>10^{23}$ cm$^{-2}$ or, possibly, Compton-thick) sources, if they are non-spurious detections (see Sect. 4.2). We combine the sub-samples of the 140 detections [with 0.5–2 keV counterparts] and of 69 detections [without counterparts], giving a total of 209 hard-spectrum sources.

The nearest neighbours (NN) distribution of these hard-spectrum sources shows departure from a uniform distribution at the 99.4 per cent probability level (based on the K-S test). The $\log N$–$\log S$ and auto-correlation function (ACF) of these 209 sources is plotted in Fig. 7. $\omega$ is clearly positive in the first two bins, representing a $3.0\sigma$ excess assuming Hamilton's formula (or $3.4\sigma$ assuming that of Efstathiou et al.). Though a power-law is not a good description of the signal, a full fit with the model of Eq. (3) results in a normalization $\theta_0 = 42^{+7}_{-13}$ arcsec and slope $\gamma = 3.1^{+1.1}_{-0.5}$ (uncertainties are, again, those relevant for one parameter of interest based on Poisson errors).

A similar auto-correlation analysis of the 260 sources with *soft* spectral count ratios ($HR < -0.2$) shows no such excess at scales less than 100 arcsec. We also note that the correlation signal of hard-spectrum sources is not dominated by the very hardest sources alone (the arrows in Fig. 6). Though the sample size is small, the NN statistic of only the 140 hard-spectrum sources *with* soft-band counterparts gives a K-S null hypothesis probability of only $10^{-3}$ compared to 100 random catalogues, again suggesting departure from a uniform distribution.

The main results of the correlation analysis on various sub-samples described above are summarized in Table 5.

### 3.5. On the significance level of observed correlations

Since the above constraints on clustering are relatively weak (a characteristic of angular clustering studies of sparse samples), it is pertinent to examine the significance levels quoted.

Initial tests above computed the distribution function of nearest neighbour separations, and the results were found to be consistent with the strength of the angular correlation function, as estimated by simple pair-counting over relevant scales. The Poisson errors used in the ACF fits, however, are strictly valid only for uncorrelated data. Bootstrap re-sampling (Barrow et al. 1984) is widely used to assess the internal reliability of a correlated dataset. Yet, as shown by Fisher et al. (1994), for sparse samples, this can over-estimate the true uncertainties by factors of ∼2 (and up to 4). For uncorrelated, and weakly correlated samples, such as ours, Poisson errors approximate the true errors, despite being a lower-limit. This approximation is likely to break down for the sample of hard-spectrum sources which shows the largest deviation from a random distribution. In this case, we re-sampled the entire ensemble of random catalogues a large number of times (≥50) with replacement in order to compute the bootstrap errors. We find $\theta_0 = 42'' \pm 22$ and $\gamma = 3.1 \pm 1.3$, implying $\sim 2\sigma$ constraints on both the normalization and the slope of the correlation function fit of the previous section (errors denote dispersion amongst the re-samples).

**Table 5.** Basic results of the auto-correlation analysis for various samples. "K-S" refers to the null hypothesis probability of the data and control samples being drawn from the same distribution. Power-law fits (Eq. (3)) to the ACF are listed in the final column, where computed or found to be significant.

| Selection criteria | | | K-S | ACF |
|---|---|---|---|---|
| $S/N > 3$ (B+G) | 0.5–2 keV | | 0.001 | $\theta_0 = 6.3'' \pm 3$; $\gamma = 2.2 \pm 0.2$ |
| $S/N > 3$ (B+G) | 2–10 keV | | 0.55 | – |
| $S/N > 2$ (G) | $''$ | ; $1 \geq HR > -0.2$ | 0.006 | $\theta_0 = 42''^{+7}_{-13}$; $\gamma = 3.1^{+1.1}_{-0.5}$ |
| $''$ | $''$ | ; $1 > HR > -0.2$ | 0.001 | – |

Since the bins used in the ACF analysis are themselves correlated at different scales, we compute the covariance matrix $\mathrm{Cov}(\theta_i, \theta_j)$ of $\omega$ returned by the above bootstrap method between all pairs of bins $(\theta_i, \theta_j)$ used for the fit. The correlation matrix is then calculated; it is simply the covariance matrix scaled to the diagonal elements as follows:

$$\mathrm{Corr}(\theta_i, \theta_j) = \mathrm{Cov}(\theta_i, \theta_j) / \sqrt{\mathrm{Cov}(\theta_i,\ \theta_i)\,\mathrm{Cov}(\theta_j,\ \theta_j)}.$$

To estimate the strength of the off-diagonal correlations, we follow Scranton et al. (2002) and form the scaled product of elements in each row (or column) $i$ of the correlation matrix: $P(i) = \prod_{r=1}^{N} |\mathrm{Corr}(\theta_i, \theta_r)|^{1/N}$. This product will be equal to one in the case of perfect correlation, while we find $P \lesssim 0.2$ at all scales of the correlation matrix, indicative of relatively small correlations. Thus, the "true" significance level for the ACF of the hard-spectrum sources is likely to be straddled by the 2- (bootstrap) and 3- (Poisson) $\sigma$ levels, as calculated above.

## 4. Discussion

The XMM-LSS survey is the widest, medium-deep, high galactic-latitude X-ray survey carried out by *XMM-Newton*. The survey currently covers a full area of 5.7 deg$^2$ and is designed to provide the best constraints on X-ray detected clusters and their evolution out to $z = 1$. The wide, contiguous coverage also gives ample opportunity to study the distribution of AGN in this field.

### 4.1. Comparison with other works

We find a positive two-point angular clustering signal with small correlation length in the 0.5–2 keV band, consistent within the errors with previous measurements using similar analysis (Vikhlinin & Forman 1995; Basilakos et al. 2005). We also note that the value of the projected correlation scale that we found is close to the size of the *XMM-Newton* PSF ($\sim 6''$), implying that some level of amplification bias may artificially result in an overestimation of the true clustering scale. The finite PSF can lead to confusion and an effective smoothing of the real source distribution, resulting in a larger correlation length for the density peaks of the observed, smoothed distribution (Kaiser 1984). We do not expect this bias to have a major effect, though, because the areal density of sources in any spatial resolution element (PSF) is small in relatively-shallow surveys such as ours (see also Basilakos et al. 2005).

In the hard (2–10 keV) band, we do not detect any clustering signal of the projected source distribution on the sky, unlike results by Yang et al. (2003) and Basilakos et al. (2004). The main differences of our survey with respect to the other works are: i) the analyzed region covers 4.2 deg$^2$ which is at least a factor of two larger than that of Basilakos et al. (2004) and $\sim$10 times larger than Yang et al. (2003); ii) this coverage is contiguous (or at least pseudo-contiguous, covering 83 per cent of the area of a $2 \times 2.5$ deg$^2$ rectangle); iii) in terms of flux, while we probe sources slightly deeper than the nominal threshold hard band flux level of $\sim 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ of Basilakos et al. (2004), the *Chandra* survey of Yang et al. (2003) probes down to $3 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$, approximately 2–3 times fainter than us[5].

Are we then simply seeing a "truer" picture of the distribution of AGN in X-rays as sky coverage is increased and improved? Or is the XMM-LSS area special in some way? We first note that $\log N$–$\log S$ of the XMM-LSS field is slightly lower than other results published in the literature at fluxes brighter than $\sim 2 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ in both bands. The field for the XMM-LSS survey was explicitly defined so as to avoid previously-known, bright X-ray sources and the deficit seen could simply reflect cosmic variance in the X-ray sky. Indeed, cosmic variance is known to cause uncertainties in the normalization of the cosmic X-ray background level of $\sim$30 per cent above 2 keV (Cowie et al. 2002; Barcons et al. 2000). Fluctuations in the *HEAO 1 A-2* X-ray background map on scales of a few degrees have been observed (Boughn et al. 2002; Fabian & Barcons 1992) and it may be plausible that our field sits in a comparatively large "void" of large scale structure. The fraction the hard X-ray background resolved out in our field is, however, consistent with that found by other surveys (Sect 2.4). Yang et al. (2003) also noted that cosmic variance manifests as "voids", but on scales much less than a sq. deg.

Due to the sharper *Chandra* PSF and the deeper flux limit of their survey, the positive result of Yang et al. (2003) may be understood in terms of their much better sensitivity to detect obscured AGN and any associated clustering signal. Not enough details are available for an assessment of any other differences (such as pipeline selection effects, differences in ACF fitting etc.) with respect to Basilakos et al. (2004). But a crude comparison shows a definite deficit of source pairs at small separations in our field. Using their Fig. 1 and Eq. (1), we infer that their fields contain $\sim$9 source pairs with a separation of between 50 and 60 arcsec over their area of $\approx$2 deg$^2$ at a similar (or slightly higher) flux limit, while we detect exactly 9 such pairs over our full area of 4.2 deg$^2$. Another difference is that we use the average $\log N$–$\log S$ of the XMM-LSS field itself to generate the fluxes of the random catalogues, while they use the distribution from another field (that of Baldi et al. 2002), though their data is consistent with the distribution used (at least at fluxes brighter than $5 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$; see their Fig. 1). We also note that Basilakos et al. (2004) do not include shot noise (Sect. 3.1) in the measurement of their random fluxes, but we found that switching off this extra Poisson noise during the generation of our random fluxes had only minor effects on the determination of the ACF. Finally, any bins with a negative $\omega$ contribution at small scales (which often occur in sparse, weakly-clustered samples) are neglected by them. Full source characterization and follow-up in our field should help to discern the correct cause of the observed differences.

[5] Both quoted limits are in the 2–8 keV band. The corresponding flux limit in 2–10 keV should be $\approx$20 per cent brighter, assuming a photon-index $\Gamma = 1.7$; the difference in band definitions is not a dominant source of discrepancy.

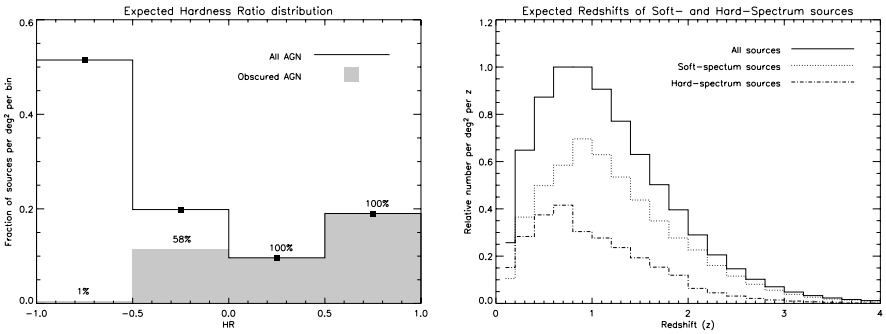


**Fig. 8.** Predictions of the distribution and classification of AGN as soft- and hard-spectrum sources in the XMM-LSS. (*Left*) The plot shows the expected hardness ratio histogram of AGN with fluxes $>8\times10^{-15}$ erg s$^{-1}$ cm$^{-2}$ in the 2–10 keV band, assuming the XLF and $N_{\rm H}$ distribution of Ueda et al. (2003). The EPIC-pn camera response with the Thin filter is assumed. The fraction of obscured AGN ($N_{\rm H} > 10^{22}$ cm$^{-2}$) is shaded in grey and their relative percentage of all sources in each bin is labelled. While most ($\gtrsim$50 per cent) of the detected sources have very low values of *HR* (and lie in the first bin), the plot shows that at high *HR* values, obscured AGN completely dominate. (*Right*) The plot shows the expected redshift distribution of AGN classified as soft-spectrum and hard-spectrum sources above the same flux limit, assuming the luminosity-dependent density evolution model of Ueda et al. Hard-spectrum sources peak at $z \sim 0.7$.

### *4.2. Selection and correlation of obscured AGN*

We have made predictions for the expected distribution of *HR* that would be observed for AGN spread in redshift and affected by different columns of intrinsic obscuring gas. We assumed that our sample follows a luminosity function (XLF) and obscuring column density ($N_{\rm H}$) distribution as calculated in recent work (Ueda et al. 2003), and folded the number counts predicted for a limiting flux $S_{\rm lim}^{2-10} = 8\times10^{-15}$ erg s$^{-1}$ cm$^{-2}$ through the EPIC response function (for simplicity, we used the pn on-axis response only). As input templates for this calculation, we used power-law X-ray spectra with a fixed, intrinsic photon-index of $\Gamma = 1.9$ (e.g., Mateos et al. 2005) and modelled photoelectric absorption and Compton reflection (Gandhi & Fabian 2003), but did not include Compton-thick sources ($N_{\rm H} > 10^{24}$ cm$^{-2}$). Obscured AGN are defined as those with $N_{\rm H} > 10^{22}$ cm$^{-2}$. The first plot in Fig. 8 shows that the dominant contributors to the sample of objects with $HR > -0.2$ (hard-spectrum sources) will be obscured AGN. Intrinsic photon-index variations may push some *un*obscured AGN into the high *HR* regime. Moreover, obscured AGN at high redshift are likely to have low *HR* values (due to positive *k*-correction into the soft band), so this threshold definitely does not select all obscured AGN uniquely; but given the relatively-shallow depth of our survey, obscured AGN will still dominate this sample. Also note that our observed sample of hardness ratios in Fig. 6 shows an increase of *HR* with decreasing flux, as is typically associated with obscured AGN in which the soft-band flux is depleted due to photo-electric absorption in the torus.

Over the deepest region of our survey, we detect marginal, but positive correlation at the $\sim2-3\sigma$ level for the sample of hard-spectrum sources, at angular separations below $\sim$100 arcsec. Interestingly, the values of $\omega$ at these separations match well those inferred by Basilakos et al. (2004) and Yang et al. (2003) for the full hard band, which might suggest cosmic variance of obscured AGN specifically. Since obscured AGN begin to dominate the source counts only at faint fluxes, we had to include sources with a less stringent detection significance ($S/N > 2$) for this analysis. We can be confident that most of these are not spurious because a large fraction (85 per cent) also have associated 0.5–2 keV counterparts (see Fig. 6). The most suspect detections are the hardest, faint ones with no 0.5–2 keV counterparts, but the correlation signal is not adversely affected by removal of these sources (Sect. 3.4). Finally, we note that if this sample were dominated by spurious fluctuations, any intrinsic clustering signal would have been diminished, rather than enhanced as we observe, due to their Poisson nature.

#### 4.2.1. Inversion of the Limber equation

An estimate of the real space correlation function can be obtained by inversion of the Limber equation that connects the true space correlation scale $r_0$ with the projected scale $\theta_0$ (see, e.g., Peebles 1980; Wilman et al. 2003). For this, an estimate of the source redshift distribution and a model for the clustering evolution is required. We assume that the AGN in our survey follow the expected redshift distribution based on the Ueda et al. (2003) model as described above. Predictions can be made separately for sources that would be classified as either soft-spectrum, or hard-spectrum AGN, based on their count ratios, irrespective of whether they are intrinsically obscured or unobscured. This redshift distribution is shown in Fig. 8. Hard-spectrum sources peak at $z = 0.7$, with a detectable tail extending out to $z \sim 2$ at least. Though the power-law model of Sect. 3.4 is not a good model to fit, assuming the best-fit parameters found therein, and inverting the Limber equation, we find $r_0 = 6(\pm3)h^{-1}$ Mpc [$H_0 = 100h$ km s$^{-1}$ Mpc$^{-1}$] by assuming co-moving clustering evolution and the redshift distribution of hard-spectrum sources in Fig. 8. This scale is similar to that of local, optically-selected galaxies (Davis & Peebles 1983) or that of optically-selected QSOs at $z \sim 1$ (Croom et al. 2001). On the other hand, it is also consistent at the $\sim 2\sigma$ level with the stronger clustering of extremely-red objects and powerful radio galaxies (e.g., Röttgering et al. 2003 and references therein). Given the weakness of the contraints above, however, we defer further discussion on this until source follow-up is complete.

### 4.3. Implications

The typical luminosities of AGN probed in medium-deep surveys such as the XMM-LSS will be $L_{\rm X-ray} \gtrsim 10^{43.5}$ erg s$^{-1}$ (e.g., Gandhi et al. 2004). A small correlation length is then consistent with X-ray selected Seyferts being unbiased tracers of structure. This is also in qualitative agreement with studies such as that of Waskett et al. (2005), who found no difference in the environments of AGN as compared to those of inactive galaxies. Wake et al. (2004) also arrived at a similar conclusion from a large sample of optically-selected AGN.

On the other hand, if the auto-correlation of hard-spectrum X-ray sources is indeed larger, it would suggest that obscured AGN are preferentially associated with higher density peaks in the underlying matter distribution; the large amount of gas and dust present in the environment not only triggers AGN activity, but also hides the AGN itself. Galaxy mergers could provide the gas necessary to achieve both (e.g., Hopkins et al. 2005), as is the case for the population of Ultra Luminous Infra Red Galaxies. In fact, assuming that AGN found in X-rays are correlated with powerful-infrared, obscured starbursts known to peak at $z \sim 0.7$ (e.g., Chary & Elbaz 2001), models can be constructed to explain the X-ray background spectrum as well as X-ray source counts (Franceschini et al. 2002; Gandhi & Fabian 2003). The mechanism that could drive the gas and dust from the large-scale environment to the scales of galactic nuclei remains unclear, however. Assuming a median redshift of 0.7, angular separations of 50–100 arcsec (the first two bins of the ACF in Fig. 7) correspond to projected physical separations of ∼350–700 kpc, too large for these systems to be bound mergers in advanced stages. Associations with filaments and groupings in the large scale matter distribution is a possibility. Interactions with any enhanced density of minor galaxies also associated with the large scale structure could provide the necessary torques to trigger gas inflows towards the nucleus, via bars or bound mergers, for instance (Shlosman et al. 1989; de Robertis et al. 1998).

In contrast, more recent X-ray luminosity function determinations and background synthesis models (over shallower areas of sky at the relevant flux limits, but with extensive redshift coverage) do not require separate evolutionary and/or formation scenarios for obscured and unobscured AGN (e.g., Treister et al. 2004; Ueda et al. 2003): these would predict no difference in the comparative correlation of obscured and unobscured AGN. Given the low overall significance of our detected correlation signal for hard-spectrum sources, we can certainly not rule out this possibility. Indeed, in a recent spatial clustering analysis, Yang et al. (2006) found no difference in the correlation properties of hard- and soft-spectrum AGN (see also Gilli et al. 2005), in contrast to their previous results in angular coordinates for the same sample. This difference might be the result of dissimilar redshift distributions of the two classes of sources; whatever the reason, it underscores the need for complete studies over larger areas of sky.

Other on-going works which will be able to measure angular correlations include the AXIS (Carrera et al. 2006) and the COSMOS surveys (Hasinger et al. 2006). The forthcoming expansion of the XMM-LSS survey itelf, along with its deep, multi-wavelength follow-up will provide good constraints on the spatial distribution and clustering of AGN as well as clusters. With an area of 10 deg$^2$, we expect to find at least 1000 AGN above a 2–10 keV flux limit of $8 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ (corresponding to our $S/N > 3$ criterion). Additionally, the numbers of obscured AGN detected should double compared to our present sample. The uncertainties on clustering statistics will decrease by a further factor of ∼2, giving a proper determination of the slope and scale length and a much better account of the cosmic variance.

*Acknowledgements.* P.G. thanks the European Southern Observatory Fellowship programme, and the X-ray group at the Institute of Astronomy in Cambridge for their hospitality and computer support during the final stages of this work. We thank the anonymous referee for suggestions that improved the robustness of our procedures. X-ray processing and extensive simulations were performed at the Centre de Calcul de l'IN2P3 in Lyon. The Liège team acknowledges support from PRODEX (XMM). Part of the research was also performed in the framework of the IUAP P5/36 project, supported by the OSTC Belgian Federal services. L.C. and D.M. acknowledge financial contribution from contract ASI-INAF I/023/05/0.

## References

Andreon, S., Valtchanov, I., Jones, L. R., et al. 2005, MNRAS, 359, 1250
Arnaud, K. A. 1996, in Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby and J. Barnes, ASP Conf. Ser., 101, 5, 17
Baldi, A., Molendi, S., Comastri, A., et al. 2002, ApJ, 564, 190
Barcons, X., Mateos, S., & Ceballos, M. T. 2000, MNRAS, 316, L13
Barrow, J. D., Bhavsar, S. P., & Sonoda, D. H. 1984, MNRAS, 210, 19P
Basilakos, S., Georgakakis, A., Plionis, M., & Georgantopoulos, I. 2004, ApJ, 607, L79
Basilakos, S., Plionis, M., Georgakakis, A., & Georgantopoulos, I. 2005, MNRAS, 356, 183
Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393
Boughn, S. P., Crittenden, R. G., & Koehrsen, G. P. 2002, ApJ, 580, 672
Brandt, W. N., Alexander, D. M., Bauer, F. E., & Vignali, C. 2004, Physics of Active Galactic Nuclei at All Scales, ed. D. Alloin, R. Johnson, P. Lira (Berlin: Springer-Verlag), [arXiv:astro-ph/0403646]
Brandt, W. N. & Hasinger, G. 2005, ARA&A, 43, 827
Carrera, F. J., Barcons, X., Fabian, A. C., et al. 1998, MNRAS, 299, 229
Carrera, F. J., Ebrero, S., Mateos, S., et al. 2006, in preparation
Chary, R. & Elbaz, D. 2001, ApJ, 556, 562
Chiappetti, L., Tajer, M., Trinchieri, G., et al. 2005, A&A, 439, 413
Cohen, A. S., Röttgering, H. J. A., Kassim, N. E., et al. 2003, ApJ, 591, 640
Cowie, L. L., Garmire, G. P., Bautz, M. W., et al. 2002, ApJ, 566, L5
Crawford, C. S., Gandhi, P., Fabian, A. C., et al. 2002, MNRAS, 333, 809
Croom, S. M., Shanks, T., Boyle, B. J., et al. 2001, MNRAS, 325, 483
Davis, M., & Peebles, P. J. E. 1983, ApJ, 267, 465
De Luca, A., & Molendi, S. 2004, A&A, 419, 837
de Robertis, M. M., Yee, H. K. C., & Hayhoe, K. 1998, ApJ, 496, 93
Efstathiou, G., Bernstein, G., Tyson, J. A., Katz, N., & Guhathakurta, P. 1991, ApJ, 380, L47
Fabian, A. C., & Barcons, X. 1992, ARA&A, 30, 429
Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. 1994, MNRAS, 266, 50
Franceschini, A., Braito, V., & Fadda, D. 2002, MNRAS, 335, L51
Gandhi, P., & Fabian, A. C. 2003, MNRAS, 339, 1095
Gandhi, P., Crawford, C. S., Fabian, A. C., & Johnstone, R. M. 2004, MNRAS, 348, 529
Gehrels, N. 1986, ApJ, 303, 336
Gilli, R., Daddi, E., Zamorani, G., et al. 2005, A&A, 430, 811
Hamilton, A. J. S. 1993, ApJ, 417, 19
Hasinger, G., Cappelluti,, N., Brunner, H., et al. 2006, ApJS, submitted
Hickox, R. C., & Markevitch, M. 2006, ApJ, submitted [arXiv:astro-ph/0512542]
Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2005, ApJ, 630, 705
Joyce, M., Montuori, M., & Labini, F. S. 1999, ApJ, 514, L5
Kaiser, N. 1984, ApJ, 284, L9
Le Fèvre, O., Mellier, Y., McCracken, H. J., et al. 2004, A&A, 417, 839
Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., et al. 2003, PASP, 115, 897
Maiolino, R., & Rieke, G. H. 1995, ApJ, 454, 95
Manners, J. C., Johnson, O., Almaini, O., et al. 2003, MNRAS, 343, 293
Mateos, S., Barcons, X., Carrera, F. J., et al. 2005, A&A, 433, 855
Matt, G., Fabian, A. C., Guainazzi, M., et al. 2000, MNRAS, 318, 173
Mullis, C. R., Henry, J. P., Gioia, I. M., et al. 2004, ApJ, 617, 192
Mushotzky, R. F., Cowie, L. L., Barger, A. J., & Arnaud, K. A. 2000, Nature 404, 459
Nandra, K., Georgantopoulos, I., Brotherton, M., & Papadakis, I. E. 2004, MNRAS, 347, L41
Pacaud, F., Pierre, M., Refregier, A., et al. 2006, MNRAS, accepted

Padovani, P., Allen, M. G., Rosati, P., & Walton, N. A. 2004, A&A, 424, 545
Peebles, P. J. E. 1980, The large-scale structure of the universe (Princeton University Press)
Pierre, M., Pacaud, F., Duc, P.-A., et al. 2006, MNRAS, accepted
Pierre, M., Valtchanov, I., Santos, S. D., et al. 2004, J. Cosm. Astropart. Phys., 9, 11
Read, A. M., & Ponman, T. J. 2003, A&A, 409, 395
Refregier, A., Valtchanov, I., & Pierre, M. 2002, A&A, 390, 1
Rosati, P., Borgani, S., & Norman, C. 2002, ARA&A, 40, 539
Röttgering, H., Daddi, E., Overzier, R., & Wilman, R. 2003, New Astron. Rev., 47, 309
Schmidt, M., Hasinger, G., Gunn, J., et al. 1998, A&A, 329, 495
Scranton, R., Johnston, D., Dodelson, S., et al. 2002, ApJ, 579, 48
Setti, G., & Woltjer, L. 1989, A&A, 224, L21
Shlosman, I., Frank, J., & Begelman, M. C. 1989, Nature, 338, 45
Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175
Szokoly, G. P., Bergeron, J., Hasinger, G., et al. 2004, ApJS, 155, 271
Treister, E., Urry, C. M., Chatzichristou, E., et al. 2004, ApJ, 616, 123
Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, ApJ, 598, 886
Valtchanov, I., Pierre, M., Willis, J., et al. 2004, A&A, 423, 75
Vikhlinin, A. & Forman, W. 1995, ApJ, 455, L109
Wake, D. A., Miller, C. J., Di Matteo, T., et al. 2004, ApJ, 610, L85
Waskett, T. J., Eales, S. A., Gear, W. K., et al. 2005, MNRAS, accepted [arXiv:astro-ph/0508102]
Webster, A. 1976, MNRAS, 175, 61
Willis, J. P., Pacaud, F., Valtchanov, I., et al. 2005a, MNRAS, 363, 675
Willis, J. P., Pacaud, F., Valtchanov, I., et al. 2005b, MNRAS, 364, 751
Wilman, R. J., Röttgering, H. J. A., Overzier, R. A., & Jarvis, M. J. 2003, MNRAS, 339, 695
Worsley, M. A., Fabian, A. C., Barcons, X., et al. 2004, MNRAS, 352, L28
Yang, Y., Mushotzky, R. F., Barger, A. J., et al. 2003, ApJ, 585, L85
Yang, Y., Mushotzky, R. F., Barger, A. J., & Cowie, L. L. 2006, ApJ, accepted, [arXiv:astro-ph/0601634]
Zacharias, N., Urban, S. E., Zacharias, M. I., et al. 2004, AJ, 127, 3043

## 7.7 Comparaison des relevés X et lensing : "Combined analysis of weak lensing and X-ray blind surveys"

### Contexte/motivation

Les résultats de Pacaud et al. (2007) démontrent clairement la nécessité de connaître la fonction de sélection d'un échantillon d'amas pour en étudier les relations d'échelles, ainsi que l'impact de l'indétermination de ces relations d'échelles sur les contraintes cosmologiques. Pour ces deux raisons, il semble profitable de combiner des relevés en aveugle dans différentes longueurs d'onde (optique, SZE, cisaillement gravitationnel). Cela permet à la fois de mieux contraindre les propriétés de chaques sources dans l'échantillon, mais également de valider la modélisation cosmologique par la combinaison de différentes fonctions de sélection.

### Résumé

Dans cet article, nous présentons une étude conjointe des effets de lentille gravitationnel faibles et de l'emission X basée sur 4 deg$^2$ des relevés CFHTLS et XMM-LSS. L'analyse du cisaillement gravitationnel présentée ici est la première à utiliser, pour des données réelles, la technique de nouvelle génération dite des *shapelets*. Nous créons des cartes de masse projetée, et isolons 6 amas de galaxies. Nous montrons que leur comptage peut être utilisé pour contraindre la normalisation du spectre de puissance $\sigma_8$ à une valeur de $\sigma_8 = 0.92^{+0.26}_{-0.30}$ pour $\Omega_m = 0.24$. Une grande dispersion est généralement observée pour la relation $M$-$T_X$ déduite des mesures de masse par cisaillement gravitationnel. Pour autant, des contraintes précises peuvent être obtenues sur sa pente et sa normalisation $M_\star$ avec un nombre limité de source pourvu que l'intervalle de masse couvert soit suffisamment large. Ajouter les amas de Bardeau et al. (2007) à notre échantillon permet ainsi de mesurer $M_\star = 2.71^{+0.79}_{-0.61}\,10^{14}h^{-1}M_\odot$. Bien que dominées par le bruit de grenaille et la variance cosmique, nos mesures sont compatibles avec les valeurs actuellement favorisées, et préfigurent les possibilités des relevés à venir. Nous explorons donc la dépendance de ces contraintes envers la taille, la profondeur et le temps d'exposition pour des relevés simultanées de cisaillement gravitationnel et d'émission X. Nous montrons que les relevés profond devraient se focaliser sur l'étude de la physique des groupes et amas de galaxies. Pour un temps d'exposition équivalent, les relevés à grande échelle génèrent en effet un nombre de détection plus élevé et sont donc preferables pour la mesure des paramètres cosmologiques comme $\sigma_8$ et $M_\star$. Des relevés de plusieurs centaines de degrés carrés sont requis pour améliorer notre connaissance de ces paramètres. Des relevés plus ambitieux recouvrant $\sim$7000 deg$^2$ permettraient d'atteindre une précision de 1% sur la mesure des normalisations du spectre de puissance et de la relation $M$-$T_X$.

### Contribution personnelle

L'article se base sur mes catalogues C1. Outre les réflexions communes sur l'ensemble du contenu et la mise à disposition des données X, j'ai particulièrement contribué à la mesure de la relation $M$-$T_X$ et à la discussion des résultats.

Mon. Not. R. Astron. Soc. **000**, 1–14 () Printed 12 December 2007 (MN LATEX style file v2.2)

# Combined analysis of weak lensing and X-ray blind surveys⋆

Joel Bergé [1]†, Florian Pacaud [1,2], Alexandre Réfrégier [1], Richard Massey [3], Marguerite Pierre [1], Adam Amara [1], Mark Birkinshaw [4], Stéphane Paulin-Henriksson [1], Graham P. Smith [5,3], Jon Willis [6]

[1] *Laboratoire AIM, CEA/DSM - CNRS - Université Paris Diderot, DAPNIA/SAp, 91191 Gif-sur-Yvette, France*
[2] *Argelander Institute für Astronomy, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany*
[3] *California Institute of Technology, 1200 E. California Blvd, Pasadena CA 91125, USA*
[4] *Department of Physics, University of Bristol, Tyndall Avenue, Bristol BS8 ITL, UK*
[5] *School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK*
[6] *Department of Physics and Astronomy, University of Victoria, Elliot Building, 380 Finnerty Road, Victoria, V8V 1A1, BC, Canada*

Accepted . Received ; in original form

**ABSTRACT**

We present a joint weak lensing and X-ray analysis of 4 deg$^2$ from the CFHTLS and XMM-LSS surveys. Our weak lensing analysis is the first analysis of a real survey using shapelets, a next generation weak lensing analysis method. We create projected mass maps of the images, and extract 6 weak-lensing-detected clusters of galaxies. We show that their counts can be used to constrain the power spectrum normalisation $\sigma_8$ to be $\sigma_8 = 0.92^{+0.26}_{-0.30}$ for $\Omega_m = 0.24$. We show that despite the large scatter generally observed in the M-T relation derived from lensing masses, tight constraints on both its slope and normalisation $M_*$ can be obtained with a moderate number of sources provided that the covered mass range is large enough. Adding clusters from Bardeau et al. (2007) to our sample, we measure $M_* = 2.71^{+0.79}_{-0.61}\, 10^{14} h^{-1} M_\odot$. Although they are dominated by shot noise and sample variance, our measurements are consistent with currently favoured values, and set the stage for future surveys. We thus investigate the dependence of those estimates on survey size, depth, and integration time, for joint weak lensing and X-ray surveys. We show that deep surveys should be dedicated to the study of the physics of clusters and groups of galaxies. For a given exposure time, wide surveys provide a larger number of detected clusters and are therefore preferred for the measurement of cosmological parameters such as $\sigma_8$ and $M_*$. We show that a wide survey of a few hundreds square degrees is needed to improve upon current measurements of these parameters. More ambitious surveys covering 7000 deg$^2$ will provide the 1% accuracy in the estimation of the power spectrum and the M-T relation normalisations.

**Key words:** gravitational lensing - surveys - dark matter - large-scale structure of Universe - cosmological parameters - X-rays: galaxies: clusters

## 1 INTRODUCTION

In the currently-favoured hierarchical model of structure formation, clusters of galaxies have formed from the collapse of gravitational potential wells (e.g. Peebles 1980; Padmanabhan 1993; Lacey & Cole 1993; Lokas 2001) and are powerful cosmological probes. For instance, since they are sensitive to the expansion history of the Universe, their abundance and spatial distribution (e.g. Viana & Liddle 1996; Wang & Steinhardt 1998; Horellou & Berge 2005; Nunes, da Silva & Aghanim 2006; Manera & Mota 2006) and their mass function (e.g. Lokas, Bode & Hoffman 2004) depend on cosmo-

⋆ Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS. It makes use of photometric redshifts produced jointly by TERAPIX and VVDS teams.

† E-mail address : joel.berge@cea.fr

logical parameters, such as the dark energy equation of state parameter $w$ (e.g. Basilakos 2003; Maor & Lahav 2005; Basilakos & Voglis 2007), or the power spectrum normalisation $\sigma_8$ (e.g. Seljak 2002; Pierpaoli et al. 2003). Several observational methods now permit the use of clusters of galaxies as cosmological probes, such as X-ray observations, weak gravitational lensing and the Sunyaev-Zeldovich effect.

Due to improvements in telescopes and techniques, X-ray studies are able to constrain cluster physics and the mass scaling relation with ever greater precision. For instance, the self-similarity for clusters of galaxies (Eke, Navarro & Frenk 1998; Arnaud, Aghanim & Neumann 2002) has been observationally verified. Nevertheless, adiabatic simulations still predicts a mass-temperature relation with double the observed normalisation (e.g. Nevalainen, Markevitch & Forman 2000; Finoguenov, Reiprich & Böhringer 2001) and the self-similarity assumption could break down at low temperatures ($T \leqslant 3$ keV). Thus, a steepening of the mass-temperature (M-T) relation is expected if galaxy groups underwent a preheating by supernovae, or a surge of entropy, in their early days (Bialek, Evrard & Mohr 2001; Muanwong et al. 2002). Recent evidence for this steepening was found by e.g. Nevalainen, Markevitch & Forman (2000); Finoguenov, Reiprich & Böhringer (2001), or Arnaud, Pointecouteau & Pratt 2005 (APP05 hereafter), but could not be seen by e.g. Ettori, De Grandi & Molendi (2002); Castillo-Morales & Schindler (2003); Vikhlinin et al. (2006). Moreover, the M-T normalisation estimation is currently limited by systematics in measuring cluster masses from their X-ray profiles. This limitation can be lifted by using probes which are independent of the physical state of the cluster.

Beyond galaxy cluster physics, the M-T relation is needed by X-ray experiments to estimate the power spectrum normalisation $\sigma_8$. Measuring this parameter has triggered much efforts in several observational areas. For instance, CMB experiments (e.g. Spergel et al. 2007) tend to a low value for $\sigma_8$ ($\leqslant 0.8$), weak lensing experiments tend to higher values ($\geqslant 0.8$), and X-ray observations provide intermediate values.

Gravitational lensing does not depend on the underlying physics of clusters of galaxies or dark matter, but only on their potential wells, and on the Universe's geometry. Strong gravitational lensing has been first used for galaxy clusters physics (e.g. Mellier, Fort & Kneib 1993; Kneib et al. 1995, 1996; Smail et al. 1997; Luppino et al. 1999; Smith et al. 2005) and measurement of $\sigma_8$ (e.g. Smith et al. 2003). Weak gravitational lensing is more difficult to measure (for reviews, see e.g. Mellier 1999; Bartelmann & Schneider 2001; Refregier 2003a; Munshi et al. 2006)), and has taken longer to be detected (Bacon, Refregier & Ellis 2000; van Waerbeke et al. 2000; Wittman et al. 2000; Rhodes, Refregier & Groth 2001). Since then, particular attention has been given to cosmic shear, i.e. statistical cosmological weak lensing (e.g. Bacon et al. 2003; Heymans et al. 2005; Massey et al. 2005, 2007a; Hoekstra et al. 2006; Semboloni et al. 2006; Schrabback et al. 2007), in attempts to measure $w$ and $\sigma_8$. It has also begun to be used as a tracer of the cosmic web (e.g. Massey et al. 2007b), and a way to detect and catalogue mass overdensities (e.g. Wittman et al. 2006, Gavazzi & Soucail 2007, GS07 hereafter, and Miyazaki et al. 2007). Beside the constraints it can bring to cosmology, it can be used as a complement to X-ray analyses of clusters of galaxies. Thanks to the physics-independent estimation of cluster masses, it appears as a unique method to calibrate the mass-temperature relation for clusters of galaxies (e.g. Hjorth, Oukbir & van Kampen 1998; Huterer & White 2002; Pedersen & Dahle 2006; Bardeau et al. 2007). It has been shown that the uncertainty in the normalisation of the mass-temperature relation is the largest source of error in $\sigma_8$ measurements inferred from X-ray cluster analyses (Seljak 2002; Pierpaoli et al. 2003). An accurate mass-temperature relation, obtained from combined weak lensing and X-ray analyses, will thus provide new insights not only on the $\sigma_8$ discrepancy, but also on galaxy cluster physics.

In this paper, we present the first joint analysis of weak gravitational lensing and X-ray wide-area surveys of a randomly-selected patch of sky. The weak lensing survey is derived from the CFHTLS, and the X-ray survey from the XMM-LSS. The weak lensing analysis uses shapelets (Refregier 2003b; Refregier & Bacon 2003; Massey & Refregier 2005), a new generation shear measurement technique, which has been shown to achieve a few percent accuracy in shear measurement from ground based telescopes (Massey et al. 2007c). We analyse one square-degree of the CFHTLS Deep survey (the D1 field) and four contiguous square degrees of the CFHTLS Wide survey, which enclose the D1 field. We create convergence maps for this region of the sky and give a catalogue of detections. We show how counting weak-lensing-selected clusters can provide an estimate of the power spectrum normalisation $\sigma_8$. We then show how the combination of weak lensing and X-ray analyses of clusters provides an estimate of the mass-temperature relation normalisation $T_*$, independent of clusters physical state. Finally, we investigate the impact of a joint weak lensing and X-ray survey strategy on the accuracy of the $\sigma_8$ and $T_*$ measurement. We consider deep and wide weak lensing surveys, with the CFHTLS characteristics, combined with a blind X-ray survey of the same region of the sky.

The organisation of the paper is as follows. Section 2 presents the surveys used in this paper, namely the CFHTLS and the XMM-LSS. The methods that we use are described in section 3. We show how we estimate the weak lensing effect using shapelets, and how we generate convergence maps and detect clusters. We also briefly describe how the X-ray properties of clusters are obtained. Section 4 presents the convergence map we inferred from our weak lensing analysis, and gives a catalogue of the galaxy clusters that we detect. We then give our estimates of the normalisation of the power spectrum and of the M-T relation. We then show in section 5 that combined blind surveys are necessary to get the best insights about those normalisations. The impact of survey strategy on those parameters estimations is discussed in section 6. We conclude in section 7.

## 2 DATA

### 2.1 Weak lensing : CFHTLS

The "Canada-France-Hawaii Telescope Legacy Survey" [1] (CFHTLS), a joint France-Canada project, consists of three different surveys, namely the Very Wide Survey, the Wide

[1] http://www.cfht.hawaii.edu/Science/CFHLS/

Synoptic Survey (referred to as "Wide survey" hereafter), and the Deep Survey. Once complete, the Wide Survey will cover 170 deg$^2$ (divided into four distinct patches ranging from 49 deg$^2$ to 72 deg$^2$) in five filters (u*,g',r',i',z'), down to a magnitude $i' \approx 24.5$. Its main goal is the study of large scale structures by weak gravitational lensing and galaxy counts. The Deep Survey covers 4 different fields, each with an area of 1 deg$^2$, in the same five filters, down to $i' \approx 28.5$. It is primarily intended for Type Ia Supernovae studies but it is also useful for measuring large-scale structures. The CFHTLS images were obtained from observations with the MegaCam camera, made of a 36 CCD mosaic, of 2048 × 4196 pixels each, with a 1 deg$^2$ field of view (Boulade et al. 2003).

In this paper, we present the weak gravitational lensing based mapping of 4 deg$^2$ of the Wide Survey (W1 patch), which include the 1 deg$^2$ field of the Deep Survey (D1 field), using both W1 and D1 images. The geometry of the fields that we use is shown in Figure 1. The data processing (astrometry, photometric calibration, stacking of images) has been done by the CFHT community and Terapix [2]. We use W1 images optimised for weak lensing : each field is the combination of 7 stacked images, each of 620 seconds exposure time. We use the T0003 release of the D1 field, consisting of 275 stacked images, with a total 37.4 hour exposure time. The average seeing is 0.7 arcsec. We masked parts of the images with saturated stars and/or too high a noise, by hand, so as not to bias our weak lensing results. This operation removes 10% of the original area covered by the data. We do not mask the ghosts created by spurious reflections on the telescope optics around saturated stars, but we eventually remove the galaxies that they cover from our catalogues, since they are too noisy. The average galaxy density is 28 arcmin$^{-2}$ in the D1 image, and 13 arcmin$^{-2}$ in the W1 images.

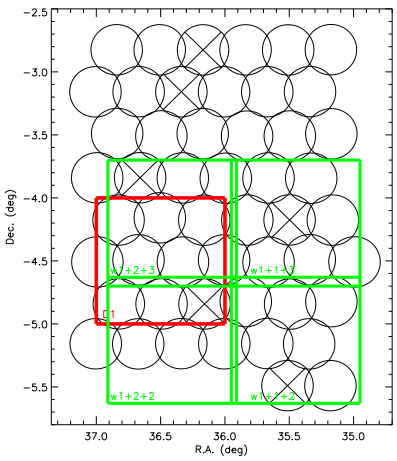


**Figure 1.** Layout of the surveys. The red square is the CFHTLS D1 field. The four green squares are the four CFHTLS W1 fields used in this paper. Circles represent the XMM-LSS pointings available in that region, prior to XMM AO5 (Pacaud et al. 2007). Here, we only used those lying within the optical data. Those marked by a cross are strongly affected by flares and are being re-observed.

### 2.2 X-ray : XMM-LSS

Designed to cover an area of several tens of square degrees up to a redshift $z = 1$, the XMM-LSS survey aims at detecting a significant fraction of clusters of galaxies, in order to constitute a sample suited to cosmological studies (Pierre et al. 2004). Its nominal exposure times are 10 ks, and have been raised up to 20 ks for the *XMM* Medium Deep Survey (XMDS : Chiappetti et al. 2005), a 2 deg$^2$ region included in the XMM-LSS, which covers the CFHTLS D1 field. In this paper, we use 4 deg$^2$ of the XMM-LSS which cover our 4 deg$^2$ CFHTLS data. The *XMM* pointings are shown on Figure 1. The raw X-ray observations reduction is presented in Pacaud et al. (2006).

## 3 METHOD

### 3.1 Weak lensing cluster detection

Introduced in Refregier 2003b, Refregier & Bacon 2003 and Massey & Refregier 2005, shapelets are a complete, orthogonal, set of basis functions with which one can analytically decompose galaxy shapes. They can be seen as fundamental shapes : a particular galaxy can be represented by a particular sum of shapelets basis functions $\chi_{n,m}$, weighted by coefficients $f_{n,m}$. Their rich formalism provides an intuitive and analytic form for geometrical transformations (such as smear, shear, rotation) and for (de)convolution. Hence, they allow one to analytically describe the smearing of the Point Spread Function (PSF) and the shear of galaxies, properly correcting for the PSF. The shear estimation they provide has been shown to reach the needed accuracy for the CFHTLS specifications by the STEP project (Massey et al. 2007c).

Our full pipeline will be described in an upcoming paper (Bergé et al. in prep). We briefly summarise it hereafter. Each sufficiently bright and non-saturated star is first decomposed into shapelets. A polynomial interpolation of each shapelet coefficient then provides a model of its spatial variations across the image. We are thus able to reconstruct the shape of the PSF at the position of each galaxy, the condition necessary for deconvolving it from the galaxies' shapes. Several stringent tests then validate our PSF model. In particular, we require that the ellipticity distribution, and the two point correlation functions of the ellipticity, of the residuals between observed stars and their shapelet models are consistent with zero. We also require that the cross-correlation between stars and galaxies ellipticity, when corrected from the PSF, is consistent with zero. The shape of galaxies is finally measured by decomposing them into shapelets, while deconvolving from the PSF, as shown in Massey & Refregier 2005.

[2] http://terapix.iap.fr

A shear estimator is created from the shapelet decompositions of galaxies, as prescribed by Massey et al. 2007d:

$$\gamma = \frac{f_{2,2}}{P^{\gamma}} \quad (1)$$

where the shear susceptibility $P^{\gamma} =< f_{0,0} - f_{4,0} > /\sqrt{2}$ is fitted on the magnitude-size plane for galaxies. The coefficients are complex numbers. Hence, the shear $\gamma$ of equation (1) is the complex notation for shears, $\gamma = \gamma_1 + i\gamma_2$.

To increase the signal to noise ratio of our measurements, we give to each galaxy $g$ a minimum variance weight $w_g = (\sigma^2_{\varepsilon,g} + \sigma^2_{P_\gamma,g} + \sigma^2_{\rm int})^{-1}$, where $\sigma_{\varepsilon,g}$ is the error on shape measurement for galaxy $g$, $\sigma_{P_{\gamma,g}}$ the error on the measurement of its shear susceptibility, and $\sigma_{\rm int}$ is the intrinsic ellipticity dispersion, set to $\sigma_{\rm int} = 0.3$. Slightly changing $\sigma_{\rm int}$ would be equivalent to give more or less weight to our measurement errors, and would mostly affect the error bars in the shear measurement. The eventual peak detection would not be affected by such slight changes. This weighting scheme is equivalent to smoothly selecting the most useful galaxies for shear measurement. For instance, the faintest are down-weighted. It therefore provides us with effective densities of $n_{\rm eff} \approx 20$ and 9 useful galaxies per square arcminute, in the D1 and W1 images, respectively. Then, a direct inversion in Fourier space of the pixelised shear map allows us to infer a convergence (i.e. projected mass) map (Kaiser & Squires 1993) of the images. Structures in this mass map are extracted from the noise using a Gaussian filter. Figures 2 and 3 show the convergence maps that we inferred from our data, and are described in section 4.1. While constructing a convergence map, we also create a signal-to-noise map, the signal-to-noise ratio being defined as

$$\nu(x,y) = \frac{\kappa(x,y)}{\sigma_\kappa(x,y)}, \quad (2)$$

where $\kappa(x,y)$ is the convergence at the $(x,y)$ sky coordinates, and $\sigma_\kappa(x,y)$ its r.m.s error. The r.m.s error $\sigma_\kappa(x,y)$ is computed using Monte-Carlo simulations in which the input galaxies are positioned like the observed ones but with randomised shape orientations.

Structures are then searched for in the filtered convergence map, and their astrometry provided, by the SExtractor software (Bertin & Arnouts 1996). They are extracted according to their signal-to-noise peak, read from the signal-to-noise map. Hereafter, we define a 'significant structure' as a detection with a signal-to-noise ratio greater than 2.5. Their mass is related to their integrated convergence through the lensing geometry, and can be estimated when their redshift and the redshift distribution of background galaxies are known. To account for the latter, we use the normalised distribution

$$n(z) = \frac{\beta}{z_s \Gamma(\frac{1+\alpha}{\beta})} \left(\frac{z}{z_s}\right)^{\alpha} \exp\left[-\left(\frac{z}{z_s}\right)^{\beta}\right] \quad (3)$$

where the parameters $\alpha, \beta, z_s$ are given for the Wide images by Benjamin et al. (2007) $(\alpha, \beta, z_s) = (0.836, 3.425, 1.171)$. To account for $n(z)$ in the D1 image, we fit Ilbert et al. (2006)'s photometric redshift distribution in the CFHTLS D1 field, and obtain $(\alpha, \beta, z_s) = (0.828, 1.859, 1.148)$. van Waerbeke et al. (2006) have shown that errors in the $n(z)$ fit are subdominant compared to Poisson noise and sample variance for the measurement of cosmological parameters. We thus neglect them hereafter.

We measure a cluster's virial mass by averaging its convergence in an aperture large enough that we can assume that the entire cluster is captured. The aperture corresponds to the region enclosed in the $2\sigma$ level of the cluster. This technique is similar to using a $\zeta$-statistic (Fahlman et al. 1994), with infinitely large annulus around the cluster, provided that the convergence in the entire field averages to 0. We verified this latter point, thus validating our choice. Note that because of the small number density of background sources, we cannot reliably fit a shear profile around clusters (see e.g. Paulin-Henriksson et al. (2007) for an example of mass estimation using two profile fits around the galaxy cluster Abell 209). We then convert the virial mass into $M_{200,c}$, the mass enclosed in the sphere of mean overdensity 200 times higher than the critical density, using the recipe from Hu & Kravtsov (2003). Hereafter, we will note $M_{200,c}$ more simply $M_{200}$.

Weak lensing is affected by the entire mass distribution along the line of sight. As a result, the weak lensing mass measurement of one cluster is biased by projection effects. It has been shown, using different mass estimators, that large-scale structures in the line of sight, and near the target cluster, introduce errors ranging from a few percent (Reblinsky & Bartelmann 1999; Hoekstra 2001, 2003; Clowe, De Lucia & King 2004) to a few tens percent (Metzler, White & Loken 2001; de Putter & White 2005). In this paper, we assume that they produce a 20% error, added in quadrature to the shear measurement error.

### 3.2 X-ray cluster detection and analysis

The X-ray cluster detection pipeline has been described in Pacaud et al. (2006). It takes account of the Poisson nature of the X-ray images, to extract and analyse clusters of galaxies in a two-step procedure. Clusters are first detected by a multi-resolution wavelet filter (Starck et al. 1998). Then, each source is analysed using a maximum likelihood profile fitting procedure, and its X-ray properties assessed. Three classes of extended sources have been defined (Pacaud et al. 2006; Pierre et al. 2006) : (1) the C1 class contains the highest surface brightness sources, and is uncontaminated ; (2) the C2 class allows for 50 % contamination, and contains less bright extended sources ; (3) finally, the C3 class contains optically confirmed sources with apparent X-ray emission, that were not selected as C1 or C2. In this paper, we only consider C1 class detections, representative of the most massive objects seen in the XMM-LSS. The redshift of detected clusters has been measured using spectroscopic observations from a number of telescope and instrument combinations detailed in Table 2 of Pierre et al. (2006). Their temperature estimation is described in Willis et al. (2005).

Pacaud et al. (2007) have extracted and analysed 29 C1 clusters from 5 deg$^2$ of the XMM-LSS data (shown on Fig. 1), that contain our 4 deg$^2$ optical data. Among other things, they have measured their luminosity and temperature. Here, we take into account their 16 clusters which are enclosed in the fields of our CFHTLS data, making use only of their temperature and redshift. They are listed in Table 1. Note that the cluster XLSSC053 is in the G12 XMM-LSS pointing (shown by the cross in the D1 field, on Fig. 1), which was

not used when Pacaud et al. (2007) analysed the XMM-LSS observations. This pointing has been re-observed, and the X-ray characteristics of XLSSC053 are listed in Table 1.

## 4 RESULTS

In this section, we give the properties of clusters of galaxies detected with our weak lensing pipeline. Counting the weak lensing detections allows us to constrain the matter power spectrum normalisation $\sigma_8$. We then use the weak lensing mass of the detected groups to calibrate the mass-temperature relation for clusters of galaxies.

### 4.1 Convergence maps and cluster catalogue

Figure 2 shows the 4 deg$^2$ of the W1 field that we considered. No significant overdensity (i.e. with $\nu \geqslant 2.5$) has been detected. As we will quantitatively show in section 4.2.1, this is consistent with the expected cluster counts for this survey. The black square in the image shows the position of the D1 field. Since there are around 20 galaxies per arcmin$^2$, we expect more significant detections in this deep field (see section 4.2.1). The upper panel of Fig. 3 shows the convergence map that we obtained from our weak lensing analysis of the D1 field. Due to the varying level of noise in our map, which varies independently of $\kappa$, two peaks with the same convergence $\kappa$ value do not necessarily have the same significance. That results in the rejection of seemingly significant structures, such as the peak around $(\alpha, \delta)$=(36.75°, −4.75°). Significant structures are marked out by the white contours, which start at $2.2\sigma$, with an increment of $0.5\sigma$. The lower panels show individual candidate clusters in more details. In these, contours start at $2.2\sigma$, with an increment of $0.2\sigma$. Even though we consider as significant a structure with $\nu \geqslant 2.5$, we plot the $2.2\sigma$ contours as a way to show the extension of our detections. Dashed circles mark X-ray clusters, and dotted circles GS07's weak lensing detections. GS07 measured the shear in the D1 field using the KSB method (Kaiser, Squires & Broadhurst 1995).

Table 1 summarises the measured characteristics of the clusters that we detect, together with all X-ray and GS07's detections in the region. The clusters that we detect through our shapelets weak lensing analysis are labelled with WL*id*, where *id* runs from 00 to 07, and are sorted by decreasing significance. Their labels are listed in the first column. Their official *XMM* names are given in column (2), and GS07's IDs in the third column. X-ray clusters marked by a $^\dagger$ are outside the D1 field. Columns (4) and (5) give their position. Column (6) lists their spectroscopic redshifts, except for clusters WL03, for which tomographic redshift is given, and WL04 for which photometric redshift is given (Aussel et al. in prep). The significances of the weak lensing detections are listed in column (7), in D1 and W1. A '-' means that the cluster is not detected ; a 'X' means that the cluster is outside the D1 field. Columns (8) and (9) give their weak lensing mass $M_{200}$(WL) and X-ray temperature, respectively. Column (10) gives the mass estimate from the X-ray data, $M_{200}(X)$. As in Pacaud et al. (2007), these were evaluated under the assumption of an isothermal $\beta$-model gas distribution in hydrostatic equilibrium with the cluster's potential well. In this earlier work, the associated statistical errors were generally dominated by the temperature measurement uncertainty ($\delta M/M \approx \delta T/T$ of order 10-25%). Here, the error on the emission profile can also become quite significant because we estimate the masses within $R_{200}$ instead of $R_{500}$ where the X-ray emission starts to vanish. Moreover, it was shown by Vikhlinin et al. (2006) and Rasia et al. (2006) that the isothermal $\beta$-model assumption leads to an underestimation of the total mass, by up to 40% for low mass systems. For these reasons, we decided not to provide error bars for our X-ray masses. Finally, column (11) gives some details about weak lensing detections, explaining for instance why we chose not to take them into account, or why we do not detect a cluster seen by another method. Among the rejection criteria are the proximity to an edge or to a masked region, the mass inversion procedure being sensitive to missing data and to edge effects. In that sense, and for clarity, the detections closest to edges have been removed from Fig. 2 and Fig. 3.

Clusters WL01, WL04 and WL05 have univocal counterparts both in our X-ray catalogue and in GS07's KSB one. No significant C1 X-ray source has been selected around WL02, and it remains invisible to GS07. Moreover, a visual inspection of the optical images does not show any galaxy overdensity around it. No significant C1 X-ray source has been detected at the position of WL03, even though it is also seen by GS07. Cluster WL06 lies just below our detection threshold ($\nu = 2.48$). Nevertheless, since GS07 detect it (though just below the detection threshold they use for their analyses) and since it is also detected by our X-ray analysis, we decided to list it, and to assess its weak lensing characteristics. The significance of cluster WL07 is even lower ($\nu = 2.42$). Since it is found close to the XLSSC022 cluster (which coincides with GS07's Cl07), we show its contours in Fig. 3 and list it in Table 1. However, it lies near an edge and a mask, so that its weak lensing characteristics are likely to be biased. We thus do not measure its mass, and will not take it into account in what follows. Cluster XLSSC025 (GS07's Cl05) is under a mask, and cannot be detected by our pipeline.

In summary, out of our 7 shapelet weak lensing detections, 4 (WL01, WL04, WL05 and WL07) have a counterpart both in our X-ray catalogue and in the KSB weak lensing catalogue by GS07, even though we remove WL07 from our subsequent analyses. One detection (WL02) appears only in our catalogue. One (WL06) has an X-ray counterpart, and appears in GS07's map, but just below the detection threshold they use for their analysis. Finally, one (WL03) has a counterpart in GS07's catalogue, but is not selected as a C1 X-ray cluster. This proves a good agreement between the three clusters detection methods used in those observations. X-ray clusters XLSSC005 and XLSSC029 are at too high a redshift to be detected with our surveys. Cluster XLSSC011 is too close and not massive enough to be detected by a weak lensing method, as will be shown in section 4.2.1.

Figure 4 compares the significance of our detections with that given by GS07. We make use of WL06, even though GS07 did not use it, but gave its significance. Given our weak lensing measurement characteristics ($n_{\rm eff}$ = 19 arcmin$^{-2}$, $\sigma_{\rm int} = 0.3$) and theirs ($n_{\rm eff} = 25.3$ arcmin$^{-2}$, $\sigma_{\rm int} = 0.23$), Eq. (4) below allows us to compute the expected proportionality factor between our detections' sig-

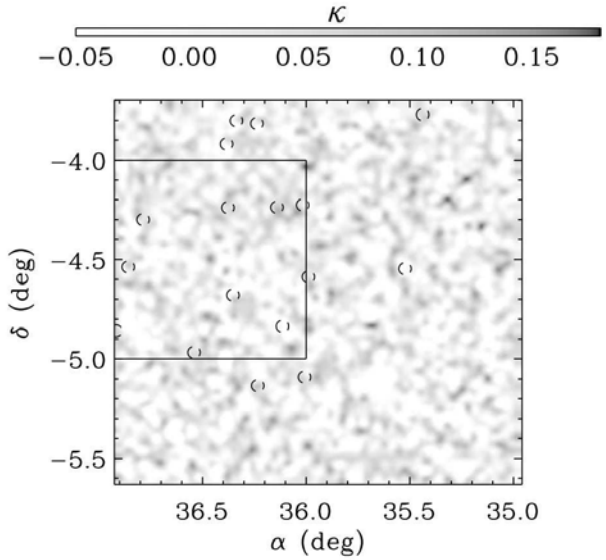


**Figure 2.** Convergence map inferred from our weak lensing measurement of the W1 field. The square in the W1 map shows the boundaries of the D1 field (Fig. 3). The map is smoothed by a 2.3' FWHM Gaussian. Dashed circles mark C1 X-ray clusters.

nificance $\nu_{\rm shapelets}$ and theirs, $\nu_{\rm GS07}$. We expect $\nu_{\rm GS07} = 1.47\ \nu_{\rm shapelets}$. This relation is shown by the dashed line on Fig. 4. The significance of clusters in both catalogues scale as expected.

While X-ray masses listed in column (10) of Table 1 must be taken with caution, they can be compared to the weak lensing masses listed in column (8). Although one can notice an order of magnitude agreement between $M_{200}$(WL) and $M_{200}$(X), masses estimated from X-ray data seem slightly underestimated. This is consistent with the previously mentioned bias arising from the isothermal beta-model parametrisation.

While they do not provide strong statistics, our detections can be used to estimate $\sigma_8$, as shown below. Four detections have an X-ray counterpart and can thus be used to constrain the mass-temperature relation, provided that we add clusters from another catalogue. This is described below.

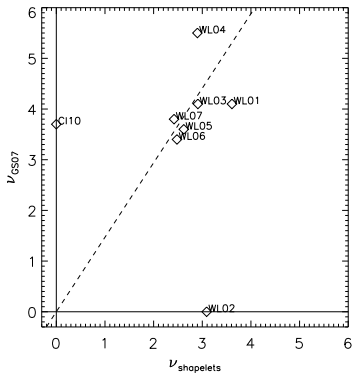


**Figure 4.** Comparison between our detection significances and those of GS07. Symbols are data points. Cl10 is detected by GS07, but remains invisible to our pipeline. The dashed line features the expected relation between clusters' significance in both analyses, $\nu_{\rm GS07} = 1.47\ \nu_{\rm shapelets}$.

### 4.2 Clusters number counts

#### *4.2.1 Weak lensing selection function*

A weak gravitational lensing selection function can be computed analytically (see e.g. Hamana, Takada & Yoshida 2004; Marian & Bernstein 2006) from the signal-to-noise ratio of a halo in a given cosmology and weak lensing survey parameters. We derive such a selection function, using an optimal match filter, in Bergé, Amara & Réfrégier (in prep). In an observation characterised by a number density of background galaxies $n_g$, an NFW halo of convergence $\kappa$ has signal-to-noise ratio :

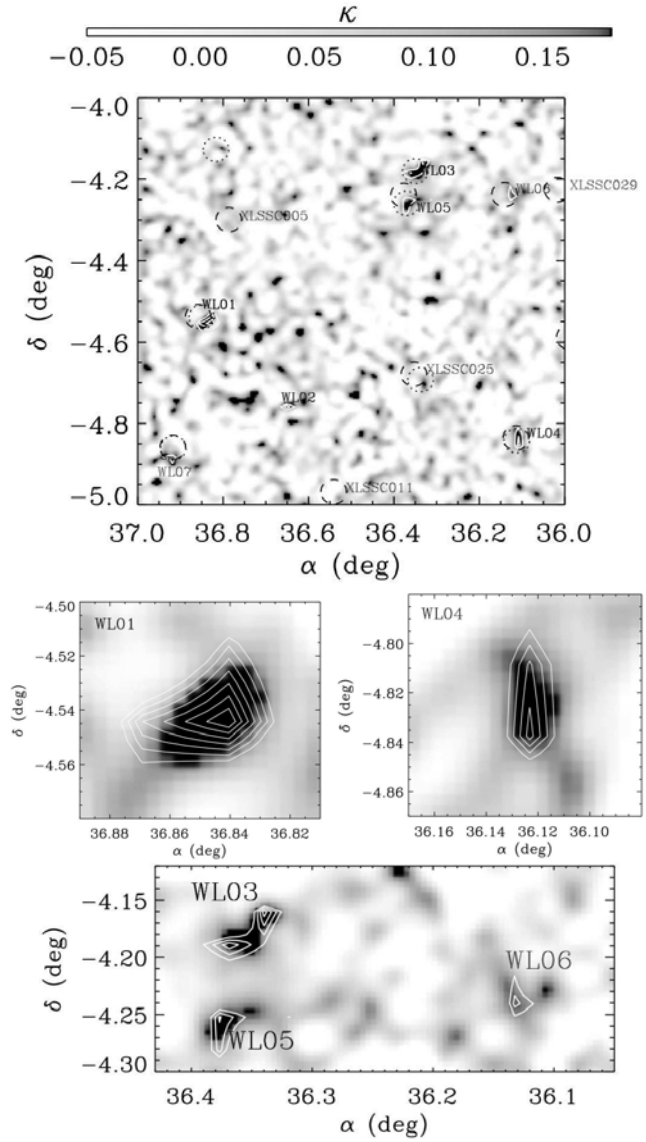


**Figure 3.** Convergence map inferred from our weak lensing measurement of the D1 field. The convergence $\kappa$ is shown for the entire field on the upper panel. Lower panels show zooms on cluster candidates WL01, WL04, and the region surrounding WL03, WL05 and WL06. In the upper panel, contours levels start at $2.2\sigma$, with an increment of $0.5\sigma$. In the lower panels, they start at $2.2\sigma$, with an increment of $0.2\sigma$. The maps are smoothed by a 1.1' FWHM Gaussian. In the upper panel, dashed circles mark X-ray clusters, and dotted circles show GS07's KSB weak lensing detections. Clusters detected by our shapelets weak lensing measurement are labelled WL*id*, and X-ray clusters not detected by weak lensing are labelled by their *XMM* name. All clusters are listed in Table 1. For clarity, false detections near edges are not shown.

**Table 1.** Clusters catalogue. Besides the clusters that we detect through our shapelet weak lensing measurement, we also list clusters seen by GS07, and selected as C1 clusters in the XMM-LSS. Weak lensing detections' significance is given for D1 and W1, even if no detection appears in W1. $M_{200}$(WL) is the cluster weak lensing mass. $M_{200}$(X) is the mass estimated from X-ray profile, extrapolated from $R_{500}$ to $R_{200}$, with respect to $M_{500}$ masses given by Pacaud et al. (2007), and must be used with caution (see text).

| Weak lensing cluster ID | XLSSC number | GS07 ID | RA (deg) | Dec (deg) | $z$ | Significance D1/W1 | $M_{200}$(WL) ($10^{13}h^{-1}M_\odot$) | $T_X$ (keV) | $M_{200}$(X)[c] ($10^{13}h^{-1}M_\odot$) | Notes |
|---|---|---|---|---|---|---|---|---|---|---|
| WL01 | 013 | Cl03 | 36.8497 | -4.5481 | 0.31 | 3.61 / - | $8.2^{+2.5}_{-1.9}$ | $1.0^{+0.1}_{-0.1}$ | 2.1 | |
| WL02 | - | - | 36.6589 | -4.7516 | - | 3.09 / - | - | - | - | |
| WL03 | - | Cl04 | 36.3628 | -4.1886 | 0.32 [a] | 2.91 / - | $8.9^{+2.6}_{-2.2}$ | - | - | |
| WL04 | 053 | Cl02 | 36.1229 | -4.8341 | 0.50 [b] | 2.90 / - | $10.3^{+3.0}_{-2.6}$ | $3.4^{+3.1}_{-1.0}$ | 5.0 | XMM-LSS pointing not observed in Pacaud et al. (2007) |
| WL05 | 041 | Cl14 | 36.3723 | -4.2604 | 0.14 | 2.62 / - | $4.9^{+1.6}_{-1.2}$ | $1.3^{+0.1}_{-0.1}$ | 3.5 | |
| WL06 | 044 | - | 36.1389 | -4.2384 | 0.26 | 2.48 / - | $7.2^{+2.3}_{-1.7}$ | $1.3^{+0.1}_{-0.2}$ | 3.7 | just below detection threshold in GS07's catalogue |
| WL07 [d] | 022 | Cl07 | 36.9167 | -4.8606 | 0.29 | 2.42 / - | - | $1.7^{+0.1}_{-0.1}$ | 5.3 | near a mask |
| - | 025 | Cl05 | 36.3375 | -4.6925 | 0.26 | - / - | - | $2.0^{+0.2}_{-0.2}$ | 6.5 | under a mask |
| - | - | Cl10 | 36.8167 | -4.1269 | - | - / - | - | - | | |
| - | 029 | - | 36.0172 | -4.2260 | 1.05 | - / - | - | $4.1^{+0.9}_{-0.7}$ | 13.9 | too high redshift |
| - | 011 | - | 36.5410 | -4.9680 | 0.05 | - / - | - | $0.64^{+0.06}_{-0.04}$ | 1.0 | |
| - | 005 | - | 36.7866 | -4.2995 | 1.05 | - / - | - | $3.7^{+1.7}_{-1.}$ | 16.5 | too high redshift |
| - | 006† | - | 35.4382 | -3.7717 | 0.43 | X / - | - | $4.8^{+0.6}_{-0.5}$ | 30.4 | near an edge |
| - | 040† | - | 35.5232 | -4.5463 | 0.32 | X / - | - | $1.6^{+1.1}_{-0.3}$ | 6.8 | |
| - | 049† | - | 35.9892 | -4.5880 | 0.49 | X / - | - | $2.2^{+0.9}_{-0.5}$ | 5.0 | |
| - | 018† | - | 36.0079 | -5.0903 | 0.32 | X / - | - | $2.0^{+0.7}_{-0.4}$ | 8.0 | |
| - | 021† | - | 36.2338 | -5.1340 | 0.08 | X / - | - | $0.68^{+0.04}_{-0.02}$ | 1.8 | |
| - | 001† | - | 36.2378 | -3.8156 | 0.61 | X / - | - | $3.2^{+0.4}_{-0.3}$ | 14.3 | |
| - | 008† | - | 36.3367 | -3.8014 | 0.30 | X / - | - | $1.3^{+0.7}_{-0.2}$ | 2.1 | |
| - | 002† | - | 36.3841 | -3.9198 | 0.77 | X / - | - | $2.8^{+0.8}_{-0.5}$ | 9.6 | |

[a] Tomographic redshift (GS07)
[b] Photometric redshift (Aussel et al. in prep)
[c] Rough estimates based on the isothermal assumption and extrapolated from $M_{500}$ given by Pacaud et al. (2007).
[d] X-ray coordinates.
† Outside D1.

$$\nu = \frac{\sqrt{n_g}}{\sigma_\gamma}\sqrt{\int \mathrm{d}^2x\,\kappa^2(x)} \tag{4}$$

where $\sigma_\gamma$ is the r.m.s shear error per galaxy, and where we neglect projection effects and sample variance, which have been shown to have subdominant effects (Marian & Bernstein 2006).

Our selection function is shown in Figure 5 in the mass-redshift plane, for our Deep ($n_g = 20$ arcmin$^{-2}$, $\sigma_\gamma = 0.3$, thick black) and our Wide ($n_g = 9$ arcmin$^{-2}$, $\sigma_\gamma = 0.4$, red) surveys, in a cosmological model based on the three-year *Wilkinson Microwave Anisotropy Probe* results (WMAP3; Spergel et al. 2007), $(h, \Omega_m h^2, \Omega_b h^2, \sigma_8, w)$=(0.73,0.127,0.0223,0.76,-1). We use the redshift distributions for background galaxies given by equation (3).

Figure 5 shows, from bottom to top, the minimum detectable mass for a halo at a $2\sigma$, a $3\sigma$ and a $4\sigma$ detection threshold. The Deep and Wide selection functions have different slopes, as a consequence of their different redshift distribution for background galaxies. Symbols represent the position, in the redshift-mass plane, of clusters listed in Table 1. We use the weak lensing mass $M_{200}$(WL) for WL01, WL03, WL04, WL05 and WL06 (thick square symbols) and the X-ray mass $M_{200}$(X) for other clusters (diamonds). Although we detect cluster XLSSC022 (WL07), we do not assess its gravitational mass, and thus show its X-ray mass in Fig. 5. Cluster XLSSC025 should be detectable (and is detected by GS07), but it is under a mask in our analysis. Triangle symbols (labeled † in Table 1) correspond to C1 X-ray clusters in W1 that are outside the D1 region. It is clear from Figure 5 that they can not be detected by our weak lensing analysis of W1. Only XLSSC006 should be seen, at the $2\sigma$ level. However, its detection is plagued by its proximity to the edge of the image.

Also shown on Fig. 5 are the X-ray selection functions, for 50% and 80% detection probabilities (dashed lines) as estimated by Pacaud et al. (2007).

Figure 5 shows an excellent agreement between the clusters characteristics and their predicted detectability by weak gravitational lensing.

#### *4.2.2 Number counts*

From equation (4), the expected number of haloes detected above a certain significance can be computed using a Press-Schechter approach (Press & Schechter 1974). For this purpose, we use the Jenkins et al. (2001) mass function to es-

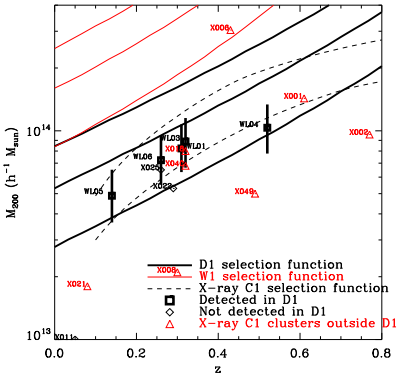


**Figure 5.** Weak lensing selection function for a survey like D1 (thick black ; $\sigma_{\rm int}$ = 0.3, $n_g$=20 arcmin$^{-1}$, redshift distribution as Eq. (3)) and W1 (red ; $\sigma_{\rm int}$ = 0.4, $n_g$=9 arcmin$^{-1}$, redshift distribution as Eq. (3)), assuming a WMAP3 cosmology in each case. From bottom to top, lines correspond to $2\sigma$, $3\sigma$ and $4\sigma$ significance. Dashed lines show the X-ray selection function, corresponding to 50%, and 80% detection probability (Pacaud et al. 2007 Fig. 18, lower and upper curves, respectively). Thick square symbols are our detections in the D1 data, labeled by their ID ; they are not detectable in the W1 data. Diamonds are clusters detected either by GS07 or by X-ray analysis, in D1, that we do not detect for reasons listed in the text. Red triangles are C1 X-ray clusters lying outside the D1 region. Except for the thick square symbols (for which we use the weak lensing mass $M_{200}$(WL)), we use the X-ray mass $M_{200}$(X). Except XLSSC006 (labeled for visibility as X006), they cannot be detected by a weak lensing experiment in the W1 data. XLSSC006 is not detected because of its proximity to an edge.

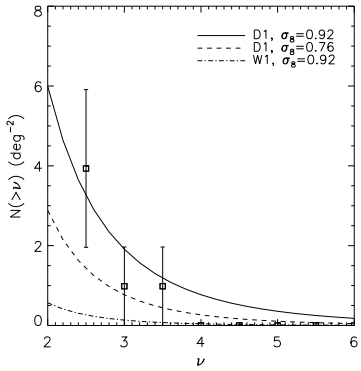


**Figure 6.** Cumulative cluster number density as a function of weak lensing detection significance, in the D1 data. The error bars include shot noise and sample variance. The dashed line shows the expected number counts in a WMAP3 cosmology, for the survey's characteristics. The solid line is our best fit, when varying $\sigma_8$ ($\sigma_8$ = 0.92). The dot-dashed line is the expected number counts for the Wide survey, with $\sigma_8$ = 0.92.

timate the number of haloes that we can detect, as a function of significance threshold. Curves on figure 6 show such counts for different $\sigma_8$ and survey depths. Miyazaki et al. (2002) already used this statistic to discriminate between halo profile models. It is used here to measure $\sigma_8$.

Most of our detections are validated by corresponding objects either in the catalogue of GS07 or our X-ray C1 clusters catalogue. Nevertheless, despite its relative high significance, WL02 does not have such independent support. We indeed consider it as a false detection, and do not take it into account for cluster counts. We then estimate the number of false detections from Monte Carlo simulations. For this purpose, we input galaxies at the position of the actual ones, but randomise their shear, and look for detections with significance higher than $2.5\sigma$. The convergence maps that we infer from them show only false detections. We find that, in this particular experiment, we expect only one false detection above the $2.5\sigma$ level. This is thus consistent with removing WL02. For this counting purpose, we remove WL06 and WL07 from our catalogue, since they do not reach the $2.5\sigma$ level. The symbols on Fig. 6 represent our cumulative counts, corrected from false detections. Their error bars include the effects of shot noise and sample variance, computed from Hu & Kravtsov's (2003) analytic formula. We then fit the expected number counts to our data as a function of $\sigma_8$, keeping all other parameters constant. In order to avoid covariance between our data points in the cumulative counts depicted by Fig. 6, we performed the fit on the expected differential number counts $\mathrm{d}N/\mathrm{d}\nu(\nu)$. We find $\sigma_8 = 0.92^{+0.26}_{-0.30}$ (at the 68.3% confidence limit), for $\Omega_m = 0.24$. Despite large error bars, we can set interesting constraints thanks to the strong dependence of these counts on $\sigma_8$, as shown by the difference between the solid and dashed curves in Figure 6. The dashed line shows the expected cumulative number counts for the Deep survey in a WMAP3-like universe ($\sigma_8 = 0.76$). The solid line is our best fit ($\sigma_8 = 0.92$). The dot-dashed line shows the expected number density on the Wide survey, with $\sigma_8 = 0.92$.

### 4.3 Mass-temperature relation

Under the virial equilibrium assumption, the mass and temperature of a cluster are related by the following scaling relation (Pierpaoli et al. 2003)

$$\frac{M_{\rm vir}(T,z)}{10^{15}h^{-1}M_\odot} = \left(\frac{T}{T_*}\right)^{3/2}\left[\Delta_c(z)E(z)^2\right]^{-1/2}\left[1-2\frac{\Omega_\Lambda(z)}{\Delta_c(z)}\right]^{-3/2} \quad (5)$$

where $M_{\rm vir}$ is the virial mass, $T$ is the virial temperature, $T_*$ is a normalisation factor, and $E(z)^2 = \Omega_m(1+z)^3 + \Omega_\Lambda + \Omega_k(1+z)^2$. $\Delta_c(z)$ is the overdensity inside the virial radius, in units of the critical density. We compute it using the fitting formula by Weinberg & Kamionkowski (2003) for $\Delta_{\rm vir} = \Delta_c/\Omega_m$, which is very similar to an earlier approximation by Nakamura (1996) given by Kitayama & Suto (1996) for a universe with arbitrary $\Omega_m$.

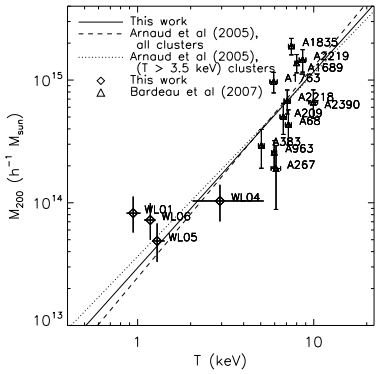


**Figure 7.** Mass-temperature relation, normalised to $z = 0$, for our group sample (diamonds). To increase our statistic, we added clusters from Bardeau et al. 2007 (triangles). We use X-ray temperature and weak lensing mass. The solid line is our best fit (Eq. 7). The dashed and dotted line are APP05 M-T relation, when they consider all clusters or only those with $T > 3.5\,\mathrm{keV}$, respectively.

A more general relation often used to fit observations makes use of a related normalisation factor $M_*$ and is given, at redshift $z = 0$, by

$$M_{200} \approx M_* \left(\frac{T}{4\ \mathrm{keV}}\right)^{\alpha} \tag{6}$$

where $M_{200}$ is the mass inside the sphere of mean overdensity 200 times higher than the critical density and $\alpha = 3/2$ in the hydrostatic equilibrium assumption (e.g. APP05). Hereafter, to account for redshift evolution, we normalise all our temperatures to $z = 0$ by dividing them by $E(z)^{2/3}$.

Measuring $\sigma_8$ from X-ray counts is affected by the degeneracy $\Omega_m^{0.6}\sigma_8 \propto T_*^{-0.8}$ (Pierpaoli et al. 2003). Pierpaoli et al. (2003) have shown that the uncertainty in $M_*$ is the main concern in measurements of $\sigma_8$ from X-ray cluster observations alone. Such data is limited by the requirement that the cluster masses be inferred from the X-ray profiles. Smith et al. (2003) have also shown that unrelaxed clusters, being hotter than relaxed clusters, provide a supplementary bias to the $\sigma_8$ estimate. It is thus important to have a mass estimate independent of the hydrostatic equilibrium assumption. Weak gravitational lensing gives such an estimate. Combined with X-ray temperature, it can be efficiently used to constraint the M-T relation, independently of the cluster physical state. Hjorth, Oukbir & van Kampen (1998); Pedersen & Dahle (2006); Bardeau et al. (2007) have already used it to measure the M-T relation normalisation.

As described above, we have the weak lensing mass and X-ray temperature of only four groups. Hence, to increase our statistics, we add Bardeau et al. (2007)'s clusters to our catalogue, providing us with 11 additional clusters. Bardeau et al. (2007) estimated cluster masses by fitting an NFW model to their tangential shears. We should note here that since Bardeau et al. (2007)'s mass estimation and ours are based on different techniques, our subsequent analysis of the mass - temperature relation could be slightly biased due to possible calibration differences. Figure 7 shows the relation between the temperature and the weak lensing mass $M_{200}$ for the combined catalogues. Diamonds are our groups, labelled WL*id*, triangles are Bardeau et al. (2007)'s clusters, labelled A*id*. Bardeau et al. (2007) proceeded to the weak lensing analyses of massive haloes, the temperature of which were obtained by Zhang et al. (2007) and Ota & Mitsuda (2004) ; particularly, they estimated their weak lensing mass and measured the scale relations for those clusters. They fitted their sample by varying both $\alpha$ and $M_*$, and found a large slope, far from the hydrostatic equilibrium assumption, $\alpha = 4.6 \pm 0.7$. Doing the same analysis on the larger range in mass that the addition of both catalogues probes, from galaxy groups to galaxy clusters, we find :

$$\frac{M_{200}}{10^{14}h^{-1}M_{\odot}} = 2.71^{+0.79}_{-0.61}\left(\frac{T}{4\ \mathrm{keV}}\right)^{1.60\pm0.44}, \tag{7}$$

which is in good agreement with APP05 (whether they use all clusters or only the most massive ones), Bardeau et al. (2007) (when they fix $\alpha = 1.5$) or Pedersen & Dahle (2006). The solid line on Fig. 7 is our best fit. The dashed line is the best fit from APP05, when they consider all clusters in their catalogue. The dotted line is their best fit when they consider their most massive clusters ($T > 3.5$ keV).

Our result is consistent with self-similarity evolution for galaxy clusters down to low temperatures. It is also consistent with previous measurements which observe a steepening of the M-T relation at the low mass end, due to the expected self-similarity breaking for such masses (e.g. Nevalainen et al. 2000; Finoguenov et al. 2001; APP05). Moreover, one must be aware that the galaxy groups we consider were detected just above our weak lensing selection function (Fig. 5). Due to the expected scatter in the M-T relation, those groups can represent only the most massive ones with temperature ranging about 1 keV. Our group sample could thus bias our fit towards a flat slope for the M-T relation. The analysis of more low temperature groups will be needed to further explore this issue.

## 5 DISCUSSION

The power spectrum normalisation $\sigma_8$ has been measured with different probes, such as X-ray clusters of galaxies, CMB, and cosmic shear (i.e. statistics of weak gravitational lensing). Some discrepancies have emerged between the preferred value from those measurements. Recent CMB observations favour a low $\sigma_8$ and cosmic shear used to emphasise a high value (see e.g. Refregier 2003a for a review). X-ray clusters provide intermediate measurements (see e.g. Pierpaoli et al. 2003 for a review). The dominant discrepancy between cosmic shear and X-ray clusters has recently been reduced by Jarvis et al. (2006), who measured $\sigma_8 \approx 0.81$ for $\Omega_m = 0.26$ when using cosmic shear alone, followed by Benjamin et al. (2007), who used Ilbert et al. (2006)'s improved galaxies photometric redshifts, and measured $\sigma_8 = 0.84$ for $\Omega_m = 0.24$. To clarify these discrepancies, one needs to measure both the power spectrum and the M-T relation normalisations $\sigma_8$ and $M_*$, as we discuss here.

The regions allowed for by different measurements on

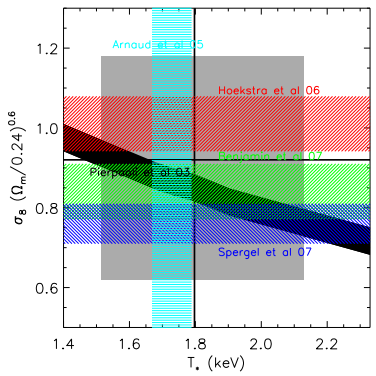


**Figure 8.** Domain allowed for by different measurements, in the $T_*$ - $\sigma_8(\Omega_m/0.24)^{0.6}$ plane. The shaded region shows the constraints given by our $\sigma_8$ and $T_*$ measurements. Thick lines are our best fits. The slanted black region correspond to the $1\sigma$ constraints on the $\Omega_m^{0.6}\sigma_8 \propto T_*^{-0.8}$ relation from Pierpaoli et al. 2003. The vertical, light blue, shaded region shows the $1\sigma$ error on $T_*$ from APP05. CMB derived constraints of $\sigma_8$ (Spergel et al. 2007) are shown by the horizontal dark blue shaded region. Cosmic shear $\sigma_8$ estimations from Hoekstra et al. 2006 and Benjamin et al. 2007 are marked by the red and green horizontal shaded regions. Hoekstra et al. 2006's constraints are typical of cosmic shear results. They are higher than X-ray estimations, marked by the intersection between Pierpaoli et al. 2003's and APP05's allowed domains.

the $T_*$ - $\sigma_8(\Omega_m/0.24)^{0.6}$ plane are shown by Figure 8. The shaded region shows the constraints given by our $\sigma_8$ and $T_*$ measurements. Our best fits are shown by the thick lines. The slanted black band on the figure is the 68.3% bound on the $\Omega_m^{0.6}\sigma_8 \propto T_*^{-0.8}$ relation from Pierpaoli et al. (2003) using X-ray clusters. Its intersection with the vertical light blue band (APP05's $T_*$ estimation) gives the current value for $\sigma_8$ favoured by X-ray cluster observations, $\sigma_8 \approx 0.77 \pm 0.06$ for an $\Omega_m = 0.3$ universe (e.g. Pierpaoli et al. 2003), which corresponds to $\sigma_8(\Omega_m/0.24)^{0.6} \approx 0.88 \pm 0.05$. This value is higher than that measured by Spergel et al. (2007) from CMB analyses of WMAP3 (dark blue), but lower than most cosmic shear analysis, like that of Hoekstra et al. (2006) made with CFHTLS Wide data (red). This highlights the discrepancy between X-ray and weak lensing estimates of $\sigma_8$ mentioned above. However, Benjamin et al. (2007) give a lower estimate for $\sigma_8$, consistent with X-ray measurements (green). This could be a sign that other cosmic shear analyses did not take some systematics into account, and have thus overestimated $\sigma_8$. According to Benjamin et al. (2007), previously published analyses made use of insufficiently-accurate galaxy photometric redshifts. Using Ilbert et al. (2006)'s redshifts yielded a lower value of $\sigma_8$ both for cosmic shear (Benjamin et al. 2007) and for our cluster count analysis. We found a 5% decrease in our $\sigma_8$ estimation when going from previous redshift distributions to Ilbert et al. (2006)'s ones. This is less than the change reported by Benjamin et al. (2007), and our best fit still tends to favour a higher value for $\sigma_8$, but is limited by low statistics. Smith et al. (2003) have analysed the bias from unrelaxed clusters in $\sigma_8$ measurement using lensing clusters and the M-T relation. They found that unrelaxed clusters are 30% hotter than relaxed clusters : using unrelaxed clusters can provide 20% overestimates of $\sigma_8$. This is enough to explain the large range of measured $\sigma_8$, from $\approx 0.6$ to $\approx 1$. They estimated $\sigma_8(\Omega_m/0.24)^{0.6} = 0.86 \pm 0.23$. Estimates from X-ray alone can also be affected by systematics, such as the mass estimate from X-ray profiles of clusters. For example, a slight decrease of $T_*$ would cause an increase of the X-ray estimate for $\sigma_8$. A better insight into this will come from an accurate measurement of $T_*$, preferably with mass estimation methods independent of cluster physics. Large combined weak lensing and X-ray surveys will be needed to disentangle the situation. They will provide both independent constraints on $\sigma_8$, and insights on $T_*$.

## 6 PROSPECTS FOR FUTURE SURVEYS

In the following, we investigate the impact of future combined blind weak lensing and X-ray surveys on the measurement precision of the power spectrum and the mass-temperature relation normalisations. We take the WMAP3 (Spergel et al. 2007) cosmology as our fiducial model. We consider two different ground based survey strategies for our weak gravitational lensing analysis : deep and wide surveys similar to the CFHTLS Deep and Wide surveys. We use their observed weighted number density of useful background galaxies to be $n_g = 20$ arcmin$^{-2}$ and 9 arcmin$^{-2}$, respectively, distributed according to Eq. (3). We also assume the intrinsic ellipticity and shape measurement error to be $\sigma_\gamma = 0.3$ in both cases. Following the CFHTLS scheduling, we take for exposure times 40 hours per square degree for the deep survey and 1 hour per square degree for the wide survey.

### 6.1 $\sigma_8$ measurements

We first investigate the impact of future surveys on the $\sigma_8$ measurement. Using the Press-Schechter approach described in section 4.2.2, we estimate the number of weak lensing detections with significance higher than 2.5, taking into account shot noise and sample variance. We assume that all clusters have a spherically symmetrical NFW profile. We thus neglect the effect of haloes' asphericity shown by Clowe et al. (2004) : triaxial haloes oriented along the line of sight appear more massive than triaxial haloes of the same mass, but perpendicular to the line of sight, and thus have a higher signal-to-noise ratio. Clowe et al. (2004) have shown that this approximation does not yield any difference in the mass measurement dispersion. Figure 9 shows the 68.3% relative error on $\sigma_8$ that can be reached by counting weak lensing detected clusters as a function of their significance, for a deep (thick solid line) and a wide (thick dashed line) surveys, as a function of survey's size and observing time. Because of the higher number density of clusters it allows one to detect, a deep survey provides errors 3 times lower than a wide survey of the same size. However, for a given exposure time, a wide survey provides errors 2.1 times lower than a deep

one. That means that the gain due to the coverage (and detectable clusters number) increase is faster than the one due to depth increase. A larger coverage is also advantageous in that it makes sample variance fall down rapidly. Moreover, a wide survey detects the most massive haloes, the physics of which is better understood. Consequently, in a survey strategy driven by exposure time, one should prefer a wide survey. The flat dashed-dot line shows the current constraints provided by cosmic shear analyses (Hoekstra et al. 2006; Benjamin et al. 2007). The flat dashed-dot-dot line shows the current constraints from the combination of X-ray $M_*$ measurement (APP05) and X-ray cluster counts (Pierpaoli et al. 2003).

Detecting and counting clusters on a 10 deg$^2$ deep survey will be competitive with current cosmic shear measurements, whereas 20 deg$^2$ of coverage is needed to compete with current X-ray clusters measurements. Those figures transform as 100 deg$^2$ and 200 deg$^2$ for a wide survey. That is, to compete with current cosmic shear surveys, one needs 400 hours of deep survey exposure, or 100 hours of wide survey exposure. Double these times are required to compete with X-ray surveys. A wide survey, less demanding in exposure time than a deep one, should then be used. Counting clusters on the entire planned CFHTLS Wide Survey 170 deg$^2$ will provide a 6% fractional error on the $\sigma_8$ measurement. Reaching the 1% fractional error will require a 7000 deg$^2$ wide survey, or a 700 deg$^2$ deep survey. Future surveys (e.g. Pan-STARRS[3], LSST[4], DUNE[5]) will be able to achieve such errors.

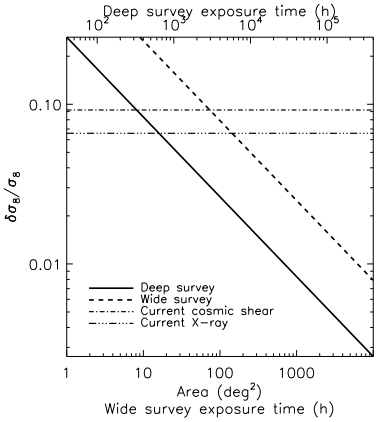


**Figure 9.** Relative errors on $\sigma_8$, from clusters counts in a weak gravitational lensing survey, as a function of survey size and integration time. All other parameters are kept constants. We assume that 1 deg$^2$ of wide requires 1 hour of observation time, and 1 deg$^2$ of deep requires 40 hours of observation time. That is, the lower x-axis shows area as well as the wide survey exposure time ; the upper x-axis shows the deep survey exposure time. The thick solid line corresponds to a deep survey, and the thick dashed line to a wide survey. The flat lines show the current error measurement from cosmic shear statistics (dash-dot, Hoekstra et al. 2006; Benjamin et al. 2007), and from X-ray clusters (dash-dot-dot, Pierpaoli et al. 2003; APP05).

### 6.2 $T_*$ measurements

We now turn to the precision that can be reached on $T_*$ measurements by future joint surveys. What matters here is not cluster counts as a function of significance, but as a function of mass. As we see from figure 5, low mass clusters of galaxies cannot be seen through weak gravitational lensing since they do not create high enough signal-to-noise ratios. A deep survey captures lower mass clusters than a wide survey, but deep and wide surveys give access to the same number of massive clusters. Therefore, a deep survey is naturally focused on the physics of galaxy groups (e.g. it can probe similarity breaking at the low mass end of the mass-temperature relation). A wide survey gives the same statistics on massive haloes, generally used to measure the mass-temperature relation normalisation : for this purpose, one should then choose a wide survey. To compare the merits of both deep and wide survey on the $T_*$ estimation's precision, we simulate M-T relations for both types of survey. We take a realistic scatter into account, $\sigma_{\log,\mathrm{int}} = 0.051$ for the logarithmic M-T relation (APP05). We assume that masses are measured through weak gravitational lensing. We measure $T_*$ and the error on its estimate, by assuming (1) that our cluster sample is complete, (2) that we only make use of those clusters detected in our blind survey, and (3) that we know the X-ray temperature of each of them. We also investigate the influence of the mass estimation fractional error $\delta M/M$.

Figure 10 shows the 68.3% error on $T_*$ that can be reached from a combined blind X-ray and deep (solid line) or wide (dashed line) weak gravitational lensing surveys, as a function of survey's size and integration time. Here again, as for the error on the power spectrum normalisation, a deep survey gives errors 2.3 times lower than a wide one with the same sky area coverage. On the other hand, a wide survey gives errors 2.7 times lower than a deep one with the same exposure time. The dependence on area underlines the $T_*$ estimation's reliance on the number of useable haloes for the M-T relation fitting. Figure 10 also shows the sensitivity of the $T_*$ estimation to the mass measurement errors. Black lines assume $\delta M/M = 0.2$, and the red ones $\delta M/M = 0.3$, which are the current fractional errors from weak lensing mass measurement from weak lensing. Going from $\delta M/M = 0.3$ to $\delta M/M = 0.2$ allows one to reduce the error on $T_*$ by a factor of 1.3 (resp. 1.2) for a deep (resp. wide) survey of a given sky area. The flat solid line represents the current error on $T_*$ from X-ray clusters (APP05). Assuming a 20% error measurement on weak lensing masses, one needs a 50 (resp. 300) deg$^2$ weak lensing deep (resp. wide) survey to reach the current error. Reaching the 1% fractional error (for our fiducial model with $T_* = 1.9$) will require a 2500 deg$^2$ wide survey, or a 500 deg$^2$ deep survey. Weak lensing surveys like LSST or DUNE combined with X-ray surveys like eROSITA will be able to reach such a limit.

[3] http://panstarrs.ifa.hawaii.edu
[4] http://www.lsst.org
[5] http://www.dune-mission.net

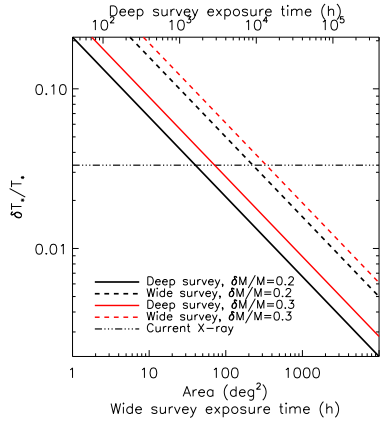


**Figure 10.** Relative errors on $T_*$ from combined weak gravitational lensing and X-ray surveys, as a function of survey size and integration time. All other parameters are kept constants. The lower x-axis shows area as well as the wide survey exposure time; the upper x-axis shows the deep survey exposure time. Solid lines show the errors for a deep survey, assuming the fractional error on weak lensing mass measurement is 20% (thick black) and 30% (red). Dashed lines, with the same colour indexing, show the errors for a wide survey. The current error measurement from X-ray clusters is shown by the flat dashed-dot-dot line (Pierpaoli et al. 2003; APP05). We made the same assumption about the relation between survey area and observation time as in Fig. 9.

In a survey strategy driven by exposure time, a wide survey of 2500 deg$^2$ (2500 hours) will be able to reach the 1% accuracy both on $\sigma_8$ and $T_*$, at a much cheaper expense than a deep survey. Nevertheless, a deep survey will still be useful to probe high redshift regions ($z \geqslant 0.8$), and to study low mass clusters of galaxies ($M \leqslant 10^{14} h^{-1} M_\odot$).

## 7 CONCLUSION

We have presented the first shapelet analysis of weak gravitational lensing surveys. We have constructed convergence maps of the CFHTLS Deep D1 field, and of 4 deg$^2$ of the CFTHLS Wide W1 field, which include the D1 field. We have detected six clusters of galaxies, through the lensing signal they generate. Our D1 map is in good agreement with that of GS07, precedently created using the KSB shear measurement method. We combined our weak lensing data with the X-ray analysis of XMM-LSS C1 clusters lying in the same region of the sky (Pacaud et al. 2007). These three clusters catalogue are consistent. All our shapelet detections have either an X-ray counterpart or a KSB detection. Counting our detections and accounting for the weak lensing selection function allowed us to constrain the power spectrum normalisation $\sigma_8(\Omega_m/0.24)^{0.6} = 0.92^{+0.26}_{-0.30}$. The combination of lensing masses and X-ray temperatures provided us with a new measurement of the mass-temperature relation normalisation $T_*$ (or equivalently $M_*$) for clusters of galaxies, $M_* = 2.71^{+0.79}_{-0.61}\, 10^{14} h^{-1} M_\odot$. Our results, though limited by low statistics and sample variance, are consistent with other current estimates. We also measured the slope of the mass-temperature relation, and found it consistent with self-similarity for low mass clusters, $\alpha = 1.60 \pm 0.44$. We have shown that one must measure both $\sigma_8$ and $T_*$ from combined weak lensing and X-ray surveys to investigate the discrepancy between independent measurements of $\sigma_8$ from different probes.

Weak lensing surveys are becoming more and more effective, and are currently being optimised for best extracting cosmological information. Optimal surveys will allow us to provide more accurate estimates of $\sigma_8$ and $T_*$, and to disentangle the current $\sigma_8$ issue (Amara & Refregier 2007). We have compared the merits of weak lensing deep and wide blind surveys, based on the CFHTLS, at estimating $\sigma_8$. We also looked at their merits at estimating $T_*$ while combined with an X-ray survey on their region of the sky. We found that for experiments driven by exposure time constraints, a wide survey will give $\approx 3$ times lower errors on the estimates of both $\sigma_8$ and $T_*$. To secure the measurement of $\sigma_8$ and $M_*$ with the current statistical accuracy, a 200 deg$^2$ and a 300 deg$^2$ wide surveys will be needed respectively. We finally found that a 7000 deg$^2$ wide survey will be able to reach the 1% accuracy both on the power spectrum and mass-temperature relation normalisations.

## ACKNOWLEDGEMENTS

The authors wish to thank Hervé Aussel, Krys Libbrecht, Jean-Baptiste Melin, Yannick Mellier, Sandrine Pires, Trevor Ponman, Jean-Luc Starck, Geneviève Soucail and Romain Teyssier for useful discussions, and Cathy Horellou for comments on the first version of the paper.

## REFERENCES

Amara A., Refregier A., 2007, MNRAS, 381, 1018
Arnaud M., Aghanim N., Neumann D.M., 2002, A&A, 389, 1
Arnaud M., Pointecouteau E., Pratt G.W., 2005, A&A, 441, 893 (APP05)
Bacon D., Refregier A., Ellis R., 2000, MNRAS, 318, 625
Bacon D., Massey R., Refregier A., Ellis R., 2003, MNRAS, 344,673
Bardeau S., Soucail G., Kneib J.P., Czoske O., Ebeling H., Hudelot P., Smail I., Smith G.P., , 2007, A&A, 470, 449
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Basilakos S., 2003, ApJ, 590, 636
Basilakos S., Voglis N., 2007, MNRAS, 374, 269
Benjamin J. et al., 2007, MNRAS submitted, astro-ph/0703570
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bialek J.J., Evrard A.E., Mohr J.J., 2001, ApJ, 555, 597
Boulade O. et al., 2003, SPIE, 4841, 72
Castillo-Morales A., Schindler S., 2003, A&A, 403, 433
Chiappetti L. et al., 2005, A&A, 439,413
Clowe D., De Lucia G., King L., MNRAS, 350, 1038
de Putter R., White M., 2005, New Astronomy, 10, 676
Eke C.R., Navarro J.F., Frenck C.S., 1998, ApJ, 503, 569
Ettori S, De Grandi S., Molendi S., 2002, A&A, 391, 841

Fahlman G., Kaiser N., Squires G., Woods D., 1994, ApJ, 437, 56
Finoguenov A., Reiprich T.H., Böhringer H., 2001, A&A, 368, 749
Gavazzi R., Soucail G., 2007, A&A, 462, 459 (GS07)
Hamana T., Takada M., Yoshida N, 2004, MNRAS, 350, 893
Heymans C. et al., 2005, MNRAS, 361, 160
Hjorth J., Oukbir J., van Kampen E., 1998, MNRAS, 298, 1L
Hoekstra H., 2001, A&A, 370, 743
Hoekstra H., 2003, MNRAS, 339, 1155
Hoekstra H. et al., 2006, ApJ, 647, 116
Horellou C., Berge J., 2005, MNRAS, 360, 1393
Hu W., Kravtsov A.V., 2003, ApJ, 584, 702
Huterer D., White M., 2002, ApJ, 578, 95L
Ilbert O. et al, 2006, A&A, 457, 841
Jarvis M., Jain B., Bernstein G., Dolney D., 2006, ApJ, 644, 71
Jenkins A., Frenck C.S., White S.D.M, Colberg J.M., Cole S., Evrard A.E., Couchman H.M.P., Yoshida N., 2001, MNRAS, 321, 372
Kaiser N., Squires G., 1993, ApJ, 404, 441
Kaiser N., Squires G., Broadhurst T., 1995, ApJ, 449, 460
Kitayama T., Suto Y., 1996, ApJ, 469, 480
Kneib J.-P., Ellis R.S., Smail I., Couch W.J., Sharples R.M., 1996, ApJ, 471, 643
Kneib J.-P., Mellier Y., Pello R., Miralda-Escude J., Le Borgne J.-F., Boehringer H., Picat J.-P., 1995, A&A, 303, 27
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lokas E.L., 2001, Acta Phys. Polon. B, 32, 3643
Lokas E.W., Bode P., Hoffman Y., 2004, MNRAS, 349, 595
Luppino G.A., Gioia I.M., Hammer F., Le Fèvre O., Annis J.A., 1999, A&A, 136, 117
Marian L., Bernstein G., 2006, PhReviewD, 73, 123525
Maor I., Lahav O., 2005, JCAP, 07, 003
Manera M., Mota D.F., 2006, MNRAS, 371, 1373
Massey R. et al., 2007c, MNRAS, 376, 13
Massey R., Refregier A., 2005, MNRAS, 363, 197
Massey R., Refregier A., Bacon D., Ellis R., Brown, M.L., 2005, MNRAS, 359, 1277
Massey R. et al., 2007b, Nature, 445, 286
Massey R. et al., 2007a, ApJ in press, astro-ph/0701480
Massey R., Rowe B., Réfrégier A., Bacon D., Bergé J., 2007d, MNRAS in press
Mellier Y., 1999, ARAA, 37, 127
Mellier Y., Fort B., Kneib J.-P., 1993, ApJ, 407, 33
Metzler C.A., White M., Loken C., 2001, ApJ, 547, 560
Miyazaki S. et al., 2002, ApJ, 580, 97L
Miyazaki S., Hamana T., Ellis R.S., Kashikawa N., Massey R.J., Taylor J., Refregier A., 2007, ApJ, 669, 714
Muanwong O., Thomas P.A., Kay S.T., Pearce F.R., 2002, MNRAS, 336, 527
Munshi D., Valageas P., van Waerbeke L., Heavens A., 2006, astro-ph/0612667
Nakamura T.T., 1996, Master's thesis, Univ. Tokyo
Nevalainen J., Markevitch M., Forman W., 2000, ApJ, 532, 694
Nunes N.J., da Silva A.C., Aghanim N., 2006, A&A, 450, 899
Ota N., Mitsuda K., 2004, A&A, 428, 757
Pacaud F. et al., 2006, MNRAS, 372, 578
Pacaud F. et al., 2007, doi: 10.1111/j.1365-2966.2007.12468.x, astro-ph/07091950
Padmanabhan T., 1993, Structure Formation in the Universe, Cambridge Univ. Press, Cambridge
Paulin-Henriksson S., Antonuccio-Delogu V., Haines C.P., Radovich M., Mercurio A., Becciani U., 2007, A&A, 467, 427
Pedersen K., Dahle H., 2007, ApJ, 667, 26
Peebles, P.J.E., 1980, The Large-Scale Structure of the Universe, Princeton Univ. Press, Princeton, NJ
Pierpaoli E., Borgani S., Scott D., White M., 2003, MNRAS, 342, 163
Pierre M. et al., 2004, JCAP, 9, 11
Pierre M. et al., 2006, MNRAS, 372, 591
Press,W.H., Schechter P., 1974, ApJ, 187, 425
Rasia E. et al., 2006, MNRAS, 369, 2013
Reblinsky K., Bartelmann M., 1999, A&A, 345, 1
Refregier A., 2003a, ARAA, 41, 645
Refregier A, 2003b, MNRAS, 338, 35
Refregier A., Bacon D., 2003, MNRAS, 338, 48
Rhodes J., Refregier A., Groth E.J., 2001, ApJ, 552, L85
Schrabback T. et al., 2007, A&A, 468, 823
Seljak U., 2002, MNRAS, 337, 769
Semboloni E. et al., 2006, A&A, 452, 51
Smail I., Ellis R.S., Dressler A., Couch W.J., Oemler A.J., Sharples R.M., Butcher H., 1997, ApJ, 479, 70
Smith G.P., Edge A.C., Eke V.R., Nichol R.C., Smail I., Kneib J.-P., 2003, ApJ, 570, L79
Smith G.P., Kneib J.-P., Smail I., Mazzotta P., Ebeling H., Czoske O., 2005, MNRAS, 359, 417
Spergel D.N. et al., 2007, ApJS, 170, 377
Starck J.L., Murtagh F., Bijaoui A., 1998, Image Processing and Data Analysis : The Multiscale Approach. Cambridge Univ. Press, Cambridge
van Waerbeke L. et al., 2000, A&A, 358, 30
van Waerbeke L., White M., Hoekstra H., Heymans C., 2006, Astropart. Phys., 26, 91
Viana P.T.P, Liddle A.R., 1996, MNRAS, 281, 323
Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S.S., Van Speybroeck L., 2006, ApJ, 640, 691
Wang L., Steinhardt P.J., 1998, ApJ, 508, 483
Weinberg N.N., Kamionkowski M., 2003, MNRAS, 341, 251
Willis J.P. et al., 2005, MNRAS, 363, 675
Wittman D.M., Dell'Antonio I.P., Hughes J.P., Margoniner V.E., Tyson J.A., Cohen J.G., Norman D., 2006, ApJ, 643, 128
Wittman D.M., Tyson J.A., Kirkman D., Dell'Antonio I., Bernstein G., 2000, Nature, 405, 143
Zhang Y.Y., Finoguenov A., Böhringer H., Kneib J.P., Smith G.P., Czoske O., Soucail G., 2007, A&A, 467, 437

## 7.8 Prospectives sur les relevés X :

## "The XMM-LSS cluster sample and its cosmological applications. Prospects for the XMM next decade"

### Contexte/motivation

Au vu des résultats préliminaires du XMM-LSS et de la difficulté à obtenir du temps d'observation pour un projet aussi large, nous avons profité du workshop de prospectives sur l'héritage d'XMM, organisé par l'European Space Astronomy Center, pour dresser le bilan des travaux effectués et envisager le potentiel de ce type de relevé dans l'avenir.

### Résumé

La fonction de sélection bien caractérisée du relevé XMM-LSS permet de modéliser simultanément les comptages d'amas observés et l'évolution de la relation $L_X$-$T$. Nous récapitulons dans cet article les résulats relatifs aux premiers 5 $deg^2$ pour un échantillon contrôlé comprennant 30 objets : ils sont compatibles avec le jeu de paramètres de WMAP3 associés à une évolution auto-similaire des amas.

Nous montrons ensuite que l'extension d'un tel relevé à 200 $deg^2$ pourrait (1) permettre la distinction entre les principaux scenarii d'évolution de la relation $L_X$-T et (2) fournir une détermination unique et auto-suffisante de $\sigma_8$ et $\Gamma$ avec une précision de $\sim$5% and 10% respectivement. Il faut pour cela y ajouter des informations sur la masse des objets provenant d'observations du cisaillement gravitationnel ou de l'effet Sunyaev-Zel'dovich.

### Contribution personnelle

Le bilan érigé sur les premières années du relevé repose pour l'essentiel sur les travaux développés dans Pacaud et al. (2007) et Pierre et al. (2007). J'ai de plus estimé les contraintes prévisionnelles sur l'évolution de la relation $L_X$-$T$ et participé aux discussions sur les contraintes cosmologiques à présenter.

Astron. Nachr. / Solicited talk given at **XMM-Newton: The Next Decade**, ESAC, Madrid 4-6 June 2007
**000**, No. 00, 1 – 4 (0000) / **DOI *please set DOI!***

# The XMM-LSS cluster sample and its cosmological applications. Prospects for the XMM next decade

**M. Pierre**[1,⋆], **F. Pacaud**[1,2], **J.B. Melin**[3], and **the XMM-LSS consortium**

[1] DAPNIA/Service d'Astrophysique, Laboratoire AIM CNRS, CEA-Saclay, F-91191 Gif-sur- Yvette, France
[2] Argelander-Institut für Astronomie, University of Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
[3] DAPNIA/Service de Physique des Particules, CEA Saclay, F-91191 Gif-sur-Yvette, France.

Received : September 2007, accepted : November 2007

**Key words** X-rays: galaxies: clusters – cosmological parameters

The well defined selection function of the XMM-LSS surveys enables a simultaneous modelling of the observed cluster number counts and of the evolution of the L-T relation. We present results pertaining to the first 5 $\deg^2$ for a well controlled sample comprising 30 objects : they are compatible with the WMAP3 parameter set along with cluster self-similar evolution. Extending such a survey to 200 $\deg^2$ would (1) allow discriminating between the major scenarios of the cluster L-T evolution and (2) provide a unique self-sufficient determination of $\sigma_8$ and $\Gamma$ with an accuracy of $\sim$ 5% and 10% respectively, when adding mass information from weak lensing and S-Z observations.



## 1 Introduction

It has been recognised for a long time that clusters of galaxies, as the most massive entities of the universe, can be used as cosmological probes. They provide key information on the normalisation of the power spectrum and are potentially suitable for studying the properties of dark energy. They represent important and independent constraints in addition to those from the CMB and supernovae because they involve very different physics. It is also critical to ensure consistency between the cosmological constrains from the early and local universe.

Main statistical tools for cluster studies are the cluster number counts ($dn/dz$) and the cluster two-point correlation function ($\xi$). This requires that, whatever the detection wavelength, the samples must be in some sense, complete and uncontaminated and thus, requires well understood detection and selection procedures. Cluster physics evolution is a key ingredient in interpreting the observed cluster density as function of redshift. It is usually modelled in the form of scaling laws relating observable quantities such as flux, richness, luminosity or temperature to cluster masses. The cluster scaling laws are, however, still poorly known beyond the local universe.

With its mosaic of overlapping XMM pointings ($10^4$ s), the XMM Large-Scale Structure survey (XMM-LSS, Pierre et al 2004) has been designed to detect a significant fraction of the cluster population out to $z = 1$, over an area of several tens of $\deg^2$, so as to constitute a sample suitable for cosmological studies. We present below the procedures developed to detect the clusters and to further analyse their number counts along with their temperature and luminosity distribution in a self-consistent approach. In light of the results obtained so far, we discuss the cosmological impact of a future 200 $\deg^2$ XMM wide survey.

⋆ Corresponding author: e-mail: mpierre@cea.fr

## 2 Detecting and selecting clusters in the XMM-LSS survey

In the redshift range of interest, although the cluster apparent sizes ($20'' < R_c < 100''$) are significantly larger than the XMM PSF and source confusion can be considered as negligible, cluster detection is a very specific task since our objects are weak sources (count-rate from 0.3 to 3 counts/min). We developed a two-step procedure combining wavelet multi-resolution analysis and maximum likelihood fits both using Poisson statistics. The pipeline was extensively tested using simulations which allowed us to define a sub-region in the `extent` vs `extent likelihood` parameter space, where the contamination level by point-sources is lower than 1%. This constitutes the *class one* (C1) cluster sample. Strictly speaking, this selection is not flux limited, but allows the construction of well controlled and uncontaminated cluster samples significantly larger than a simple flux limit would allow (Pacaud et 2006). Our C1 sample shows a density of $\sim 6$ clusters per $\deg^2$ .

## 3 Current cluster results

The XMM-LSS currently covers 10 $\deg^2$. It is located in the W1 area of the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) and associated with a number of surveys in the radio, infrared and UV domains (Fig. 1). In these pro-

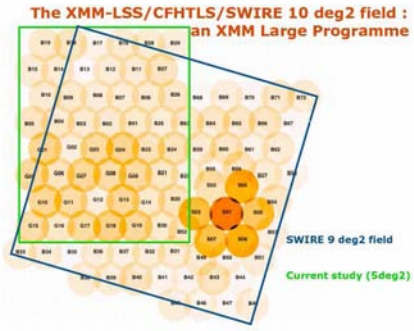


**Fig. 1 Layout of the 98 XMM-LSS pointings** including the Subaru Deep Survey. The colour scale indicates the effective exposures, from 80 ks (SDS01) to $\sim 0$. The green rectangle delineates data obtained prior to the AO5. In addition to the coverage by the SWIRE and CFHT Legacy Surveys, observations from the VLA, Integral, UKIDSS and Galex are available in the field. Full coverage by SCUBA2 and Herschel is planned.

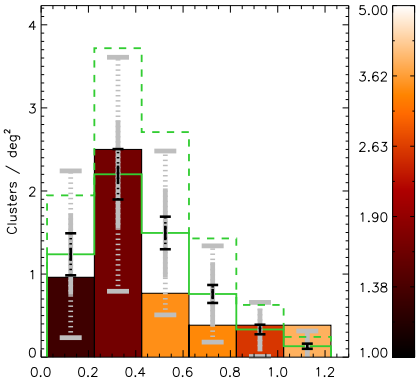


**Fig. 2 The current C1 cluster redshift distribution over the first 5 $\mathrm{deg}^2$ of the XMM-LSS**. The colour scale indicates the cluster mean temperature for each bin (unweighted mean of the individual cluster temperatures in keV.) The solid green histogram shows the expectations of our cosmological model (WMAP3: $\sigma_8$ = 0.74 and self- similar evolution for the Lx-T relation) along with the Poisson error bars. The dash-line histogram shows the expectations for a model with WMAP 1st year cosmological parameters ( $\sigma_8$ = 0.85) and a non-evolving Lx-T relation. Fluctuations around the mean expectation are represented by the solid and dotted error bars for the shot noise and sample variance (estimated from Hu & Kravtsov 2003) respectively. The grey error bars are for 5 $\mathrm{deg}^2$, while the black ones for 200 $\mathrm{deg}^2$.

ceedings, we summarise the results from the first 5 $\mathrm{deg}^2$ which are presented in detail by Pacaud et al 2007 (P07).

### 3.1 Modelling the cluster number counts

Some 30 C1 clusters are found in the first 5 $\mathrm{deg}^2$. They all have been spectroscopically confirmed. Their redshift distribution is displayed on Fig. 2. P07 performed an ab initio modelling of the observed number counts as follows: assume a (1) cosmological framework ($\Lambda$CDM) and a power spectrum along with a transfer function; (2) a mass function, (3) a halo model, (4) various scaling evolutionary relations for cluster physics. Then, for each redshift and mass range, the predicted luminosities are transformed into XMM count-rates using a dedicated plasma code. The C1 selection criteria are finally applied, resulting in a simulated observed redshift distribution (Fig. 2). The effect of the degeneracy between cluster scaling law evolution and cosmology is clearly illustrated on the figure; however, our results, despite the still small size of the sample, favour the WMAP3 parameter sets along with cluster self-similar evolution.

### 3.2 Evolution of the L-T relation

Each C1 cluster undergoes dedicated spatial and spectral fits in order to derive reliable luminosity and temperature estimates. Willis et al (2005) have shown that using a well adapted binning procedure, it is statistically possible to obtained a 20% temperature accuracy with only 200 photons for groups up to 2 keV. It turns out that for all C1 clusters, we obtain temperature measurements with satisfactory accuracy. The average temperature of the C1 clusters as a function of redshift is displayed on Fig. 2. Because of the very tight relation between X-ray temperature and luminosity, the mean temperature of the detected clusters appears to increases with redshift (Malmquist bias). This redshift - temperature distribution shows that the XMM-LSS survey unveils for the first time the population of low-mass groups (T= 2 keV) around $z = 0.3$, which constitute the building blocks of the present day clusters.

Assuming that all clusters, whatever their mass, follow the same evolutionary scaling laws, we have used our data to constrain the evolution of the L-T relation. The observed luminosity enhancement, with respect to the local expectation, is computed for the 30 clusters distributed in four redshift bins (Fig 3). The raw data suggest a rather strong evolution, best fitted with a two-parametre model. However, the inclusion of the survey selection function in the fit suggests a much more mild evolution, fully compatible with the predictions of the self-similar model. The fact that selection effects were not systematically considered in the former L-T(z) studies may explain the discordant results obtained so far (see a compilation in P07): because the cluster mass function is so steep, most clusters are detected around the survey limiting sensitivity, hence favouring the compilation

 

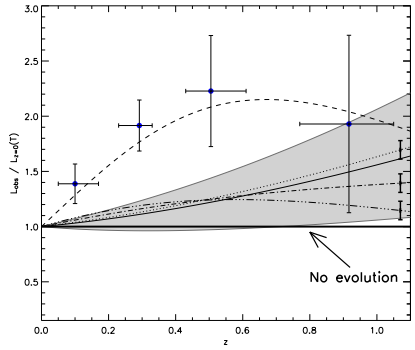


**Fig. 3** **The L-T relation**. The graph shows the cluster X-ray luminosity enhancement with respect to the expectation at $z = 0$. Our clusters are sorted in 4 redshift intervals. The dash line is the result of an ad hoc two-parameter fit to the raw data points: $(1+z)^{\alpha} \times E(z)^{\beta}$ with $\alpha = 4.7,\ \beta = -5.4$. When the survey selection function is taken into account, the best fit is the solid line, $(1+z)^{\alpha}E(z)$ with $\alpha = -0.07$, which is very close to the evolution predicted by the self-similar model (dotted line, $\alpha = 0$ ). The grey region delineates the $1\sigma$ confidence interval for the surveyed $5\ \mathrm{deg}^2$. The dot-dash and triple-dot-dash lines are the evolutionary models by Voit (2005) including non gravitational physics. The thick error bars at the end of the three models indicate the expectations for a 200 $\mathrm{deg}^2$ survey.

of over-luminous objects at any redshift. This is certainly not a feature unique to the XMM-LSS survey, but rather of any X-ray cluster survey (provided that the detection is performed down to the capabilities of the survey). This bias is to affect any sub-sample "randomly" selected from all-sky or serendipitous surveys for subsequent deep XMM or Chandra temperature observations. We emphasise that the XMM-LSS survey represents the first attempt, not only to determine the L-T relation with survey data alone but also to explicitly include the selection effects in the determination of the evolution of the scaling laws. The latter was possible thanks to the very well modelled survey selection function.

## 4 AGNs in the XMM-LSS

The XMM-LSS provides a source density of $\sim 300\ /\mathrm{deg}^2$ down to a flux limit of $4\ 10^{-15}\ \mathrm{erg\,s^{-1}\,cm^2}$ in the [0.5-2] keV band (95% level completeness limit). This constitutes the largest deep AGN X-ray sample over a single field. The special relevance of the sample also comes from its unique multi-$\lambda$ coverage.

### 4.1 The spatial distribution

The XMM-LSS data set allowed, for the first time, the study of the angular distribution of faint AGNs over an area of $5\ \mathrm{deg}^2$. We found a significant clustering in the soft band and none in the hard band. A sub-sample of $\sim 200$ sources with hard X-ray count ratios, likely dominated by obscured AGNs, does show a positive signal allowing for a large angular correlation length at the $3\sigma$ level (Gandhi et al 2006).

### 4.2 Spectral properties

A dedicated study of some 100 AGN selected in the hard band reveals a mismatch between the classification based on the characteristics of the optical emission lines and the classification given by the X-ray spectroscopy. This led to question some aspects of the AGN unified scheme (Garcet et al 2007). The many CFHTLS and SIWRE flux data points allowed us to perform SED fits on the XMM-LSS point source population. From this, it was possible to classify the AGN (star formation, type 1 or 2 AGN, Seyfert) and to obtain photometric redshifts. Combining with the X-ray spectral data points, we demonstrate that the SED properties are continuous through the various classes (Tajer et al 2007, Polletta et al 2007).

## 5 Data releases

Cluster data (positions, redshifts, $\mathrm{L}_X$, $\mathrm{T}_X$, mass estimates) are published for the first $5\ \mathrm{deg}^2$ along with the scientific analysis (Valtchanov et al 2004 , Willis et al 2005, P07). The complete source catalogue with optical data and thumb nail images is also public (Pierre et al 2007). Data can be retrieved from the CDS or, in a more extensive form, via the consortium data bases for the cluster[1] and source[2] catalogues.

## 6 Next decade with XMM: A 200 $\mathrm{deg}^2$ wide survey

During the past years, we explored a number of issues regarding cluster detection and science with the XMM-LSS survey. In addition to the C1 clusters, a complementary cluster sample of about the same size has been identified in the XMM-LSS, but for which the selection criteria are less well defined (Adami et al in prep). We have also detected a number of $z > 1$ clusters (Andreon et al 2005, Bremer et al 2006). The determination of the luminosity, temperature and mass (hydrostatic hypothesis) of the C1clusters from survey data appears to provide quantities reliable for statistical cosmological studies[3]. A weak lensing analysis over part of the XMM-LSS area shows promising prospects for constraining independently the slope and the normalisation of the M-T relation (Berge et al 2007). A further step

[1] http://l3sdb.in2p3.fr:8080/l3sdb/

[2] http://cosmos.iasf-milano.inaf.it/ lssadmin/Website/LSS/Query/

[3] The mass of XLSS 29 at $z = 1.05$ was measured to be $1.4\ 10^{14}\ M_{\odot}$ with survey data (P07) and $1.8 \pm 0.5\ 10^{14}\ M_{\odot}$ from a subsequent 80 ks XMM pointing (Maughan et al 2007)

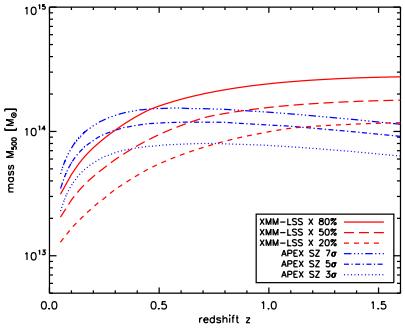


**Fig. 4 Comparison between the XMM-LSS and Sunyaev-Zel'dovich sensitivities** in terms of limiting mass. The red lines show various *measured* detection probability thresholds for the C1 clusters. The blue lines are the *predictions* for the $10\mu$K APEX survey, currently observing the XMM-LSS field. For the regime of interest ($z < 1$), the X-ray observations are at least as efficient as the S-Z ones in terms of cluster detection.

in constraining cluster masses is to be reached by the upcoming generation of Sunyaev- Zel'dovich surveys which, in principle, with a sensitivity of $10\mu K$, are well matched to mass range of the XMM-LSS survey (Fig. 4).

Further, we have shown that, in order to increase the precision of the L-T relation evolution, it is more efficient to increase the cluster sample than the accuracy on the temperature measurements - a useful tip for a proper use of future XMM observing time. This is due to the large intrinsic dispersion of the L-T relation itself (P07).

In this way, a 200 $\deg^2$ survey with 10 ks XMM pointings[4], would allow definitively discriminating not only between cluster self-similar evolution and no evolution, but also between other theoretically justified models based on non-gravitational physics (Fig. 3). Such a 200 $\deg^2$ survey would moreover overcome sample variance problems and provide more than a thousand C1 clusters as well as some 100 clusters at $z > 1$. Having determined the cluster evolution rate, an important degeneracy would be removed in the cosmological interpretation of the cluster number counts (Fig. 2) leaving a handle on the equation of state of the Dark Energy (P07). The cluster sky distribution provides additional constraints on the cosmological parameters and a 200 $\deg^2$ survey will not only allow a self-sufficient and accurate determination of $\sigma_8$, i.e. 5 %, but also of the slope of the matter power spectrum, $\Gamma$, i.e. 10% (Fig. 5).

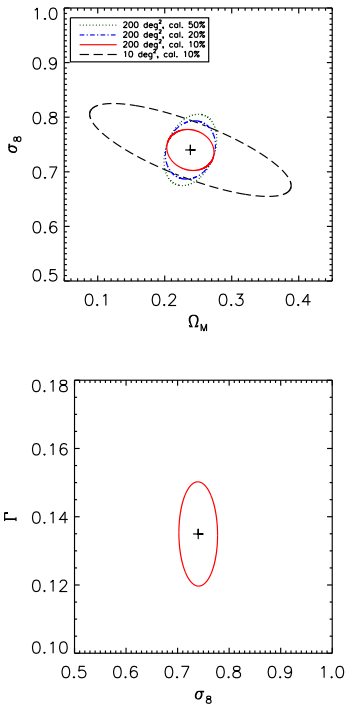


**Fig. 5 Top: Constraints on the $\sigma_8 - \Omega_m$ plane from $dn/dz + \xi(r)$.** All contours are $1\sigma$ for C1 clusters, marginalised over $\Omega_\Lambda$. (*dots*) The C1 X-ray cluster population alone ($6/\deg^2$) over 200 $\deg^2$; adding information from S-Z (*dash-dot*) and weak lensing (*solid*) mass measurements; (*dash*) same as solid but for 10 $\deg^2$ coverage. X-ray masses are taken to be accurate to 50%, adding S-Z then weak lensing data reduces this to 20% then to 10%; the latter giving an accuracy on $\sigma_8$ of 6 %. **Bottom: Constraints on the $\sigma_8 - \Gamma$ plane from $\xi(r) + dn/dz$.** Contours are 1 $\sigma$ for the C1 population, 200 $\deg^2$, mass accuracy of 10% and marginalisation over $\Omega_m$ and $\Omega_\Lambda$. The expected accuracy on $\Gamma$ is 10%.

## References

Andreon, S., et al: 2005, MNRAS 359, 1250
Berge J. et al: 2007, MNRAS submitted
Bremer, M., et al: 2006, MNRAS 371, 1427
Gandhi, P., et al : 2006, A&A 457, 393
Garcet, O., et al : 2007, A&A in press (astroph/0709.2027)
Hu & Kravtsov: 2003, ApJ 584, 702
Maughan, B., et al: 2007, MNRAS submitted (astroph/07092300)
Pacaud, F. et al : 2006, MNRAS 372, 578
Pacaud, F et al (P07): 2007, MNRAS in press (astroph/0709.1950)
Pierre, M., et al : 2004, JCAP 9, 11
Pierre, M., et al: 2007, MNRAS in press (astroph/0708.3299)
Polletta, M., et al: 2007, ApJ 663, 81
Tajer, M., et al 2007: A&A 467, 73
Valtchanov I., et al : 2004 A&A 423, 75
Voit, M.: 2005, Rev. Mod. Phys 77, 207
Willis, J., et al : 2005, MNRAS 363, 675

[4] such a survey would require some 24 Ms with the current observing settings, which could be decreased by about 1/3 when the foreseen mosaicing mode with a reduced pn overhead is implemented

 

# Troisième partie:

## Annexes

# Annexe A
# Notions d'analyse multirésolution

La transformation en ondelette présentée en section 4.3.3, est à inscrire dans le cadre plus large des méthodes dites d'*analyse multirésolution* (Mallat, 1989).

## A.1 Définition

Mathématiquement, une analyse multirésolution dyadique $M$ de $\mathcal{L}^2(\mathbb{R})$ (l'espace des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$ et de carré sommable) est une suite de sous-espaces $V_j$ de $\mathcal{L}^2(\mathbb{R})$, vérifiant les propriétés suivantes :

$$\{V_j\}_{j\in\mathbb{Z}} \begin{cases} (i) & \forall j \in \mathbb{Z},\ \ V_{j+1} \subset V_j \\ (ii) & \bigcap V_j \ = \ \{0\} \\ (iii) & \bigcup V_j \ = \ \mathcal{L}^2(\mathbb{R}) \\ (iv) & \forall j \in \mathbb{Z}, f(x) \in V_j \ \Leftrightarrow \ f(\frac{x}{2}) \in V_{j+1} \\ (v) & \forall (j,k) \in \mathbb{Z}^2, f(x) \in V_j \ \Leftrightarrow \ f(x-2^j k) \in V_j \\ (vi) & \exists\ \phi \in V_0 \mid \{\phi(x-k)\}_{k\in\mathbb{Z}} \ \textit{forme une base de } V_0. \end{cases} \tag{A.1}$$

Chacun des sous-espace $V_j$ est alors appelé *espace d'approximation*. De manière imagée, la propriété $(i)$ stipule que les sous-espaces sont imbriqués les uns dans les autres, et permettent donc de décrire une fraction d'autant plus grande de $\mathcal{L}^2(\mathbb{R})$ que $j$ est petit (ce qui justifie la dénomination *espace d'approximation*). Les relations $(iv)$ et $(v)$ permettent d'associer au paramètre $j$ une notion de résolution : l'échelle caractéristique des fonctions appartenant à un sous-espaces suit une progression géométrique. Les relations $(ii)$ et $(iii)$ impliquent qu'à résolution minimale (resp. maximale), le sous-espace d'approximation perd (resp. possède) toute aptitude à décrire l'espace $\mathcal{L}^2(\mathbb{R})$.

La fonction $\phi$ définit en $(iv)$ est appelée *fonction d'échelle*. Elle confère une structure analytique simple aux sous-espaces $V_j$. La combinaison des propriétés

$(iv)$ et $(v)$ implique en effet que pour tout $j$ appartenant à $\mathbb{Z}$, $\{2^{-j}\phi(2^{-j}x-k)\}_{k\in\mathbb{Z}}$ est une base de $V_j$.

Puisque l'analyse multirésolution définit un ensemble d'espace emboîtés, toute fonction du sous-espace $V_{j+1}$ appartient également à $V_j$ et peut donc être développée sur une base de $V_j$. En appliquant cela à la fonction d'échelle (qui génère une base de $V_j$) et à sa dilatation d'un facteur 2 (qui appartient à $V_{j+1}$), on obtient la propriété fondamentale :

$$\frac{1}{2}\phi(\frac{x}{2}) \quad = \quad \sum_k h(k)\phi(x-k). \tag{A.2}$$

Cette *relation d'échelle* permet de définir un filtre passe bas $h(k)$ qui détermine par récurence la projection d'un signal depuis un espace d'approximation vers les autres.

## A.2 Lien avec les ondelettes

Par définition de l'analyse multirésolution, on a $V_j \subset V_{j-1}$. Il est donc possible de compléter $V_j$ par son espace supplémentaire dans $V_{j-1}$, $W_j$, tel que :

$$V_j \cup W_j = V_{j-1} \; et \; V_j \cap W_j = \{0\}. \tag{A.3}$$

On notera $V_{j-1} = V_j \oplus W_j$. Ces sous-espaces permettent de décrire l'information perdue lorsque la résolution est dégradée et sont donc appelés espaces des détails.

On montre qu'il est toujours possible de définir une fonction $\psi$ appelée ondelette mère associée à l'analyse mulitrésolution $M$ dont les translations/dilatations forment des bases des espaces de détails $W_j$ :

$$\forall j \in \mathbb{Z}, \quad \left\{2^{-j}\,\psi(2^{-j}x-k)\right\}_{k\in\mathbb{Z}} \quad forme\ une\ base\ de\ W_j. \tag{A.4}$$

La base d'ondelette ainsi définie permet de construire une transformation en ondelette discrète dyadique analogue à celle présentée dans la section 4.3.3. Réciproquement, il est démontré que toute ondelette peut être inscrite dans le cadre d'une analyse multirésolution.

De la même manière que pour la fonction d'échelle, $W_1 \in V_0$ et la première dilatation de l'ondelette $\psi(x)$ peut donc être décomposée selon la base définit par les fonctions d'échelles translatées :

$$\frac{1}{2}\psi(\frac{x}{2}) \; = \; \sum_k g(k)\phi(x-k), \tag{A.5}$$

définissant ainsi un filtre passe haut $g$ qui permet la projection depuis un espace d'approximation vers l'espace des détails de la résolution inférieure.

La définition des filtres $h$ et $g$ permet de définir entièrement les fonction d'échelle et ondelette associées à une analyse multirésolution. Leur existence permet en outre de simplifier considérablement la mise en øeuvre numérique des méthodes d'ondelettes.

## A.3 Transformation discrète : l'algorithme 'à trou'

### A.3.1 Principe

La mise en oeuvre d'une transformation en ondelettes discrète (entendre par là appliquée à des données discrètes) peut se faire de plusieurs manières. Dans le cas de l'application au traitement d'image, une résolution supérieure à la taille du pixel ou inférieur à la taille de l'image n'a pas de sens. Il est donc d'usage de définir l'espace d'approximation $V_0$ comme celui correspondant à la taille du pixel, et de restreindre l'analyse multirésolution aux échelles $j \in \mathbb{N}$. De plus, le signal n'est généralement décomposé que partiellement, c'est à dire modélisé comme la somme de sa projection sur un espace d'approximation de basse résolution $p$, et de ses projections sur les espaces de détails de résolution supérieure :

$$f(x) = \sum_k < f, \phi_{p,k} > \phi_{p,k} + \sum_{j=1}^{p} \left( \sum_k < f, \psi_{j,k} > \psi_{j,k} \right), \tag{A.6}$$

où $\phi_{p,k}$ signifie $2^{-p}\phi(\frac{x-k}{2^{-p}})$.

Une méthode relativement simple à implémenter et d'utilisation courante pour obtenir ce genre de décomposition est l'algorithme 'à trou'. Celui-ci se définit à partir du spectre du signal dans la base générée par la fonction d'échelle :

$$c_0(k) = \langle f(x), \phi(x-k) \rangle \tag{A.7}$$

D'après la relation d'échelle (A.2) associée au filtre $h$, les approximations successives du signal s'obtiennent simplement par récurrence :

$$c_j(k) = \sum_l h(l) c_{j-1}(k + 2^{j-1} l) \tag{A.8}$$

Afin d'exploiter directement cette propriété, les coefficients d'ondelette s'expriment dans l'algorithme 'à trou' comme :

$$w_i(k) = c_{i-1}(k) - c_i(k) \tag{A.9}$$

Ce qui correspond à une ondelette :

$$\frac{1}{2}\psi(\frac{x}{2}) = \phi(x) - \frac{1}{2}\phi(\frac{x}{2}). \tag{A.10}$$

Le nom de l'algorithme provient des 'trous' générés par les sauts de $2^j$ pixels effectués dans le calcul des coefficients $c_{j-1}$ au cours de la sommation A.8.

Outre sa relative facilité de mise en øeeuvre, l'algorithme 'à trou' à de nombreux avantages :

- Il s'agit d'une transformation *cubique* (la transformée comporte le même nombre de point que le signal à chaque échelle), ce qui la rend très redondante, et par cela particulièrement appropriée au applications de filtrage ou débruitage.

– Contrairement à nombre d'algorithmes pour les ondelettes, son calcul ne nécessite pas le passage dans l'espace de Fourrier ce qui évite les artefacts introduits par la périodisation.
– Son temps d'éxecution est plus court que la plupart des autres méthodes de décomposition.

Enfin, le spectre du signal reconstruit s'obtient facilement par sommation :

$$c_0(k) = c_p(k) + \sum_{j=1}^{p} w_j(k) \tag{A.11}$$

Il existe cependant un défaut notable de l'algorithme 'à trou' lié à sa forte redondance. La correspondance des espaces réel et ondelette n'est plus bi-univoque. Il est alors nécesaire d'utiliser un procéssus itératif afin que la transfomée en ondelette par l'algorithme 'à trou' du signal filtré reconstruit corresponde bien aux coefficients $w_j$ calculés.

### A.3.2 Choix de la fonction d'échelle

L'algorithme ainsi définit est entièrement caractérisé par le choix de la fonction d'échelle (où de manière équivalente du filtre $h$) et peut être facilement généralisé à 2 dimensions pour le traitement d'images. Différents critères conduisent à utiliser de manière systématique les interpolations B-splines centrées comme fonction d'échelle avec l'algorithme à trou. Celles-ci sont d'abord symétrique (ce qui simplifie l'interprétation des décallages par la suite), elle sont régulières (ce qui évite les instabilités), et elles correspondent à une bande très étroite dans l'espace de Fourrier.

Mathématiquement, elles peuvent être définies à l'ordre $l-1$ par leurs transformées de Fourier :

$$\hat{B}_{l-1}(\nu) = \left(\frac{\sin \pi\nu}{\pi\nu}\right)^l \tag{A.12}$$

Dans le cas du XMM-LSS, c'est la fonction B-spline d'ordre 3 qui est utilisée. Elle correspond à un filtre 1D $h = (\frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16})$. La fonction déchelle et l'ondelette résultantes sont représentées figure A.1.

La généralisation à 2 dimensions donne :

$$\hat{\phi}(u,v) = \frac{3}{2}\hat{B}_3(4\sqrt{u^2+v^2}), \tag{A.13}$$

et

$$h = \begin{pmatrix} \frac{1}{256} & \frac{1}{64} & \frac{3}{128} & \frac{1}{64} & \frac{1}{256} \\ \frac{1}{64} & \frac{1}{16} & \frac{3}{32} & \frac{1}{16} & \frac{1}{64} \\ \frac{3}{128} & \frac{3}{32} & \frac{9}{64} & \frac{3}{32} & \frac{3}{128} \\ \frac{1}{64} & \frac{1}{16} & \frac{3}{32} & \frac{1}{16} & \frac{1}{64} \\ \frac{1}{256} & \frac{1}{64} & \frac{3}{128} & \frac{1}{64} & \frac{1}{256} \end{pmatrix}. \tag{A.14}$$

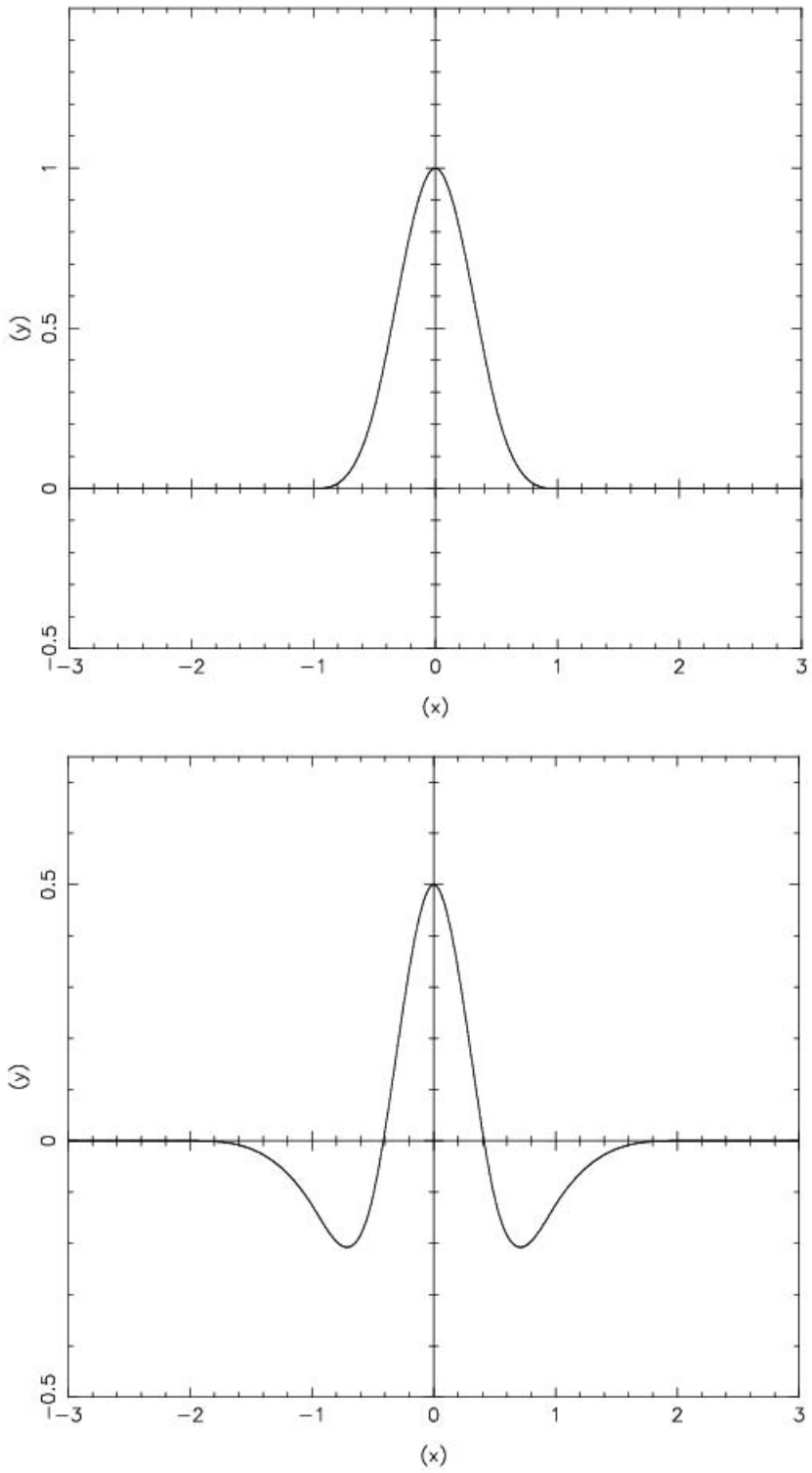


**Fig. A.1**: *Fonction d'échelle (en haut) et ondelette (en bas) associées à l'interpolation B-spline cubique.*

# Annexe B
# Contributions écrites et orales

## Publications référées

Maughan, B., Jones, L., Pierre, M., Andreon, S., Birkinshaw, M., Bremer, M., **Pacaud**, **F.**., Ponman, T., Valtchanov, I. & Willis, J., "Testing the galaxy cluster mass-observable relations with Chandar and XMM-Newton observations of XLSSC 029", soumis à MNRAS.

Bergé, J., **Pacaud**, **F.**., Réfrégier, A., Massey, R., Pierre, M., Amara, A., Birkinshaw, M., Paulin-Henriksson, S., Smith, G. P. & Willis, J., "Combined analysis of weak lensing and X-ray blind surveys", accepté par MNRAS, `ArXiv :0712.3293` (2008).

Pierre, M., **Pacaud**, **F.**., Melin, J. & le consortium XMM-LSS, "The XMM-LSS cluster sample and its cosmological applications. Prospects for the XMM next decade", accepté par Astronomische Nachrichten (proceedings du symposium de l'ESA/ESAC : "XMM-Newton : The Next Decade"), `ArXiv :0712.0262v1` (2008).

Garcet, O., Gandhi, P., Gosset, E., Sprimont, P. G., Surdej, J., Borkowski, V., Tajer, M., **Pacaud**, **F.**., Pierre, M., Chiappetti, L. et al., "The XMM large scale structure survey : optical vs. X-ray classifications of active galactic nuclei and the unified scheme", A&A, **474**, 473 (2007).

Mazure, A., Adami, C., Pierre, M., Le Fèvre, O., Arnouts, S., Duc, P. A., Ilbert, O., Lebrun, V., Meneux, B., **Pacaud**, **F.**. et al., "Structure detection in the D1 CFHTLS deep field using accurate photometric redshifts : a benchmark", A&A, **467**, 49 (2007).

**Pacaud**, **F.**., Pierre, M., Adami, C., Altieri, B., Andreon, S., Chiappetti, L., Detal, A., Duc, P.-A., Galaz, G., Gueguen, A. et al., "The XMM-LSS survey : the Class 1 cluster sample over the initial 5 deg$^2$ and its cosmological modelling", MNRAS, **382**, 1289 (2007).

Pierre, M., Chiappetti, L., **Pacaud**, **F.**., Gueguen, A., Altieri, B., Aussel, H., Gandhi, P., Garcet, O., Gosset, E., Paioro, L. et al., "The XMM-LSS catalogue : X-ray sources and associated optical data. Version I", MNRAS, **382**, 279 (2007).

Bremer, M. N., Valtchanov, I., Willis, J., Altieri, B., Andreon, S., Duc, P. A., Fang, F., Jean, C., Lonsdale, C., **Pacaud**, **F.**. et al., "XMM-LSS discovery of a z = 1.22 galaxy cluster", MNRAS, **371**, 1427 (2006).

Gandhi, P., Garcet, O., Disseau, L., **Pacaud**, **F.**., Pierre, M., Gueguen, A., Alloin, D., Chiappetti, L., Gosset, E., Maccagni, D. et al., "The XMM large scale structure survey : properties and two-point angular correlations of point-like sources", A&A, **457**, 393 (2006).

**Pacaud**, **F.**., Pierre, M., Refregier, A., Gueguen, A., Starck, J.-L., Valtchanov, I., Read, A. M., Altieri, B., Chiappetti, L., Gandhi, P. et al., "The XMM Large-Scale Structure survey : the X-ray pipeline and survey selection function", MNRAS, **372**, 578 (2006).

Pierre, M., **Pacaud**, **F.**., Duc, P.-A., Willis, J. P., Andreon, S., Valtchanov, I., Altieri, B., Galaz, G., Gueguen, A., Fèvre, J.-P. L. et al., "The XMM Large-Scale Structure survey : a well-controlled X-ray cluster sample over the D1 CFHTLS area", MNRAS, **372**, 591 (2006).

Willis, J. P., **Pacaud**, **F.**., Valtchanov, I., Pierre, M., Ponman, T., Read, A., Andreon, S., Altieri, B., Quintana, H., Dos Santos, S. et al., "The XMM Large-Scale Structure survey : an initial sample of galaxy groups and clusters to a redshift z < 0.6", MNRAS, **363**, 675 (2005).

Andreon, S., Willis, J., Quintana, H., Valtchanov, I., Pierre, M. & **Pacaud**, **F.**., "Galaxy evolution in clusters up to z= 1.0", MNRAS, **353**, 353 (2004).

Pierre, M., Valtchanov, I., Altieri, B., Andreon, S., Bolzonella, M., Bremer, M. N., Disseau, L., Dos Santos, S., Gandhi, P., Jean, C. et al., "The XMM-LSS survey. Survey design and first results", Journal of Cosmology and Astro-Particle Physics, **9**, 11 (2004).

## Conférences et posters

**Pacaud**, **F.**., "Cosmological Applications of the XMM-LSS Cluster Sample", dans *séminaire au Argelander Institut für Astronomie de l'université de Bonn* (2006).

**Pacaud**, **F.**. & Pierre, M., "Cosmology with the XMM-LSS cluster sample", dans *XLIst Rencontres de Moriond : "From Dark Halo to Light", intervention orale* (2006).

Willis, J., **Pacaud**, **F.**. & Pierre, M., "Galaxy groups and low mass clusters at z<0.6 : A perspective from the XMM Large Scale Structure survey", dans *XLIst Rencontres de Moriond : "From Dark Halo to Light"*, `ArXiv :astro-ph/0610800` (2006).

Pierre, M., **Pacaud**, **F.**. & le consortium XMM-LSS, "The XMM-LSS cluster sample and its cosmological applications", dans *proceedings of the ESA/ESAC symposium : "The X-ray Universe"*, `ArXiv :astro-ph/0511184` (2005).

Andreon, S., Willis, J., Quintana, H., Valtchanov, I., Pierre, M. & **Pacaud**, **F.**., "Colour-detected Clusters in the XMM Large Scale Structure Survey", dans Y. Giraud-Héraud & J. Tran Thanh Van (rédacteurs), *proceedings of the XXXIXth Rencontres de Moriond : "Exploring the Universe"*, p. 165–168, `ArXiv :astro-ph/0405574` (2004a).

Andreon, S., Willis, J., Quintana, H., Valtchanov, I., Pierre, M. & **Pacaud**, **F.**., "Galaxy evolution in clusters from z=1 to z=0", dans A. Diaferio (rédacteur), *IAU Colloq. 195 : Outskirts of Galaxy Clusters : Intense Life in the Suburbs*, p. 367–371, `ArXiv :astro-ph/0405384` (2004b).

**Pacaud**, **F.**., Pierre, M. & le consortium XMM-LSS, "The XMM-LSS cluster catalogue", dans Y. Giraud-Héraud & J. Tran Thanh Van (rédacteurs), *the XXXIXth Rencontres de Moriond : "Exploring the Universe", poster* (2004).

**Titre :** EXPLOITATION COSMOLOGIQUE DU RELEVÉ XMM-LSS

**Résumé :** Le modèle généralement admis de formation hiérarchique des structures de l'univers, prédit que les surdensités locales à petite échelle s'éffondrent en premier puis fusionnent pour former des systèmes de plus en plus larges. Dans ce contexte, les amas de galaxies se révèlent les structures les plus massives à avoir atteint l'équilibre et, à ce titre, fournissent de fortes contraintes cosmologiques, indépendantes et quasi-orthogonales à celles déduites du fond diffus cosmologique et des supernovae.

La détection de ces systèmes via l'émission X du gaz intra-amas permet d'assembler de larges échantillons faiblement contaminés et constitue à ce jour la méthode de sélection la plus robuste. Pour cela, le relevé XMM-LSS prévoit de cartographier 64 $deg^2$ du ciel avec XMM-Newton, le satellite X le plus sensible jamais construit, afin de constituer un échantillon sans précédent de plusieurs centaines d'amas jusqu'à z=1 et plusieurs dizaines de milliers de noyaux actifs de galaxies (AGN).

Cette thèse présente un ensemble de méthodes développées afin de sélectionner les sources étendues et les AGNs dans les données XMM, ainsi que des outils d'analyse détaillée pour les amas détectés. Les résultats obtenus avec les 5 premiers degrés carrés du relevé sont ensuite retracés : d'abord les propriétés de l'échantillon d'AGNs et l'analyse de leur corrélation angulaire ; puis, les implications de l'échantillon d'amas X pour l'évolution de la relation température - luminosité X, avec un accent particulier sur l'impact des effets de sélection, et enfin sa modélisation cosmologique dans le cadre du modèle hiérarchique.

Quelques comparaisons multi-longueurs d'onde ainsi que les perspectives pour un relevé plus étendu sont également considérées.

**Mots clés :** Cosmologie - Rayons X - Amas de galaxies - Détection - Effets de sélection - Noyaux actifs de galaxies

---

**Title :** COSMOLOGICAL EXPLOITATION OF THE XMM-LSS SURVEY

**Abstract :** The widely acknowledged model of hierarchical structure formation of the universe, predicts that local small-scale overdensities collapse first then merge to form increasingly larger systems. In this context, galaxy clusters constitutes the most massive structures that have reached equilibrium and, as such, provide tight cosmological constraints that are independent and quasi-orthogonal to those arising from the cosmic microwave background and the supernovae.

The detection of these systems through the X-ray emission of the intra-cluster medium allows the assembly of large, weakly contaminated samples and prevail to date as the most robust selection technique. For this reason, the XMM-LSS survey intends to map 64 $deg^2$ of the sky with XMM-Newton, the most sensitive X-ray satellite ever built, so as to gather an unprecedented sample of several hundreds of clusters up to z=1 and a few tens of thousands of active galactic nuclei (AGN).

This thesis presents a set of methods developed in the purpose of selecting extended sources and AGNs in XMM data, as well as detailed analysis tools for the detected clusters. The results derived from the first 5 square degrees of the survey are then depicted : first the properties of the AGN sample and the analysis of their angular correlation ; then, the implications of the X-ray cluster sample on the evolution of the temperature - X-ray luminosity relation, emphasizing the impact of selection effects, and finally its cosmological modelling in the framework of hierarchical model.

Some multi-wavelength comparisons as well as prospects for larger surveys are also considered.

**Keywords :** Cosmology - X-rays - Galaxy clusters - Detection - Selection effects - Active galactic nuclei