Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **406,** 673–688 (2010) doi:10.1111/j.1365-2966.2010.16720.x

3D-Matched-Filter galaxy cluster finder – I. Selection functions and CFHTLS Deep clusters

M. Milkeraitis,[1★] L. Van Waerbeke,[1] C. Heymans,[2] H. Hildebrandt,[3] J. P. Dietrich[4,5,6] and T. Erben[7]

[1] *University of British Columbia, Department of Physics and Astronomy, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada*
[2] *The Scottish Universities Physics Alliance, Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*
[3] *Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333CA Leiden, the Netherlands*
[4] *ESO, Karl-Schwarzschild-Str. 2, 85748 Garching b. München, Germany*
[5] *Physics Department, University of Michigan, 450 Church Street, Ann Arbor, MI 48109-1040, USA*
[6] *Michigan Center for Theoretical Physics, 450 Church Street, Ann Arbor, MI 48109-1040, USA*
[7] *Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany*

Accepted 2010 March 20. Received 2010 March 15; in original form 2009 December 7

ABSTRACT

We present an optimized galaxy cluster finder, 3D-Matched-Filter (3D-MF), which utilizes galaxy cluster radial profiles, luminosity functions and redshift information to detect galaxy clusters in optical surveys. This method is an improvement over other matched-filter methods, most notably through implementing redshift slicing of the data to significantly reduce line-of-sight projections and related false positives. We apply our method to the Canada–France–Hawaii Telescope Legacy Survey (CFHTLS) Deep fields, finding ∼170 galaxy clusters deg^{-2} in the $0.2 \leq z \leq 1.0$ redshift range. Future surveys such as LSST and JDEM can exploit 3D-MF's automated methodology to produce complete and reliable galaxy cluster catalogues. We determine the reliability and accuracy of the statistical approach of our method through a thorough analysis of mock data from the Millennium Simulation. We detect clusters with 100 per cent completeness for $M_{200} \geq 3.0 \times 10^{14}\,\mathrm{M}_{\odot}$, 88 per cent completeness for $M_{200} \geq 1.0 \times 10^{14}\,\mathrm{M}_{\odot}$ and 72 per cent completeness well into the $10^{13}\,\mathrm{M}_{\odot}$ cluster mass range. We show a 36 per cent multiple detection rate for cluster masses $\geq 1.5 \times 10^{13}\,\mathrm{M}_{\odot}$ and a 16 per cent false detection rate for galaxy clusters $\gtrsim 5 \times 10^{13}\,\mathrm{M}_{\odot}$, reporting that for clusters with masses $\lesssim 5 \times 10^{13}\,\mathrm{M}_{\odot}$ false detections may increase up to ∼24 per cent. Utilizing these selection functions we conclude that our galaxy cluster catalogue is the most complete CFHTLS Deep cluster catalogue to date.

Key words: methods: numerical – galaxies: abundances – galaxies: clusters: general – galaxies: luminosity function, mass function – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

Clusters of galaxies pinpoint the densest regions in the Universe. Housed in deep gravitational wells, clusters act as laboratories to study the influence of extreme environments on galaxy formation and evolution (Dressler 1980). The most massive clusters are also natural gravitational telescopes, lensing and magnifying light from the most distant of galaxies (Stark et al. 2007). As clusters trace the high-mass tail of the matter distribution, a complete sample can provide a very sensitive probe of cosmology (Battye & Weller 2003). Directly probing the growth of structure in the early Universe, cluster cosmology is fast becoming an important component of future dark energy surveys. As the future of optical astronomy ushers in large data sets of wide, deep sky coverage, it becomes necessary to develop automated algorithms that methodically and accurately search these data sets for galaxy clusters. Depending on scientific goals, it is desirable to have as complete a galaxy cluster sample as possible, over a range of redshifts; any intrinsic limitations of these search algorithms and the resultant biases introduced into generated galaxy cluster catalogues must be understood. The question of how to methodically select and quantify the completeness of a cluster sample is the subject of this paper.

Many different approaches to searching for galaxy clusters in astronomical data exist: current optical astronomy methods focus on finding clusters as overdensities via friends-of-friends algorithms such as Li & Yee (2008), density maps as in Adami et al. (2010) or Voronoi tesselation methods as in Lopes et al. (2004) and van Breukelen & Clewley (2009) for example. Other search methods look for large, red, elliptical cluster galaxy populations, and are

★E-mail: martham@phas.ubc.ca

© 2010 The Authors. Journal compilation © 2010 RAS

referred to as red sequence techniques (see Gladders & Yee 2000; Cohn et al. 2007; Kodama et al. 2007; Lu et al. 2009 and Thanjavur, Willis & Crampton 2009), perhaps also including the existence of bright central galaxies for example into the search algorithms (such as the maxBCG method of Koester et al. 2007 and Hilbert & White 2010). Alternatively, other methods search for clusters based on characteristic galaxy cluster luminosity and radial profiles (Postman et al. 1996; Kepner et al. 1999; Olsen et al. 1999a; White & Kochanek 2002; Kochanek et al. 2003; Gilbank et al. 2004; Dietrich et al. 2007; Grove, Benoist & Martel 2009; Menanteau et al. 2009). Each of these methods makes assumptions about general cluster properties and, resultantly, derived galaxy cluster catalogues will primarily include galaxy clusters that reflect those assumptions. Ideally a method that minimizes these biases and produces an understood and complete galaxy cluster catalogue will be best suited for statistical science of galaxy clusters and investigations into cosmology.

With this in mind, we present an optimized luminosity function and radial profile based optical galaxy cluster-finding algorithm that we call 3D-Matched-Filter (3D-MF). We present thorough tests of 3D-MF on simulations, and use the ascertained galaxy cluster selection functions to produce a galaxy cluster catalogue for the Canada–France–Hawaii Telescope Legacy Survey (CFHTLS) Deep fields. To this effect, Section 2 of this paper will discuss the Millennium Simulation catalogues used, Section 3 will discuss matched-filter methodology, as well as 3D-MF itself, and Section 4 will outline the selection functions of 3D-MF from the Millennium Simulation data (*without* and *with* photometric redshift error). Subsequently, 3D-MF will be run on the CFHTLS Deep data set and the resultant galaxy cluster catalogues will be discussed in Section 5.

A Lambda cold dark matter (ΛCDM) cosmology has been assumed throughout this paper: $H_0 = 73\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{M} = 0.25$, $\Omega_\Lambda = 0.75$ (which is also consistent with the Millennium Simulation used herein).

2 MILLENNIUM SIMULATION

2.1 Catalogues

The simulation catalogues used throughout this paper are six pencil-beam mock catalogues from Kitzbichler & White (2007) (hereafter KW07). These catalogues are light cones created from the semi-analytic galaxy catalogue from De Lucia & Blaizot (2007), which in turn is derived from the Millennium Simulation of the concordance ΛCDM cosmogony (Springel et al. 2005, hereafter S05).

The Millennium Simulation well reproduces galaxy clustering as a function of luminosity as shown in fig. 5 of S05. Furthermore, semi-analytic modelling applied to the original Millennium Simulation catalogues, in the De Lucia & Blaizot (2007) and KW07 catalogues, results in luminosity functions in similarly good agreement with observations. Consequently, these catalogues are well suited to test a cluster-finding algorithm based on luminosity, such as the 3D-MF galaxy cluster-finding algorithm presented in this paper. Conversely, there are difficulties in reproducing observational colours in simulation work (again, see fig. 5 from S05 and a discussion in De Lucia & Blaizot 2007). This fact hinders the analysis of colour-based cluster-finding algorithms, such as red sequence based techniques, on the Millennium Simulation. Cohn et al. (2007) tested their red sequence algorithm to examine changes with redshift using the Millennium Simulation. The authors found that the simulation colours did not match observations; they instead determined their red sequence by matching the simulations. In another recent example, Hilbert & White (2010) used the maxBCG method from Koester et al. (2007) (based on the red sequence technique) on the Millennium Simulation and found that the simulation did not reproduce the colours of passively evolving galaxies to the degree required for a direct application of colour–redshift relations derived from observations. Resultantly, they used altered colour–redshift relations, noting that if they had not adjusted the colour–redshift selection relations by hand, running maxBCG on the Millennium Simulation data set and subsequent modelling would have found almost no clusters with $z > 0.25$. The 3D-MF technique does *not* use Millennium Simulation colour information, and relies instead on the luminosity function of galaxy clusters; therefore, evaluating on the Millennium Simulation provides a robust test of the application of this method to real data.

Table 1. Parameters used to compile a list of known galaxy clusters from the KW07 mock light cones of the Millennium Simulation.

Millennium Simulation galaxy cluster traits
(I) Cluster members have the same friends-of-friends identification number (as described in the simulations), thus avoiding dark matter substructures
(II) Cluster members are flagged in the simulations with the M_{200} of their parent halo
(III) Clusters have $\geq$5 galaxy members

The KW07 mock catalogues used herein consist of six deep fields of $1^\circ\!.4 \times 1^\circ\!.4$ and mark the first time these light cones have been used in cluster-finding work. 3D-MF uses redshift information: this is not obtained from Millennium Simulation colours. Magnitude information is used from KW07 but only for the purpose of calculating luminosity functions as described. The SDSS magnitudes were converted to MegaCam magnitudes (according to equation 1[1]) in anticipation of the CFHT data that 3D-MF will first have the opportunity to examine in Section 5:

$$i_\mathrm{MegaCam} = i_\mathrm{SDSS} - 0.085\,(r_\mathrm{SDSS} - i_\mathrm{SDSS}). \quad (1)$$

Hereafter we refer to i_MegaCam as i'-band.

2.2 Millennium Simulation galaxy clusters

A common definition does not exist in the literature with which to define a cluster, other than the agreed upon fact that a cluster consists of multiple galaxies gravitationally bound to one another orbiting a common centre of mass. Smaller numbers of galaxies bound together are more traditionally known as *groups*, whereas larger galaxy associations are typically called *clusters*; it is the border between groups and clusters that is not always strictly defined. For the purposes of this paper, a cluster will be defined in the Millennium Simulation as a collection of galaxies that belong to the same parent halo and have $\geq$5 members (where membership is defined by examining the Millennium Simulation catalogues to $z < 1.2$); traits of *clusters* are listed in Table 1. Note that M_{200} is the mass enclosed within radius r_{200}: the radius within which the mean density is 200 times the critical density at that redshift. Based on

[1] The Canadian Astronomy Data Centre, Herzberg Institute of Astrophysics, MegaCam *ugriz* filter set as per http://www2.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/megapipe/docs/filters.html.

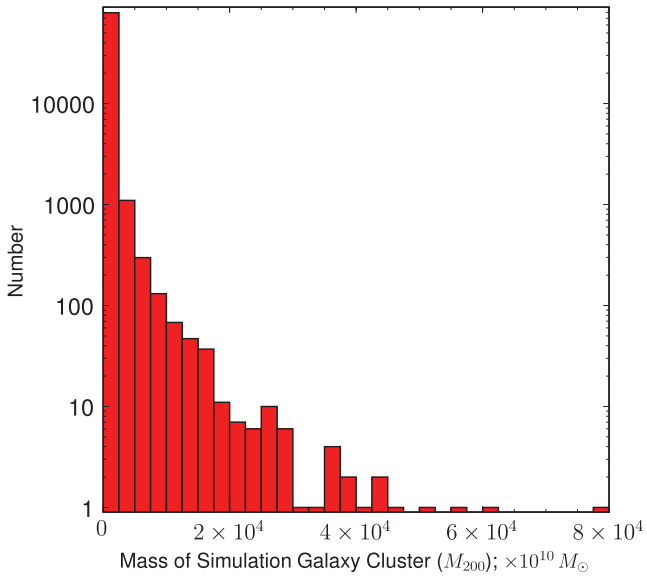

Figure 1. Mass distribution of Millennium Simulation KW07 galaxy clusters (following the cluster definition given in Table 1, and showing the cluster redshift range of $0.2 \leq z \leq 1.0$ as explained in the text).

this galaxy cluster definition, the mass distribution of KW07 galaxy clusters is shown in Fig. 1. The redshift distribution of these *known* clusters is presented in Fig. 2.

The centre for each known Millennium Simulation cluster is defined in this paper as the average position of the brightest 10 (or average of all members when the number of galaxies per cluster is <10) i'-band objects per cluster. The position of the centre is only used when matching the clusters found by 3D-MF to those known in the Millennium Simulation to determine which clusters were found, and a tolerance range is used (described in Section 4); it is not crucial to immediately determine a more exact position of the cluster centre and it will be examined further in Section 4.3.2.

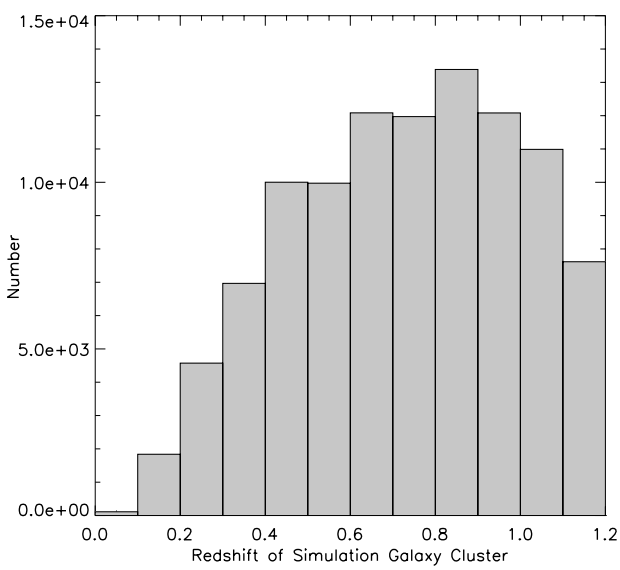

Figure 2. Redshift distribution of Millennium Simulation KW07 galaxy clusters (following the cluster definition given in Table 1, and noting the catalogue light cones extend much further than a redshift of 1.2: higher ranges are not shown here).

3 THE 3D-MF ALGORITHM

3.1 The foundations

The 3D-MF galaxy cluster-finding algorithm bases its search on the luminosity and radial profile of a galaxy cluster, appropriately sized for the redshift of the cluster. Prior to implementing our changes (described in Section 3.3), this method was based on Postman et al. (1996, hereafter P96) and the reader is directed there for a more in-depth discussion of the background of this technique. The method can use any sensible choice of luminosity or radial profile to model a galaxy cluster: 3D-MF follows P96 and currently uses a modified Schechter luminosity function (equation 2) and truncated Hubble radial profile (equation 3) to describe a fiducial galaxy cluster:

$$\Phi(M) = 0.4 \ln 10\, \Phi^* 10^{0.4(\alpha+1)(M^*-M)} \exp[-10^{0.4(M^*-M)}], \tag{2}$$

where Φ is the galaxy luminosity function, Φ^* sets the overall normalization of galaxy density, M is absolute magnitude and α is the slope of the faint end of the luminosity function. M^* is a characteristic absolute magnitude: Φ drops with increasing luminosity but at magnitudes brighter than M^* the exponentially decreasing slope of the luminosity function cuts off dramatically. The integral of the Schechter luminosity function diverges for $\alpha < -1$ at magnitudes fainter than M^*; the multiplicative term, $\exp[-10^{0.4(M^*-M)}]$, that has been added to modify the Schechter equation, is a dimensionless power-law cut-off (weighted by the flux of the galaxy) that allows the function to remain integrable.

The radial profile of the cluster is modelled as a projected truncated Hubble radial profile:

$$P\left(\frac{r}{r_c}\right) = \frac{1}{\sqrt{1+(r/r_c)^2}} - \frac{1}{\sqrt{1+(r_{co}/r_c)^2}}, \tag{3}$$

where r is the distance of a galaxy from a cluster centre, r_c is the cluster core radius and r_{co} is the cut-off radius. r_{co} is arbitrary and must be chosen such that it extends well beyond the cluster's core radius, importantly allowing non-circularly symmetric clusters to still be detected by this technique.

Once the fiducial cluster has been established by choice of luminosity and radial profile, the data are searched in one wavelength band for areas that maximally and simultaneously match both profiles, thereby *filtering* galaxy clusters out of the data (and thus the terminology 'matched-filter'). In previous applications of matched-filter methodologies, the data have not been binned per redshift as photometric redshift information has only recently become commonplace. In the absence of photometric redshift information in the past, the profiles that were being matched were themselves redetermined with an assumed cluster size and M^* value at a series of trial redshifts.

The likelihood, $\mathcal{L}$, that galaxies within r_{co} match the luminosity and radial profile of a fiducial cluster at a particular assumed redshift is calculated according to equation (4):

$$\ln\, \mathcal{L} \propto \int P\left(\frac{r}{r_c}\right) \frac{\Phi(m-m^*)}{b(m)} D(r,m)\, \mathrm{d}^2 r\, \mathrm{d}m, \tag{4}$$

where m is the apparent magnitude of a galaxy, m^* is the apparent magnitude corresponding to the characteristic luminosity of cluster galaxies, which incidentally includes a redshift-dependent k-correction, $b(m)$ is the background galaxy count and $D(r,m)$ is the total number of galaxies at a given magnitude and distance from a cluster centre.

This likelihood is calculated for every possible cluster centre in an input catalogue and can be viewed as an array or image map;

these maps will be referred to as likelihood maps. The likelihood maps are searched for peaks, which are galaxy cluster detections.

The significance of a peak, or galaxy cluster detection, is measured according to equation (5):

$$\sigma = \frac{S_{\rm p} - S_{\rm bg}}{\sigma_{\rm bg}}, \tag{5}$$

where $S_{\rm p}$ is the peak signal and $S_{\rm bg}$ is the background signal calculated by binning the distribution of all likelihood values in a likelihood map and taking the mode of this Gaussian distribution. The background dispersion, $\sigma_{\rm bg}$, is $0.741 \times (Q_3 - Q_1)$ where Q_1 and Q_3 are the first and third quartiles of the ranked likelihood value distribution. Cluster detections can be filtered based on their significance, ensuring detections are well above the background noise, and in this way a galaxy cluster list can be generated.

3.2 2D false detection rates

False detections are defined as galaxy clusters found by an algorithm that does not exist in reality: an important parameter, among others, by which to define the purity of an algorithm. Older versions of matched-filter techniques have uncomfortably high false detection rates, a feature largely attributable to line-of-sight projections.

P96 test their matched-filter algorithm on a self-created simulation (mimicking the Palomar Distant Cluster Survey) and find a 65 per cent false detection rate of galaxy clusters with peak signal thresholds $\geq 3\sigma$. P96 reduce this false detection rate to 31 per cent by adding the criteria that clusters must be found by their algorithm twice, looking separately in two colour bands. The European Southern Observatory (ESO) Imaging Survey had some success in finding clusters using their implementation of this same algorithm, but it was initially developed with the purpose of preparing a list of candidate clusters for follow-up observations and not for producing a well-defined sample for statistical analysis (Olsen et al. 1999a, hereafter O1; Olsen et al. 1999b, hereafter O2). These ESO Imaging Survey cluster papers, and subsequent work with their algorithm (Dietrich et al. 2007, hereafter D07; Olsen et al. 2007, hereafter O3; Grove et al. 2009, hereafter O4), have not tested their version of the matched-filter algorithm on simulations in terms of its ability to recover known clusters, other than to very generally quantify the ability of their finder to avoid noise detections. Furthermore, these papers aim for the same galaxy cluster density as the Palomar Distant Cluster Survey in P96 (i.e. O3 quotes $\sim$52.1 clusters deg^{-2}), and use the same method, implying a similar false detection rate to that quoted above (O3 claims 16.9 ± 5.4 false detections deg^{-2}, or 33 per cent).

Kepner et al. (1999) implement a matched-filter algorithm based on P96 as well, but do not examine cluster recovery numbers. Instead, they focus on the accuracy of their algorithm to detect a given cluster at its proper redshift, or the accuracy of determining the proper richness of a given cluster. White and Kochanek implement a matched-filter algorithm based on Kepner's version (White & Kochanek 2002, hereafter WK02) and test on self-created simulations to examine their cluster-finding abilities; however, they find their simulations contain too few galaxies compared to real observational surveys: in fact, 60 per cent too few galaxies for an R-band magnitude $<$24. This limitation affects complications caused by line-of-sight projections, as well as luminosity functions integral to the matched-filter methodology, and thus efficiency rates of their algorithm compared to real data are reported as conservative estimates. Kepner et al. (1999) and WK02 use different likelihood functions from equation (4) in this paper, as well as different peak detection methods, and the reader is directed to those papers for more details. WK02 make no adjustments to improve cluster detections at the edges of their data. These edge effects become important when using cluster finders on real data: masking bad pixels or bright stars is necessary in reality and clusters on the borders of the masked regions, if not properly weighted or accounted for in terms of the fraction available to the search algorithm, will likely be missed. In terms of false detection reports, WK02 show how their false detection rates change with redshift: for their deepest self-created simulation of $1^\circ\!.5 \times 1^\circ\!.5$ out to redshift 1.0, with a limiting R-band magnitude of 24, they have false positive rates of $>$40 per cent (averaging $\sim$60 per cent for all redshift bins) for clusters $\geq 2 \times 10^{14}\,{\rm M}_\odot$.

Proponents of other cluster-finding techniques, such as red sequence based methods, often comment upon older matched-filter methods' high false detection rates as a reason to avoid this technique (e.g. Gladders & Yee 2000 or Gilbank et al. 2004). However, the false detection rates quoted above are largely caused by line-of-sight projections in the data; removing this contamination will lead to much improved false detection rates and push down the galaxy cluster mass limit to which reliable detections can be made. Arguably, red sequence based methods have their own line-of-sight projection issues. Gladders & Yee (2000) for example introduced their red sequence based cluster finder by testing it on known CNOC2 clusters (i.e. not simulations, but real data), and found excellent false positive rates of $<$5 per cent for clusters at redshifts $<$0.5. However, as WK02 also show, false positives increase with redshift. The Gladders & Yee (2000) CNOC2 sample testing provides limitations in the higher redshift ranges: notably a 36 per cent false detection rate in the redshift range $0.5 < z < 0.6$ (where authors suggest spectroscopic data were incomplete and did not deduce anything further). Furthermore, Koester et al. (2007), using their maxBCG method (based on the red sequence method), find 16 per cent of found galaxy clusters with shallow redshifts ($0.1 < z < 0.3$) are in reality line-of-sight projections (as tested on the observational data set of the SDSS). As all optical cluster finders have false positive rates that will increase with redshift, it becomes a goal of this paper to minimize these false detections as much as possible.

Intrinsic biases will exist to some degree in all methods. For example, red sequence based methods have their own intrinsic biases about galaxy clusters, notably assuming all galaxy clusters have large enough red sequence populations to be detected based on red sequence alone. Complimentary galaxy cluster-finding techniques with different biases are therefore necessary, first to determine whether clusters can be found that are missed by other techniques, and second to be able to ultimately construct a galaxy cluster list, from observational data, that is as complete as possible, and whose biases are well understood (via testing on simulations).

3.3 3D-MF's advancements

Using the foundations of the method discussed above in Section 3.1, 3D-MF advances the matched-filter technique by incorporating photometric redshift information, and adjusts additional elements as described below, to update and improve galaxy cluster finding. With the requirement that all next-generation surveys for cosmology come with photometric redshifts, it is natural for 3D-MF to extend the matched-filter technique into this third dimension.

3.3.1 Optimal use of photometric redshifts

3D-MF cuts the input data by redshift, initially into non-overlapping redshift slices. Each slice is then matched against appropriate luminosity and radial filters scaled for that particular redshift slice. The resultant likelihood maps (described in Section 3.1 with equation 4) are run individually through a new 3D-MF peak detection pipeline to identify galaxy clusters. Upon completion of a range of sequential redshift slices the entire process is repeated with a shift in the slices (thus overlapping the redshift slicing with the initial run) in order to account for clusters or structures that exist on borders of the original slicing. For example, an input catalogue could be sliced from redshift 0.00 to 1.20 (with a bin width of 0.2, but any desired bin width could be chosen) and these would be considered the *original* slices. After running 3D-MF on these original slices, the process is repeated by slicing the whole input catalogue differently, from redshift 0.10 to 1.30 (with the same bin width for example), and these would be considered the *shifted* slices. The resulting galaxy cluster lists from a complete run of original slices and a complete run of shifted slices are then merged (the details of which are presented in Section 4). Clusters that are multiply detected can be tracked, and false detections in either the original or shifted slices can be lowered by associating them properly with clusters from the shifted or original slices, respectively. The aim of 3D-MF is to optimize cluster detections through redshift slicing, thus eliminating line-of-sight projections, but also to take into account the fact that each cluster spans a range of space and subsequently redshift slicing has to be applied carefully to not cut these clusters and count them multiple times, or conversely to miss them entirely. In optimizing 3D-MF a variety of slicing styles were examined (various slice widths, non-uniform slice widths, differing overlapping slice amounts, etc.), and the slicing method that produced the best results (as seen in Section 4.2.1) used redshift slice widths of 0.2, with shifted slices overlapping original slices by a redshift of 0.1.

3.3.2 Masks

3D-MF accounts for masked regions of data (obscured by bad pixels, star diffraction spikes or data regions of low signal-to-noise ratio for example) by rejecting potential cluster candidates that are masked above a set per cent (given as an input threshold, chosen here to be 50 per cent). In cases where an important fraction of a cluster's luminosity function is missing due to a mask, it simply cannot be rebuilt, and missing optical information such as this is a limitation of optical galaxy cluster-finding algorithms, including matched-filter methods.

Partially masked clusters are weighted by 3D-MF according to the fraction of the cluster that is masked (equation 6):

$$\mathcal{L}_{\text{Mask weighted}} = \mathcal{L}\,(1 + \text{percentage masked}/100), \tag{6}$$

where the likelihood, $\mathcal{L}$ (from equation 4), is up-weighted by the percentage of the redshift-scaled radial filter that is masked, giving $\mathcal{L}_{\text{Mask weighted}}$.

The mask-weighted likelihood maps, $\mathcal{L}_{\text{Mask weighted}}$, increase the likelihood that a partially masked cluster matches the fiducial cluster at a given redshift, as its luminosity and radial profiles are up-weighted to compensate for the information missing due to the presence of the mask. Importantly, mask-weighting also improves detections of clusters near the edges of an image as clusters previously missed in these regions were found after the addition of this criterion.

3.3.3 Background galaxy counts

The requirement for an accurate measure of the background galaxy counts [$b(m)$ in equation 4] is an issue in P96 and other implementations of matched-filter such as O1–O4 discussed in Section 3.2. These older implementations of matched-filter methodology assume a priori that $b(m)$ can be modelled by a power law in m, or they determine $b(m)$ by creating self-simulated backgrounds (following the prescription of Soneira & Peebles 1978). 3D-MF determines the background counts by binning the magnitude distribution of the input data and subsequently interpolates within this measured distribution to determine $b(m)$ for each galaxy (i.e. pulling out an interpolated background number count for a particular galaxy magnitude), including it as a data-driven $b(m)$ in the integral over cluster galaxies in equation (4). We find that this accurately models the likelihood of a cluster signal against a true data-derived background and increases a cluster's likelihood of detection with the advantage that our method does not rely on biased modelling of $b(m)$.

3.3.4 Peak detection

Previous matched-filter methods (implementations of P96 such as D07 and O1–O4) detect peaks in the likelihood maps using SExtractor; while SExtractor is optimized for the detection of isolated peaks, it tends to blend things. Although this is excellent for object detection, the likelihood maps created with a matched-filter algorithm are more akin to continuous maps and thus need an appropriate peak detection method. Furthermore, SExtractor is parameter dependent; we desire a robust method that withstands reasonable changes in parameter space. 3D-MF detects all peaks above a user-input threshold (such as $\geq 3.5\sigma$) and thereby generates a complete list of possible cluster candidates.

3.3.5 Assigned cluster redshift

Other versions of the matched-filter technique assign a redshift to a galaxy cluster where it is maximally detected (i.e. the redshift at which the fiducial cluster is maximally matched) but in practice this often assigns an incorrect redshift to the cluster because all data are considered at all redshifts. Several other factors also contribute to this incorrect redshift assignment: among them the assumed background counts discussed above, the k-correction (and the effect on this of the assumed evolution of various galaxy types) and the fact that the cluster signals become overcorrected via the normalization of the luminosity filter (as in D07 and O1–O4). These methods apply a cluster signal correction factor to counteract some of these points, but resulting redshift assignments are still, on average, incorrect and thus are not robust (nor do they claim to be; see O1 and O2).

A plethora of papers followed up on the clusters found in O1 and O2 with the goal of obtaining spectroscopic redshifts of the matched-filter-found cluster members (Ramella et al. 2000; Hansen, Olsen & Jørgensen 2002; Benoist et al. 2002; Olsen et al. 2003, 2005a,b; Grove, da Costa & Benoist 2008). Not all clusters were conclusively found in the spectroscopic analysis; of the ones that were, the spectroscopic redshifts for these systems were compared with their matched-filter-found redshifts. Unfortunately, number counts are too low to ascertain anything statistically (i.e. the sample size available to compare spectroscopic and photometric redshifts was often only a few galaxy clusters per paper).

Also investigating redshift comparisons, galaxy clusters in D07 have redshifts derived from their matched-filter implementation and a sample of these is compared to known spectroscopic redshifts from the literature. These authors find their matched-filter has a bias towards lower redshifts and conclude they are underestimating the true redshift of galaxy clusters; a discrepancy that increases with higher redshifts.

As 3D-MF requires photometric redshift information for all of the objects in a catalogue and slices the catalogue according to redshift, once each galaxy cluster is found, its associated redshift slice is automatically, immediately and more accurately known. For example, for a redshift bin width of 0.2, with overlapping bins shifted by a redshift of 0.1 from the original bins, consider a cluster that is first detected in an original bin: the cluster is assigned the redshift of the centre of the bin and this cluster centre is thereby known to within ± 0.1 in redshift (i.e. the width of the bin). After repeating the cluster-finding process with shifted slices (as described above in Section 3.3.1), this cluster may also be detected in an overlapping shifted slice, localizing the redshift of its centre even further (to roughly ± 0.05 in redshift: the width of the overlap region of the original and shifted slices). This of course assumes that photometric redshifts are without error and unbiased. Knowing a cluster's galaxy members (and their redshifts) leads to an even more localized determination of a cluster's redshift; this will be discussed in relation to 3D-MF in Section 4.3.1.

3.3.6 Parameter optimization

3D-MF has a variety of parameters that can be optimized for maximal galaxy cluster detection. Using Millennium Simulation data to test 3D-MF has proven irreplaceable in its contribution to improving 3D-MF, in particular determining optimal parameter settings to apply to real data, assuming a similar cosmology. A discussion of these parameter settings follows in Section 4.

4 3D-MF ON MILLENNIUM SIMULATION MOCK LIGHT CONES

In this section we present the application of the 3D-MF algorithm to the Millennium Simulation KW07 mock catalogues. We first assume that exact redshifts are known for all galaxies, and then investigate the impact of adding photometric redshift errors to the data, on 3D-MF's selection functions. When matching the 3D-MF-found clusters to known mock clusters, we will use the known clusters above a mass of $1.5 \times 10^{13}\,\mathrm{M}_{\odot}$, and focus on the redshift range $0.2 \leq z \leq 1.0$ for reasons that will be explained.

4.1 3D-MF parameters

The Millennium Simulation mock light cone catalogues were cut at an i'-band magnitude of 25.5 to simulate a magnitude-limited survey. The catalogues were then cut into optimal redshift slices (see Section 4.2.1): redshift slice widths of 0.2 and shifted slices shifted from original slices by 0.1 in redshift (i.e. original and shifted slices overlap by 0.1 in redshift). The mock catalogues were then analysed with 3D-MF, as described in Section 3, using the parameters listed in Table 2.

M^* and α (equation 2) were obtained by fitting a Schechter luminosity function to mock galaxies and found to be $M^*_{i'\text{-band}} = -20.69 \pm 0.24$ and $\alpha = -1.114 \pm 0.028$. We optimize 3D-MF to find clusters with masses $>10^{14}\,\mathrm{M}_{\odot}$ by setting the cut-off radius, r_{co}, to be 1 Mpc based on the r_{200} radius of a $3 \times 10^{14}\,\mathrm{M}_{\odot}$ cluster.

Table 2. 3D-MF basic run-time parameters (consult the text for descriptions of each parameter).

3D-MF input parameters	
Schechter function (equation 2):	
$M^*_{i'\text{-band}}$	-20.69
α	-1.11
Hubble profile (equation 3):	
r_{c}	0.1 Mpc
r_{co}	1 Mpc
Likelihood map (equation 4):	
Scale	0.124 arcsec pix^{-1}
Peak detection:	
Mask weighting	$\mathcal{L}$ (1 + percentage masked/100)
Cluster rejection	if $\geq$50 per cent masked, reject
Minimum significance (equation 5) for real detection	$\geq 3.5\sigma$
Multiple detections:	
2D grouping (RA, Dec.) in same redshift slice	1.5 Mpc diameter
Grouping between redshift slices	$z_{\text{3D-MF}} \pm 0.1$

This is sufficiently large to enclose non-symmetrically shaped clusters and thus allow them to be detected, but still small enough to maximally prevent neighbouring or on-the-verge-of-merging galaxy clusters from becoming blended into one cluster detection. P96 explored a range of cut-off radii and found the significance of their cluster detections dropped by up to 40 per cent when increasing r_{co} beyond 1 Mpc, but did not test decreasing this window size significantly within 1 Mpc. However, P96 notes that when cluster dimensions are significantly larger than the cut-off radius, the signal will obviously be truncated, though the degree depends on cluster shape. In the Millennium Simulation analysis that follows we find that our choice of r_{co} is very successful at detecting clusters of masses $>1 \times 10^{14}\,\mathrm{M}_{\odot}$ and has a very useful by-product of also detecting clusters with masses as low as $\sim 10^{13}\,\mathrm{M}_{\odot}$. As the set cut-off radius for these low-mass clusters is significantly larger than their extent, it will be difficult to interpret our results for these objects as we discuss further in Section 4.3. Future implementations of 3D-MF will experiment with optimizing low-mass detections using a varying cut-off radius.

4.1.1 Single galaxy cluster detection criteria

Following 3D-MF's search of the Millennium Simulation mock light cone catalogues, the output detected galaxy cluster list was compared with the known Millennium Simulation galaxy clusters (see Section 2.2). To match a galaxy cluster detection with known galaxy clusters, a cut was first made in a projected 2D radius, set at $0\overset{\circ}{.}044$. This value was chosen due to the fact that $0\overset{\circ}{.}044$ at a mid-redshift range, $z \sim 0.55$, corresponds to $\sim$1 Mpc, the assumed cluster cut-off radius (r_{co}) for an individual cluster as per Table 2, and notably at the higher redshift end, $z = 0.95$, $0\overset{\circ}{.}044$ corresponds

to ∼1.25 Mpc remaining a reasonable choice there as well. Note that sizing this tolerance per redshift did not significantly change results. Therefore, any galaxy cluster detections within 0°.044 of a known Millennium Simulation cluster were considered possible *candidates*. From the list of *candidates*, each 3D-MF-found cluster was matched in 3D space (RA, Dec., z) to the closest known cluster and that known cluster is thus considered detected. The redshift component of this matching was performed within ranges of $z_{\text{3D-MF}} \pm 0.1$ for redshifts >0.6 (where $z_{\text{3D-MF}}$ denotes the 3D-MF-derived cluster redshift, and ±0.1 indicates that this matching was done within a redshift slice), and $z_{\text{3D-MF}} \pm 0.1 \pm 10$ per cent for $z \leq 0.6$. Note that lower redshift clusters will be spread out in redshift bins more appreciably than higher redshift clusters in relation to the bin volume; we wanted to avoid missing proper matches between 3D-MF clusters and known mock clusters seemingly spread out due to this effect and thus a slight widening of the matching redshift parameter for lower redshift clusters was chosen.

4.1.2 Multiple galaxy cluster detection criteria

As per the discussion in Section 2.2, it can easily become complicated to define a galaxy *cluster*. For example, in some cases it can become difficult to realistically determine whether a multiply detected galaxy cluster is either (i) one larger known galaxy cluster detected multiple times (for instance consider a merging system in a slightly 'dumbbell' shape that still has two clear lobes from the individual galaxy clusters, each lobe separate and large enough to be detected individually despite the chosen r_{co} value) or (ii) from one detection of a cluster and a second detection is instead from smaller known nearby clusters in the lower mass ranges. For this reason, we chose to track multiple detections in 3D space, and apply grouping criteria. Grouping 3D-MF detections to 3D-MF detections by a physically motivated linking length, and calling these linkages *multiple detections*, we can track multiple detections within each redshift slice, as well as between redshift slices. We wanted to avoid falsely assigning detected clusters to the wrong known galaxy clusters and to track multiply detected clusters at all mass ranges separately from false detections. Resultantly, any 3D-MF detections within a pre-determined linking length of each other were assigned a similar grouping identification number. This linking length was chosen to be 0.75 Mpc (in two dimensions: RA and Dec.) based on the physically motivated assumption that if one 3D-MF detection could be exactly at the correct known centre position, and noting that clusters typically span 1.5 Mpc in diameter, this would put a detection at 0.75 Mpc away from the correct centre as a maximal radius by which to associate detections. The third dimension of linking length, redshift, was selected to be $z_{\text{3D-MF}} \pm 0.1$ (i.e. joint membership in an original slice and an overlapping shifted slice) and was not chosen to be 0.75 Mpc as in the 2D linking, simply because the resolution in redshift space is unfortunately not this high. WK02 implement a similar method and drop cluster candidates when a more likely cluster detection exists within a projected separation of 1 h^{-1} Mpc and redshift difference of ±0.05. 3D-MF's multiple detection grouping criteria are summarized in Table 2. It is important to note that 3D-MF detections were associated to each other and the known cluster centres were not used in this step, because we wanted to develop a method that would transfer to real data where centres are always unknown.

The *best* match in a multiple detection grouping is chosen to be the one that is closest to a known simulation cluster centre. We point out to the reader that this choice is a difficult one, as there are known issues regarding centroiding; how one decides the true centre of a cluster is an interesting science topic many papers devote themselves to in their entirety. We will provide evidence in Section 4.3.2 that serves to support our choice of best match, and plan to investigate multiple matches and centroiding in future work. Our selection function plots are free of multiple detections not considered best matches as we believe these are duplicate detections of single clusters.

4.1.3 False detection criteria

False detections were those 3D-MF cluster detections that did not match to any known galaxy clusters after the above matching and grouping criteria were applied. As described, the cluster-finding process is repeated with original and shifted slices. In order to ensure each false detection was not a 'faint' detection of a real structure (e.g. to determine whether a false detection was actually found more significantly in a different redshift slice, perhaps due to a cutting of the cluster), and to merge original and shifted slices, the false detections from the original slices were cross-referenced in RA, Dec. and redshift with real cluster detections from overlapping shifted slices (and vice versa). The remaining false detections will be quantified and are likely due to noise fluctuations or subsisting projection effects.

4.2 Modelling photometric redshift errors

Photometric redshift errors were modelled for this paper by examining the area of overlap between the CFHTLS Wide field data and the VIMOS Very Large Telescope (VLT) Deep Survey (VVDS; Le Fèvre et al. 2005) and the DEEP2 spectroscopic survey (Davis et al. 2007). Photometric redshifts were estimated for the CFHTLS Wide data from images provided by the CFHTLS-Archive-Research Survey (Erben et al. 2009) with the method described in Hildebrandt et al. (2009), and those photometric redshifts were then compared to secure spectroscopic redshifts. This process yields an accuracy of the photometric redshifts for galaxies with magnitudes i'-band < 24 of $\sigma_{\Delta z/(1+z)} \approx 0.047$ after rejection of 2.8 per cent of outliers.

We confine the redshift range of galaxies to $z \leq 1.1$ as that is the region where our photometric redshifts are most accurate and not badly affected by biases. Common objects in the overlap regions of both 'CFHTLS and VVDS' and 'CFHTLS and DEEP2' were binned according to magnitude and redshift and the photometric redshift error [i.e. the rms of ($z_{\text{photometric}} - z_{\text{spectroscopic}}$) after rejection of outliers] for each bin was calculated and weighted by the bin number counts. A 3D surface (magnitude, photometric redshift and photometric redshift rms) was then found to be best fitted with a two-degree polynomial: the rms error of any new magnitude and redshift can then be extrapolated from this model (including an extrapolation of errors for magnitudes such as i'-band > 24 where deeper spectroscopic redshift catalogues do not currently exist to allow more accurate modelling of this high-magnitude region). Using the Millennium Simulation mock light cone data, a galaxy's magnitude and exact redshift was used in conjunction with the aforementioned rms model to calculate a typical observational photometric redshift error rms for that particular galaxy, from which a Gaussian distribution was created. The assumption of a Gaussian error distribution is supported by the comparisons between photometric redshifts and spectroscopic redshifts mentioned above (no significant secondary peak or skewness was found as confirmed by the low number of outliers stated above and an overall vanishing

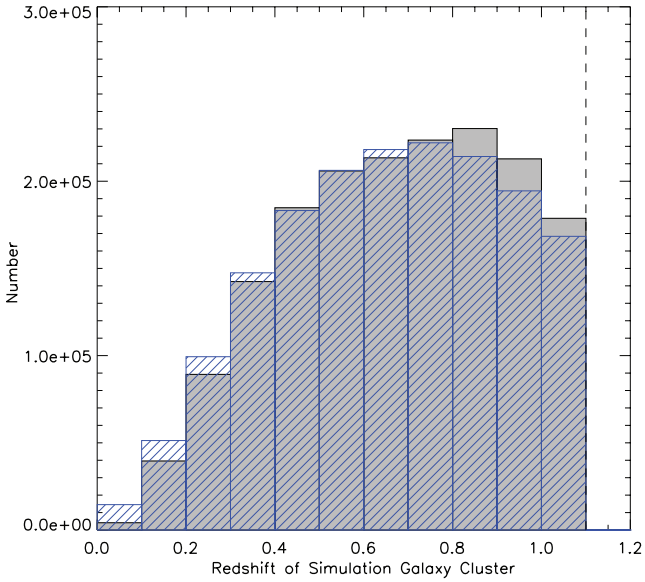

Figure 3. Redshift distribution of the six Millennium Simulation mock light cones (KW07) as exact redshifts (solid histograms) and with redshift errors modelled from the photometric redshift errors of Hildebrandt et al. (2009) (dashed histograms). Errors were derived from comparisons between the CFHTLS with VVDS and DEEP2 overlap regions; an explanation of the error modelling can be found in the text (Section 4.2). The data have been cut at a limiting i'-band magnitude of 25.5 to mimic real, observationally limited data, and are only shown out to a redshift of 1.1 as explained in the text.

photo-z bias, respectively). Randomly assigning a photometric redshift error to a given galaxy was then accomplished by randomly drawing an error value from the Gaussian distribution with the rms modelled for that galaxy. The redshift distributions of mock light cone galaxies before and after the redshift errors were applied are shown in Fig. 3 (where the data have also been cut at a limiting i'-band magnitude of 25.5 to mimic real observational data, and are only shown out to a redshift of 1.1 although, as mentioned, the light cones themselves are much deeper). The redshift error scales with redshift; selecting redshift slices of 0.2 in width ensures that our redshift slicing is larger than the rms of ∼0.1 at a redshift of 1.1.

4.2.1 Fine-tuning 3D-MF parameters

Investigating the use of various redshift slice widths with 3D-MF, as well as examining the effects of introducing redshift errors into the data, and the corresponding recovery rates of galaxy clusters, leads us to be able to fine-tune 3D-MF's parameters and select an optimal redshift slicing.

An examination of the completeness (the fraction of known clusters recovered by 3D-MF) as a function of mass *and* redshift slice width is essential to recover as many clusters as possible. Fig. 4(a) presents a few variations of redshift slice width as evidence that an analysis of the mock light cones with 3D-MF set to use redshift slice bin widths of 0.2 in size recovers the most known clusters with $M_{200} \geq 5 \times 10^{13}\,\mathrm{M}_{\odot}$. Repeating this analysis with photometric redshift errors shows that completeness tends to go up in almost all cases when running 3D-MF separately with various redshift slice widths (see Fig. 4b) but the significance of detections against the background goes down (as will be shown).

Maintaining a constant redshift slice width, it is of further interest to note the effect on completeness when the limiting magnitude of

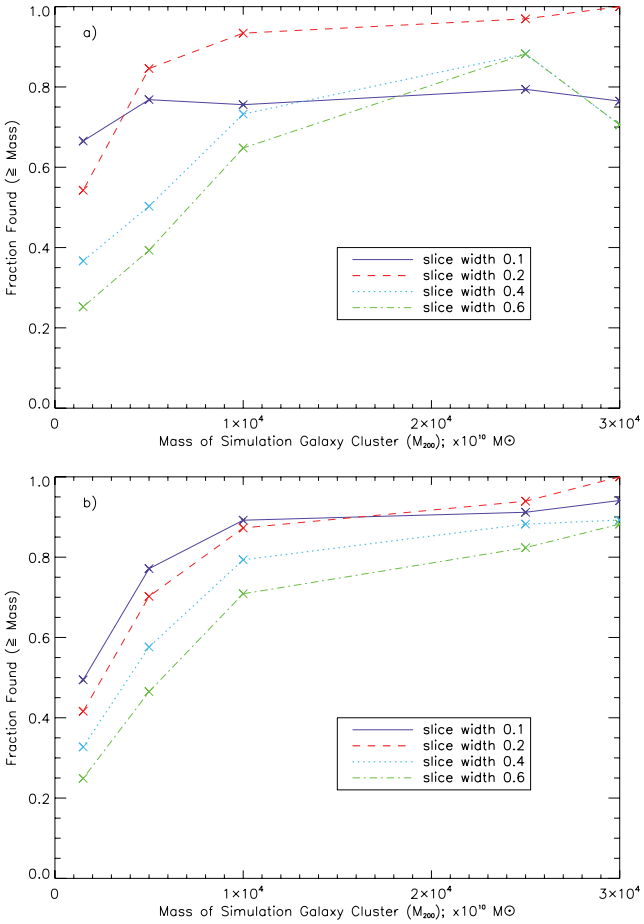

Figure 4. Completeness with respect to mass, measured as a fraction of known clusters found by 3D-MF. The input data were sliced into overlapping redshift slices as explained in the text. Overlapping slices are in each case shifted by half of the redshift slice width. Fig. 4(a) presents exact redshifts while Fig. 4(b) includes redshift errors in the data.

the survey is changed. Fig. 5 presents a constant redshift slice width of 0.2 with varying survey depths of 23.5–25.5 i'-band magnitudes. As expected, the ability to find all clusters is significantly reduced as the survey depth decreases: fewer galaxies are present in shallower data making their parent clusters improbable or impossible to optically detect. Interpretations associated with cluster masses below $\sim 5 \times 10^{13}\,\mathrm{M}_{\odot}$ will be addressed throughout Section 4.3 and the complications in this region will be addressed.

Returning to a discussion of significance levels, ignoring any detections below the 3.5σ level is a sensible decision: the probability of a 3σ positive curvature peak from a background Gaussian field, due to chance, is 4 per cent (see Van Waerbeke 2000). Effectively, desiring a contamination due to random fluctuations of less than 1 per cent, we have to be more conservative and accept only those cluster detections above 3.5σ. Fig. 6 investigates significance and clearly shows that false detections were not found at a high level of significance when compared to known clusters accurately detected by 3D-MF. Low false detection significances are encouragingly still seen in Fig. 6 with the introduction of photometric redshift errors (and Fig. 6 shows, as expected, that redshift errors reduce detection significances in general). If we cut at a higher significance level, such as 6.5σ, we could remove >95 per cent of our false detections, but we would also lose many lower significance, real detections; a trade-off must be reached and thus 3.5σ was chosen.

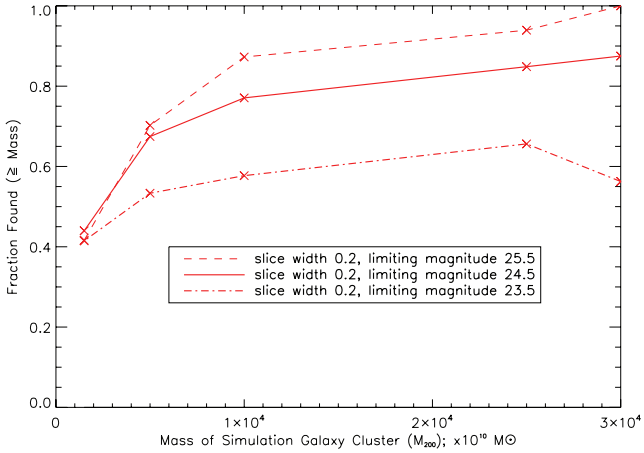

Figure 5. Similar to Fig. 4(b) (completeness with respect to mass, measured as a fraction of known clusters found by 3D-MF), but here the width of the redshift slices is held fixed and the effect (on the fraction of galaxy clusters found) caused by changing the i'-band limiting magnitude of the survey is shown.

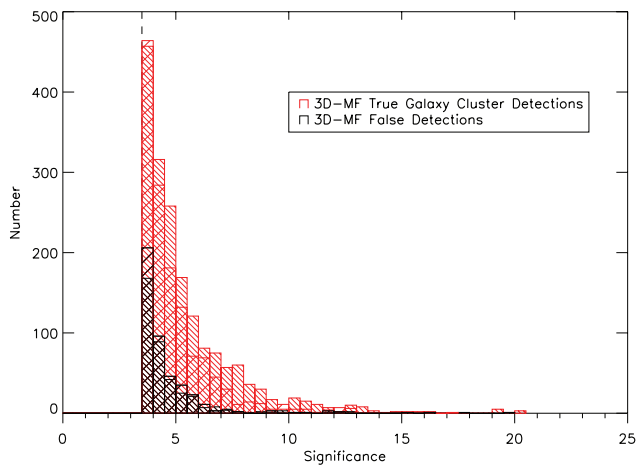

Figure 6. Comparing 3D-MF's real true, galaxy cluster detections with false detections as a function of detection significance (down-sloping lines represent no redshift errors, while up-sloping lines represent data with redshift errors).

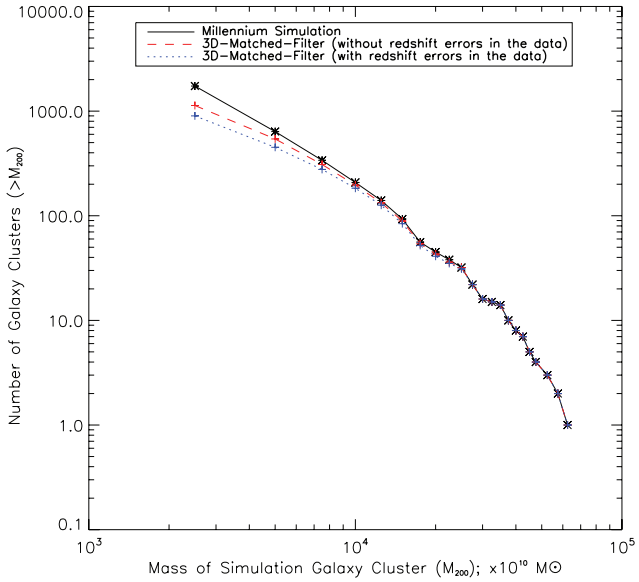

Figure 7. Cumulative mass function for galaxy clusters from the Millennium Simulation KW07 catalogue (as per Table 1; redshift errors described in the text).

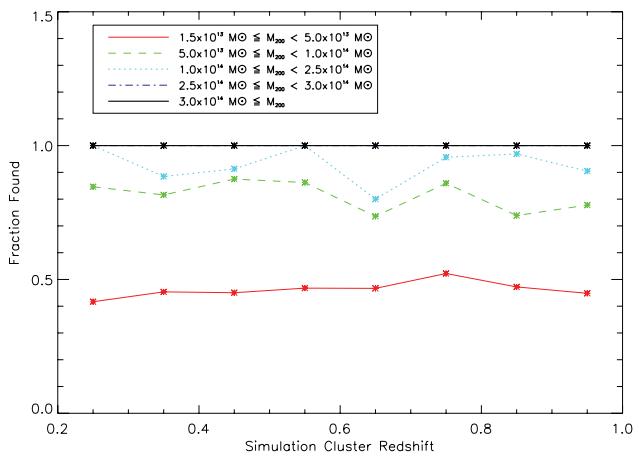

Figure 8. Completeness with respect to redshift, measured as a fraction of known clusters found.

4.3 Final selection functions

Having determined appropriate parameter settings, we now present 3D-MF's selection functions. The cumulative mass function of galaxy clusters recovered by 3D-MF is presented in Fig. 7. An examination of this figure shows the recovery rate of known clusters to be very accurate above an M_{200} of $\sim 5 \times 10^{13}\,\mathrm{M}_{\odot}$; results that are encouraging for 3D-MF, and furthermore are consistent for all redshifts (see Fig. 8).

Taking the optimal redshift bin width of 0.2 (as per Section 4.2.1) and examining completeness with respect to redshift results in fractional recovery rates of known clusters per redshift that are constant in each mass range, with excellent recovery (i.e. >88 per cent) for clusters with $M_{200} \geq 1.0 \times 10^{14}\,\mathrm{M}_{\odot}$. Completeness rates per mass, multiple detection rates and false detection information are summarized in Table 3. 3D-MF has a 100 per cent recovery for clusters with $M_{200} \geq 3.0 \times 10^{14}\,\mathrm{M}_{\odot}$, 97 per cent of clusters are found with masses above $2.5 \times 10^{14}\,\mathrm{M}_{\odot}$, 88 per cent recovery rates are seen for $M_{200} \geq 1.0 \times 10^{14}\,\mathrm{M}_{\odot}$ as mentioned and 72 per cent of clusters are found in even lower mass ranges (numbers quoted pertain to the presence of photometric redshift errors).

For clusters with redshifts in the range $0.2 \leq z \leq 1.0$ and $M_{200} \geq 1.0 \times 10^{14}\,\mathrm{M}_{\odot}$, the galaxy cluster number density in the Millennium Simulation KW07 catalogues is ∼18 galaxy clusters deg^{-2}; 3D-MF finds ∼17 galaxy clusters deg^{-2} (using exact redshifts) and ∼16 galaxy clusters deg^{-2} (using redshifts with errors) for this same redshift and mass range. Furthermore, for clusters with $1.5 \times 10^{13}\,\mathrm{M}_{\odot} \leq M_{200} < 1.0 \times 10^{14}\,\mathrm{M}_{\odot}$, the galaxy cluster number density in the Millennium Simulation KW07 catalogues is ∼258 galaxy clusters deg^{-2}; using exact redshifts 3D-MF finds ∼134 galaxy clusters deg^{-2}, and when photometric redshift errors are added, 3D-MF finds a cluster number density of ∼100 galaxy clusters deg^{-2}.

The multiple detection rate, measured as *the fraction of total 3D-MF detections matched to known clusters that are already matched to a 3D-MF detection*, for clusters with $M_{200} \geq 1.5 \times 10^{13}\,\mathrm{M}_{\odot}$, is 36.7 per cent when exact redshifts are used and 36.2

Table 3. Completeness (per cent of total known clusters found by 3D-MF) and multiple and false detection rates for 3D-MF on KW07 simulation galaxy clusters (photometric redshift errors described in the text). The detected galaxy clusters section reports the number of clusters detected at least one time, the multiple detections section reports additional detections, and the false detections section presents the fraction of total 3D-MF detections not matched to known clusters above $1.5 \times 10^{13}\,\mathrm{M}_{\odot}$.

Mass range	Number of known clusters	No photometric redshift errors Number	Completeness (per cent)	With photometric redshift errors Number	Completeness (per cent)
Detected galaxy clusters:					
$\geq 3.0 \times 10^{14}\,\mathrm{M}_{\odot}$	16	16	100.0	16	100.0
$\geq 2.5 \times 10^{14}\,\mathrm{M}_{\odot}$	32	32	100.0	31	96.9
$\geq 1.0 \times 10^{14}\,\mathrm{M}_{\odot}$	208	197	94.7	184	88.5
$\geq 5.0 \times 10^{13}\,\mathrm{M}_{\odot}$	637	548	86.0	456	71.6
$\geq 1.5 \times 10^{13}\,\mathrm{M}_{\odot}$	3240	1775	54.8	1358	41.9
Additional multiple detections:					
			Per cent of total cluster detections (per cent)		Per cent of total cluster detections (per cent)
$\geq 3.0 \times 10^{14}\,\mathrm{M}_{\odot}$		14	0.4	21	0.7
$\geq 2.5 \times 10^{14}\,\mathrm{M}_{\odot}$		28	0.8	39	1.4
$\geq 1.0 \times 10^{14}\,\mathrm{M}_{\odot}$		208	6.1	227	8.1
$\geq 5.0 \times 10^{13}\,\mathrm{M}_{\odot}$		504	14.8	439	15.6
$\geq 1.5 \times 10^{13}\,\mathrm{M}_{\odot}$		1253	36.7	1020	36.2
False detections:					
$\geq 1.5 \times 10^{13}\,\mathrm{M}_{\odot}$		383	11.2	438	15.6

per cent when redshift errors are added to the data. This multiple detection rate is what would be expected considering our redshift slices overlap by 0.1 in redshift: many clusters are included wholly in both original and shifted slices and, depending on their simulation redshift, are found by either preferentially matching our fiducial cluster (recall Section 3.1) in the original or shifted slices, or both.

False detection rates are reported as *the fraction of total 3D-MF detections not matched to known clusters above* $1.5 \times 10^{13}\,\mathrm{M}_{\odot}$; when simulations have photometric redshift errors added, the false detection rate increases a mere 4.4 per cent from the non-error photometric redshift case to 15.6 per cent false detections overall. Since photometric redshift errors scatter galaxies to random redshift slices they are not expected to induce increased clustering. False detections can be further analysed; they are shown in Fig. 9 to be a fairly uniform function of redshift, with the exception of the redshifts at the edges of our redshift range ($z = 0.2$ and 1.0). It is likely that some of these edge redshift false detections are in fact true detections of simulation clusters with redshifts <0.2 or >1.0. This needs to be investigated further, but note that regardless of the redshift range chosen, there will always be potential contamination from galaxy members of simulation clusters whose centres lie outside the range of interest but whose presence is enough to cause a 3D-MF detection. Recall that we had reasons to select this redshift range: introducing redshift errors into the simulations was confined to a redshift region where photometric redshift errors could be accurately modelled (see Section 4.2).

As previously discussed, 3D-MF measures the significance of each detection above the background galaxy distribution (equation 5). False detections as a function of the detection significance are presented in Fig. 6 with the significance levels at which known

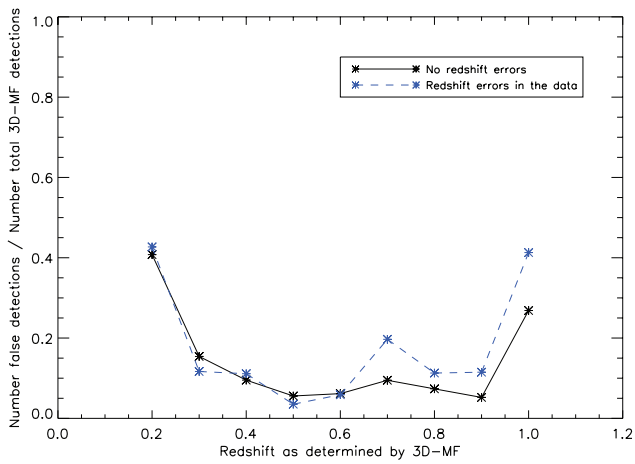

Figure 9. 3D-MF false detections, as a fraction of total 3D-MF detections, shown as a function of redshift.

clusters were accurately detected: many more highly significant clusters were found, which were real detections, rather than false detections. Introducing redshift errors does lower the significance of detections, but as Fig. 4 shows, completeness remains high.

4.3.1 Does 3D-MF have a proxy for mass?

The significance of cluster detections ($\geq 3.5\sigma$) as a function of cluster mass is presented in Fig. 10. Focusing on clusters with masses $>10^{14}\,\mathrm{M}_{\odot}$ we see higher mass clusters being detected with higher significance as one might expect. For lower mass clusters we find that the detection significance is not correlated with mass and

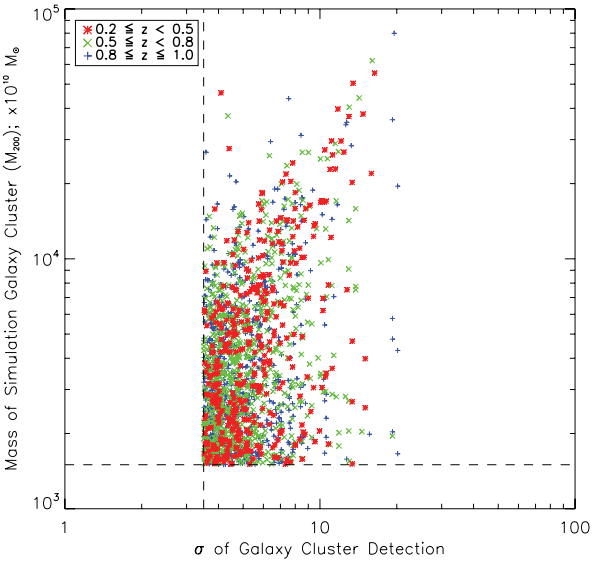

Figure 10. Significance of galaxy cluster detections in all six simulation light cones. Only detections $\geq 3.5\sigma$ are considered actual real detections. Known clusters are not considered below $1.5 \times 10^{13}\,\mathrm{M}_{\odot}$ (a sensible mass cut, well below the divergence from expectation as shown in Fig. 7). The redshift bins in the plot do not reflect redshift slicing, but rather are found clusters binned by their redshifts.

we believe this is due to a number of reasons. We looked at the few cases of $>10\sigma$ detections matched with known clusters of masses $<2 \times 10^{13}\,\mathrm{M}_{\odot}$ and found that while the closest known cluster to the detection was a low-mass cluster, in the majority of cases galaxy interlopers from a nearby more massive, $>10^{14}\,\mathrm{M}_{\odot}$, cluster were more likely responsible for the significance of the detection. In addition, as discussed in Section 4.1, 3D-MF is not fine-tuned for interpreting the detection of low-mass clusters as the radial filter is currently much larger than the extent of these low-mass clusters. 3D-MF assigns all galaxies within a 1 Mpc radius around a peak detection to a cluster; for low-mass clusters, galaxies within this radius and within a redshift slice, but outside the low-mass cluster halo, will add noise to a measurement of significance. This is demonstrated in Fig. 11 which presents the number of galaxies found by 3D-MF per known cluster as a function of cluster mass.

The left-hand panel of Fig. 11 (Fig. 11a) shows the results for exact redshifts and the right-hand panel (Fig. 11b) shows the results when photometric redshift errors are included. For clusters with $M_{200} > 5 \times 10^{13}\,\mathrm{M}_{\odot}$, 3D-MF correctly measures the number of galaxies belonging to a cluster. However, for lower mass clusters we see the measured galaxy cluster member counts unable to go below ∼30 galaxies per cluster (blue crosses), which is the average number of galaxies that fit 3D-MF's radial profile within a 1 Mpc window. Reducing the cut-off radius, r_{co}, in equation (3) would improve the association of correct galaxies to low-mass clusters but this would significantly reduce the high rate of success shown for higher mass cluster detections.

The introduction of redshift errors moves the points in Fig. 11(a) leftward, meaning that as cluster members become scattered in redshift space, 3D-MF associates less of them to their parent cluster, as would be expected. Interestingly, for the lowest mass clusters, which are the most difficult to detect, this means galaxy number counts are less likely to be overestimated while remaining detectable with 3D-MF.

Examining the number of galaxies per cluster, as determined by 3D-MF, as a potential proxy for mass is shown in Fig. 12. Again focusing on the clusters with masses $>10^{14}\,\mathrm{M}_{\odot}$ we see a weak redshift-dependent trend: the most massive clusters have the most numerous members. For the low-mass clusters however, owing to the incorrect assignment of too many galaxies to those clusters, we see no correlation.

We will investigate not having a clear proxy for mass via 3D-MF in future work. P96, D07 and O1–O4 assume that all light after background subtraction follows the Schechter function and that this can be expressed as a luminosity equivalent to the number of L^* galaxies (which they refer to as Λ_{cl}), with good results, but accurate galaxy number counts per cluster is less of an issue with the lower cluster densities and higher mass clusters their methods typically find. We plan to further investigate our currently overestimated cluster galaxy members in the low-mass ranges, by looking at galaxy number counts per cluster, and

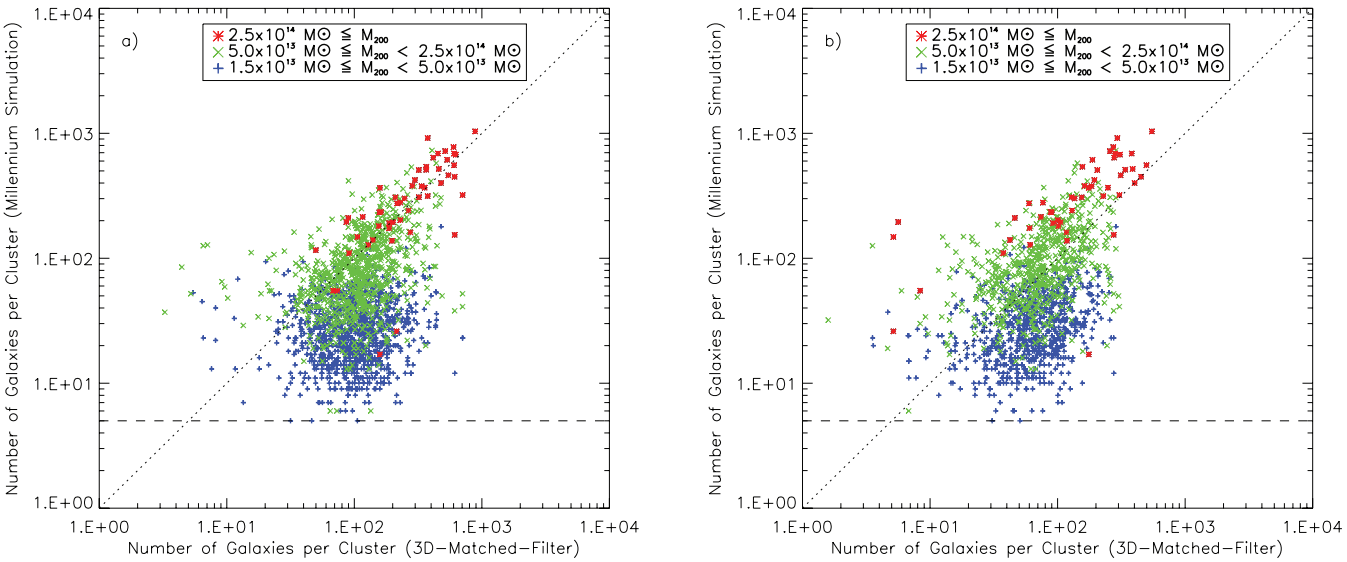

Figure 11. The number of galaxies found per known cluster as a function of cluster mass (the dashed horizontal line denotes the definition in Table 1 which requires clusters have ≥5 galaxy members). Fig. 11(a) presents exact redshifts while Fig. 11(b) includes redshift errors in the data.

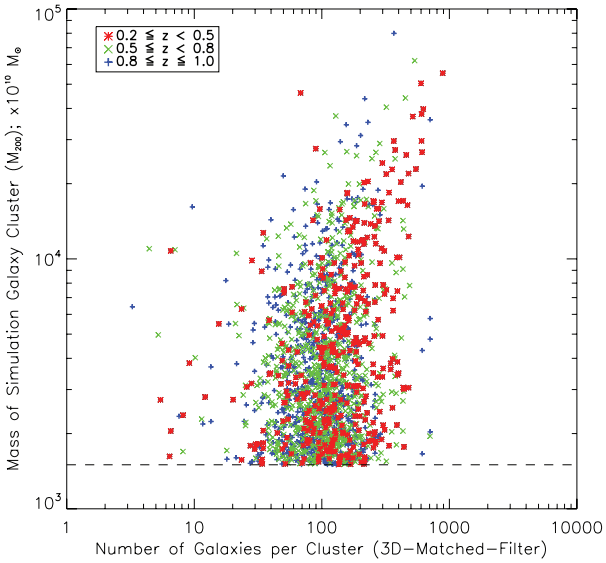

Figure 12. The range of cluster masses and the number of galaxies associated to those cluster masses by 3D-MF.

potentially optimizing a variable cut-off radius in the 3D-MF search window.

Methods complimentary to 3D-MF would work as mass estimators as well; weak gravitational lensing techniques for example can be utilized on real data to measure cluster masses (as in Hoekstra 2007). Alternatively, spectroscopic redshifts could be utilized to precisely identify galaxy cluster members and potentially lead to determining a better contamination fraction in the current relationship between 3D-MF galaxy cluster member number counts and cluster mass. X-ray information is known to be an excellent proxy for mass (Smith et al. 2005), and Sunyaev–Zel'dovich (SZ) information could be added to tell us more about the structure of clusters along the line-of-sight [as in Motl et al. (2005) and Sealfon, Verde & Jimenez (2006) for example].

4.3.2 Cluster catalogue purity

As previously discussed, it is not an easy task to define the centre position of a cluster in an automated way. By construction, 3D-MF does not have a cluster centre pre-defined. Other cluster-finding methods, such as those mentioned in Section 1, use the position of the brightest cluster galaxy member (Koester et al. 2007), or the centre of the distribution of red galaxies for example. We choose the centre of a galaxy cluster to be based on the luminosity and radial distribution of cluster galaxies. As shown in Fig. 11, for clusters with $M_{200} > 5 \times 10^{13}\,\mathrm{M}_{\odot}$, 3D-MF correctly associates galaxy members to their parent cluster, and we thereby expect the cluster centroid to be determined by this method with reasonable accuracy. For lower mass clusters however, owing to interloping galaxy members from nearby clusters (interlopers that fall within a radial profile centred near the lower mass cluster), the centroid that we measure is rather difficult to interpret.

Recall that a 3D-MF cluster detection within 0.75 Mpc of another simulation-matched 3D-MF cluster detection (a 3D-MF detection that was chosen to be the match as it was closer to the known cluster centre) is considered a secondary detection. This further complicates the issue of centroiding, as it is not clear whether the 3D-MF centroid for clusters with secondary detections should be calculated from the galaxies associated to both the primary and secondary detection or just the primary, especially considering the fact that clusters often have substructures (and perhaps in reality do not have a clear centre). In the analysis that follows we choose the primary detection as the centroid.

Fig. 13 shows the number of 3D-MF-derived clusters matched to known simulation clusters as a function of the distance between their centres. For all mass ranges we find that there are fewer 3D-MF cluster centre positions with increased radius from known cluster centres, showing that a 3D-MF centroid measurement is frequently quite accurate. The elevated plateau at a matching radius >1 arcmin, seen in the $1.5 \times 10^{13}\,\mathrm{M}_{\odot} \leq M_{200} < 5.0 \times 10^{13}\,\mathrm{M}_{\odot}$ mass range, indicates some level of contamination in our low-mass matched sample.

To investigate this, we randomized the cluster centroids found by 3D-MF in RA and Dec. and repeated our matching to known simulation cluster centre positions. The average random matches for all cluster mass ranges, from 50 randomized trials, is presented in Fig. 13 as the dotted line, which we call matching noise. Over 80 per cent of the contribution to this matching noise comes from low-mass clusters ($1.5 \times 10^{13} \leq M_{200} < 5.0 \times 10^{13}\,\mathrm{M}_{\odot}$) and is thus a major source of the contamination in the low-mass sample's plateau (red, triangle line) in Fig. 13. This is what we would expect: low-mass (and thus small) clusters contribute to the matching noise increasingly at larger radii from known KW07 centres since randomizing their 3D-MF-found positions should uniformly place their detections across the field, and with any size of matching window there will always be some of these low-mass clusters (of high number density) matched to a known cluster.

The matching noise we have shown is an overestimate of the real matches due to noise because randomizing the 3D-MF cluster positions in RA and Dec. eliminates properly identified multiple detections of single clusters, thereby overestimating single detections and maximally increasing the matching due to chance incorrect matches. For example, a cluster that is detected by 3D-MF in one original redshift slice and also in the overlapping redshift slice would be considered a multiply detected cluster. After randomization of the cluster positions, this would now appear as two separate cluster detections in the original and shifted redshift slices. To

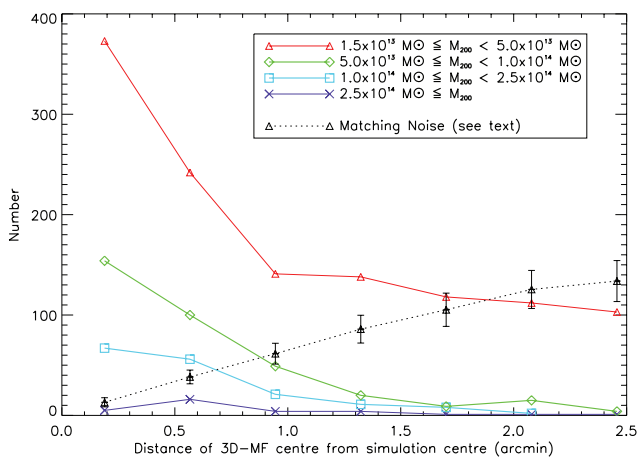

Figure 13. Distance of 3D-MF-derived cluster centres from predetermined, known simulation centres (as defined in Section 2.2). The dotted line shows results from 3D-MF-derived clusters randomized in RA and Dec. before matching to known simulation clusters according to our matching criteria: over 80 per cent of the contribution here is from the lowest (i.e. $1.5 \times 10^{13}\,\mathrm{M}_{\odot} \leq M_{200} < 5.0 \times 10^{13}\,\mathrm{M}_{\odot}$) mass range (see text for further discussion).

correct for this, we use our multiple detection rates of 36.2 per cent and conservatively report a maximal upper limit on our false detections of 24.3 per cent {i.e. [(Total number of false detections + $(1 - 0.362) \times$ Total number counts due to an upper limit of matching noise)/Total 3D-MF detections]}. As discussed at the beginning of this section, we believe the false detection rate (especially for clusters of masses $\geq 5.0 \times 10^{13}\,\mathrm{M}_{\odot}$) is closer to 15.6 per cent, but for the higher number densities of lower mass clusters this could be up to 24.3 per cent.

Due to the increased sensitivity of 3D-MF over 2D matched-filter methods, we believe 3D-MF is correctly recognizing the significance of low-mass clusters above background noise, but due to excess galaxy number counts per cluster, and oversized fiducial cluster filters, has trouble correctly choosing a cluster centroid for these low mass clusters. We can safely confirm from this centroid analysis that we have chosen a physically motivated and sensible matching radius within which to properly match 3D-MF's detections with known simulation clusters. Interestingly, the 2D matched-filter of WK02 successfully implements an iterative centroiding technique: a logical next step for improving the current implementation of 3D-MF and its ability to determine cluster centres.

5 CANADA–FRANCE–HAWAII TELESCOPE LEGACY SURVEY DEEP FIELD CLUSTER CATALOGUES GENERATED BY 3D-MF

5.1 CFHTLS Deep data and catalogue generation

For this paper we use the four $1^{\circ} \times 1^{\circ}$ CFHTLS Deep fields (D1, D2, D3, D4) and COSMOS data which were presented within the CFHTLS-Archive-Research Survey; we refer the reader to Erben et al. (2009) and Hildebrandt et al. (2009) for details on the data processing and multicolour catalogue creation.

Our five-band multicolour and photometric redshift catalogues are based on CFHTLS i'-band object detections. The seeing for this band varies between 0.71 and 0.82 arcsec within the four fields and we reach a 1σ limiting magnitude of $i'_{\mathrm{AB,lim}} \approx 29.5$ within a diameter of twice the seeing disc. Photometric redshifts were estimated with the publicly available code BPZ (see Benítez 2000 and the prescription in Hildebrandt et al. 2009). In the fields D1, D2 and D3 we quantified the accuracy of our photometric redshifts with spectroscopic redshifts from the VVDS (Le Fèvre et al. 2005), zCOSMOS (Lilly et al. 2007) and DEEP2 (Davis et al. 2007), respectively. Within the magnitude range of $17 < i'$-band < 24 we estimate unbiased photometric redshifts, and after rejecting problematic sources (we cut with the BPZ ODDS parameter for sources with ODDS > 0.9; see also Hildebrandt et al. 2009), we find a photometric–spectroscopic redshift scatter of $\sigma \approx 0.033$ of the quantity $\Delta z = (z_{\mathrm{phot}} - z_{\mathrm{spec}})/(1 + z_{\mathrm{spec}})$. Our outlier rate with $|\Delta z| > 0.15$ is 1.6 per cent. Our final Deep catalogues contain $\sim 1.23 \times 10^5$ galaxies deg^{-2} with i'-band magnitudes <25.5 (compared to the simulations with $z < 1.2$ and the same magnitude limit, which had $\sim 1.59 \times 10^5$ galaxies deg^{-2}).

5.2 CFHTLS Deep galaxy cluster catalogues courtesy of 3D-MF

3D-MF was set up according to the parameters in Table 2 but an appropriate M^* and α were derived from the CFHTLS Deep fields and found to be $M^*_{i'\text{-band}} = -22.46 \pm 0.15$ and $\alpha = -1.005 \pm 0.021$. 3D-MF was then run on the CFHTLS Deep fields. Table 4 contains eight random entries (two per field) from our CFHTLS Deep galaxy cluster catalogue; the entire catalogue is available upon request to the authors. From our Millennium Simulation tests we would expect to see ∼16–24 per cent false positives in this catalogue, distributed mostly in the lower mass ranges according to the selection functions in Section 4. Using our multiple detection criteria, we found that 37.6 per cent of Deep detections were duplicate detections of clusters (comparable to the ∼36 per cent multiple detection rate found from our Millennium Simulation tests). Grouping identification numbers in Table 4 are numbered flags: cluster detections within 0.75 Mpc of each other (recall Section 4.1.2 for details) in the same redshift slice are flagged with identical *same redshift* numbers, and cluster detections within a projected 0.75 Mpc of each other in overlapping redshift slices are separately flagged with identical *overlapping redshift* grouping identification numbers (which can be propagated through more than one overlapping redshift slice if still within 0.75 Mpc in projected radii of each other in overlapping slices). Note that grouping identification numbers in both the *same redshift* and *overlapping redshift* numbering restart at zero for each Deep field.

We use the significance of our detections to select the best galaxy cluster candidate from among multiple detections, and excise the remaining multiple detections from the following discussion and analysis. The redshift distribution of 3D-MF-found Deep galaxy clusters is shown in Fig. 14. A comparison to other published CFHTLS Deep cluster lists via older matched-filter methodology ensues.

Table 4. 3D-MF-derived CFHTLS Deep galaxy cluster catalogue. The complete catalogue is available upon request.

CFHTLS Deep galaxy clusters						
Deep field	RA	Dec.	Redshift	Significance (σ)	Grouping identification number: Same redshift	Grouping identification number: Overlapping redshift
D1	02:25:07.536	−04:01:47.892	0.2	7.16	–	–
D1	02:24:26.040	−04:52:29.676	0.9	5.21	2	58
D2	09:58:54.720	01:54:57.744	0.4	5.35	–	–
D2	09:58:53.759	02:14:16.188	0.5	5.15	–	31
D3	14:17:22.802	52:54:52.920	0.2	7.27	–	–
D3	14:18:19.440	53:05:58.200	0.5	5.61	–	–
D4	22:14:39.120	−18:09:41.760	0.6	5.23	–	37
D4	22:16:08.400	−18:11:48.840	0.4	6.73	–	–
...	...	...	...	...	...	...

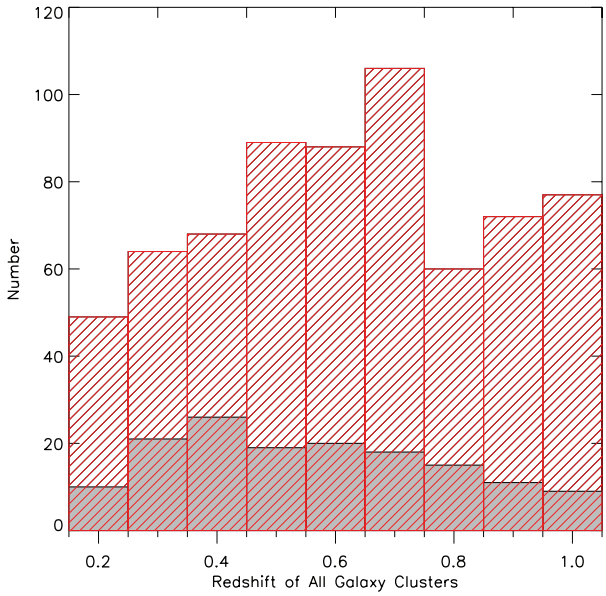

Figure 14. Redshift distribution of all 3D-MF-found (dashed up-sloping histogram) and O3-found (shaded histogram) galaxy clusters in CFHTLS Deep fields 1 through 4.

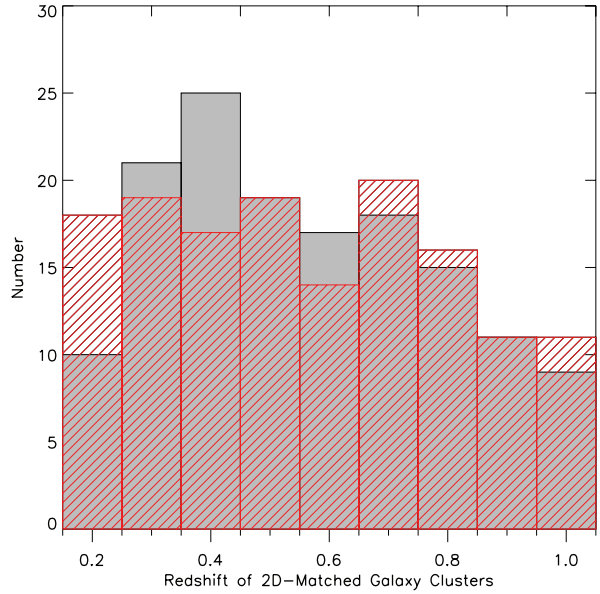

Figure 15. The redshift distribution of only those galaxy clusters found using *both* 3D-MF and O3 methods, matched to each other in 2D space (see text for details). 3D-MF results are represented by the dashed up-sloping histogram, and O3 results are the shaded histogram.

5.3 Comparison to other published CFHTLS Deep clusters found using matched-filter methods

O3 published a matched-filter-derived cluster list of the CFHTLS Deep fields. A subsequent paper in 2009 (O4) examined running their matched-filter method (which does not utilize photometric redshifts) separately on different wavelength bands and then merging lists with the 2007 paper, a technique similar to P96's efforts to reduce false detection rates. Since the resulting lists are not too different, we choose to compare the i'-band 2007 O3 cluster list to a cluster list found by 3D-MF, using the same i'-band data, in this paper. We focus on the $0.2 \leq z \leq 1.0$ range for clusters, keeping in line with our acceptable photometric redshift region for galaxies and our slicing methodology. The redshift distributions of the O3 cluster list, after their 3.5σ cut, for all four CFHTLS Deep fields is shown overlayed on the 3D-MF results in Fig. 14. Drastically fewer galaxy clusters were found by O3 at all redshifts; 3D-MF finds ∼170 galaxy clusters deg^{-2} (well within reason compared to the Millennium Simulation analysis cluster number densities in Section 4.3) compared to ∼40 galaxy clusters deg^{-2} in O3.

In order to match the clusters found using both methods, a 2D tolerance radius of 0.°044 (akin to the matching strategy in Section 4.1.1) was placed around each O3 cluster centre position and the closest 3D-MF detection in RA and Dec. was considered a match. Fig. 15 shows the redshift distribution, and Fig. 16 presents the redshift differences, of the 2D matches between the two cluster lists.

As mentioned in Section 3.3.5, older matched-filter methods tend to wrongly estimate the redshift of clusters. O3 searches the entire input catalogue (i.e. no redshift slicing) for a match to a fiducial cluster sized to match what would be expected at a particular redshift, and then repeats the process with a slightly resized fiducial cluster size (to match what would be expected for a cluster that existed at a slightly different redshift); O3 requires each cluster to be detected in two neighbouring fiducial cluster resizing searches in order to be considered further. This affects higher redshifts disproportionately, as they are less likely to be found in equally stepped fiducial cluster-resizing bins than lower redshift clusters; higher redshift space of equivalent bin widths covers much more volume. Fig. 15 shows the likely result of this: the O3 cluster sample is skewed towards lower redshifts.

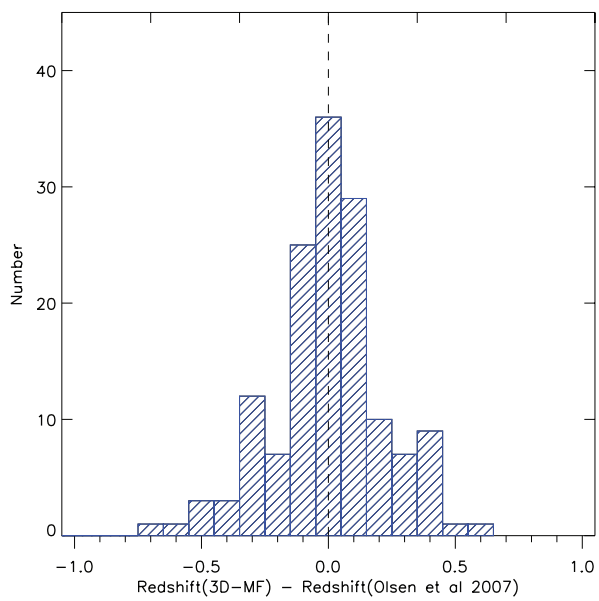

Figure 16. Differences in assigned galaxy cluster redshifts for clusters found using both 3D-MF and O3 methods (using 2D matching as described in Section 5.3).

A second separate matching was performed between the 3D-MF Deep cluster list and O3; this time a 3D tolerance range was considered around each O3 cluster. There were less matches, as seen in Fig. 17, again skewed towards the lower redshift end of the distribution.

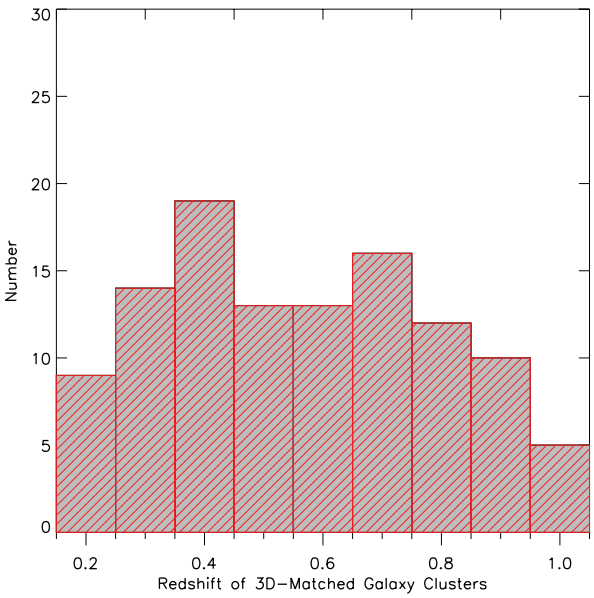

Figure 17. Redshift distribution of galaxy clusters found using both 3D-MF and O3 methods, matched to each other in 3D space (see text for details). The vertical scale of Fig. 15 is duplicated here for comparison purposes.

Comparing both the 3D-MF and O3 CFHTLS Deep cluster lists, there were 145 clusters found by both 3D-MF and O3 out of 149 total O3 clusters (97 per cent) with the 2D matching criteria described. 528 additional galaxy clusters were found with 3D-MF. As mentioned, the matching process between 3D-MF and O3 clusters was repeated considering a further third dimension in redshift space (see Table 5); in this case there were 111 clusters found by both 3D-MF and O3, suggesting 34 of the O3 clusters (23 per cent) were not assigned the correct redshift by their algorithm.

The completeness has been shown to be vastly different between the two methods: 3D-MF is 100 per cent complete down to $\sim 3 \times 10^{14}\,\mathrm{M}_{\odot}$, $\sim$90 per cent complete down to $1 \times 10^{14}\,\mathrm{M}_{\odot}$ and the false detections are likely concentrated even further down the cluster mass function at the low-mass end. It was discussed in general in Section 3.2 that older matched-filter methods have false detection rates of $\sim$30–65 per cent, suffer from line-of-sight projection issues, wrongly estimate redshifts, and have not been tested on Millennium Simulation like realistic, extensive simulations.

Table 5. A comparison of CFHTLS Deep galaxy clusters found with 3D-MF and published CFHTLS Deep galaxy clusters from O3. Cluster redshift ranges of $0.2 \leq z \leq 1.0$ are considered, keeping in line with this paper, and galaxy clusters present in both catalogues with similar redshifts (as explained in the text) are reported again in the 'O3 and 3D-MF' column.

3D cluster comparison			
CFHTLS Deep field	3D-MF	O3	O3 and 3D-MF
D1	161	43	32
D2	162	38	31
D3	190	38	25
D4	160	30	23
Total D1–D4	673	149	111

6 CONCLUSIONS

3D-MF is an optimized, automated galaxy cluster finder, ideal for large optical astronomical data sets with multiwavelength band coverage. It utilizes a fiducial galaxy cluster's radial and luminosity profiles and searches the data for matches to this over a wide range of redshifts. Future surveys, such as LSST and JDEM, can exploit 3D-MF's automated methodology and statistical approach to produce complete and reliable galaxy cluster catalogues. 3D-MF is an improvement over other matched-filter techniques due to several improvements, including

(i) the cutting of input data into overlapping redshift slices to examine it piecewise, significantly reducing line-of-sight projections;
(ii) an implementation of mask-handling capabilities, improving edge effects and cluster detections near bright stars or saturated pixels;
(iii) an accurate modelling of data-dependent background galaxy counts;
(iv) the development of a new peak detection pipeline;

as well as a list of other parameters that can be fine-tuned according to the data set being examined.

The Millennium Simulation mock catalogue light cones were used to extensively test and improve 3D-MF, and selection functions for the algorithm were presented. Redshift errors mimicking real data were modelled and added to the simulations and their effect on the selection functions was derived. With redshift errors, and focusing on the cluster range $0.2 \leq z \leq 1.0$, 3D-MF was found to recover 100 per cent of known galaxy clusters with an $M_{200} \geq 3.0 \times 10^{14}\,\mathrm{M}_{\odot}$; 97 per cent of clusters with an $M_{200} \geq 2.5 \times 10^{14}\,\mathrm{M}_{\odot}$; 88 per cent of clusters with an $M_{200} \geq 1.0 \times 10^{14}\,\mathrm{M}_{\odot}$ and 72 per cent of clusters with an $M_{200} \geq 5.0 \times 10^{13}\,\mathrm{M}_{\odot}$. 36 per cent of detections were multiple detections of clusters. False detections were determined to be occurring at a rate of 15.6 per cent of the total cluster detections, and a subsequent analysis showed this to be concentrated in the low significance, low galaxy number ($\lesssim$50) per cluster, and likely lower mass ($M_{200} < 1 \times 10^{14}\,\mathrm{M}_{\odot}$) range (potentially increased by noise up to a conservatively reported rate of 24 per cent). After selection functions were quantified (and the effects of adding redshift errors to the catalogues were analysed), 3D-MF was run on the four CFHTLS Deep fields. 3D-MF finds $\sim$170 galaxy clusters deg^{-2} in the CFHTLS Deep data set: over 400 per cent more, with a much lower false detection rate, and higher accuracy of redshift determination for true clusters, than found by other authors using 2D matched-filter methods on the same i'-band data.

For future work, there are subtle adjustments to 3D-MF that we are examining. A non-passively evolving k-correction could be applied, taking into consideration the effect of variations in galaxy types. More interestingly perhaps, as shown in Popesso et al. (2005), galaxy clusters are often better fitted with two Schechter functions as opposed to one: implementing this result may further improve 3D-MF. The radial profile used in 3D-MF's radial filter could also be tuned to try to find more low-mass, smaller clusters (if possible). Different filter bands could be used in 3D-MF when analysing the redshift slices, and results compared with the i'-band used herein, or spectroscopic redshifts could be utilized to confirm cluster detections and precisely identify galaxy cluster members. It would be sensible to try to extend the high-redshift range of the detection capabilities of 3D-MF as well, although this method is ultimately limited by the build-up of the luminosity function at redshifts $\gtrsim$1.2.

The Millennium Simulation has proved invaluable in refining and understanding 3D-MF as an automated optical cluster finder. A few points were raised that we continue to study in fastidious detail: finding a 3D-MF-derived proxy for cluster mass, refining the number of galaxies associated to a cluster, and improving cluster centroid determinations are examples. In the future we will introduce a weak gravitational lensing analysis to elucidate cluster masses and we are also investigating multiwavelength coverage in X-ray and SZ data to refine cluster information.

ACKNOWLEDGMENTS

We thank the reviewer for useful comments and feedback. MM also thanks Gabriella De Lucia for helpful discussions on the Millennium Simulation. This work involves observations 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 the Canadian Astronomy Data Centre as part of the CFHTLS, a collaborative project of NRC and CNRS. HH was supported by the European DUEL RTN, project MRTN-CT-2006-036133. TE is supported by the German Ministry for Science and Education (BMBF) through DESY under the project 05AV5PDA/3 and the German Science Foundation (DFG) through the project TR33 'The Dark Universe'. The Millennium Simulation data bases used in this paper and the web application providing online access to them were constructed as part of the activities of the German Astrophysical Virtual Observatory.

REFERENCES

Adami C. et al., 2010, A&A, 509, 81
Battye R. A., Weller J., 2003, Phys. Rev. D, 68, 083506
Benítez N., 2000, ApJ, 536, 571
Benoist C. et al., 2002, A&A, 394, 1
Cohn J. D., Evrard A. E., White M., Croton D., Ellingson E., 2007, MNRAS, 382, 1738
Davis M. et al., 2007, ApJ, 660, L1
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
Dietrich J. P., Erben T., Lamer G., Schneider P., Schwope A., Hartlap J., Maturi M., 2007, A&A, 470, 821 (D07)
Dressler A., 1980, ApJ, 236, 351
Erben T. et al., 2009, A&A, 493, 1197
Gilbank D. G., Bower R. G., Castander F. J., Ziegler B. L., 2004, MNRAS, 348, 551
Gladders M. D., Yee H. K. C., 2000, AJ, 120, 2148
Grove L. F., da Costa L., Benoist C., 2008, A&A, 490, 945
Grove L. F., Benoist C., Martel F., 2009, A&A, 494, 845 (O4)
Hansen L., Olsen L. F., Jørgensen H. E., 2002, A&A, 388, 1
Hilbert S., White S. D. M., 2010, MNRAS, 404, 486
Hildebrandt H., Pielorz J., Erben T., van Waerbeke L., Simon P., Capak P., 2009, A&A, 498, 725
Hoekstra H., 2007, MNRAS, 379, 317
Kepner J., Fan X., Bahcall N., Gunn J., Lupton R., Xu G., 1999, ApJ, 517, 78
Kitzbichler M. G., White S. D. M., 2007, MNRAS, 376, 2 (KW07)
Kochanek C. S., White M., Huchra J., Macri L., Jarrett T. H., Schneider S. E., Mader J., 2003, ApJ, 585, 161
Kodama T., Tanaka I., Kajisawa M., Kurk J., Venemans B., De Breuck C., Vernet J., Lidman C., 2007, MNRAS, 377, 1717
Koester B. P. et al., 2007, ApJ, 660, 221
Le Fèvre O. et al., 2005, A&A, 439, 845
Li I. H., Yee H. K. C., 2008, AJ, 135, 809
Lilly S. J. et al., 2007, ApJS, 172, 70
Lopes P. A. A., de Carvalho R. R., Gal R. R., Djorgovski S. G., Odewahn S. C., Mahabal A. A., Brunner R. J., 2004, AJ, 128, 1017
Lu T., Gilbank D. G., Balogh M. L., Bognat A., 2009, MNRAS, 399, 1858
Menanteau F. et al., 2009, ApJ, 698, 1221
Motl P. M., Hallman E. J., Burns J. O., Norman M. L., 2005, ApJ, 623, L63
Olsen L. F. et al., 1999a, A&A, 345, 681 (O1)
Olsen L. F. et al., 1999b, A&A, 345, 363 (O2)
Olsen L. F., Hansen L., Jørgensen H. E., Benoist C., da Costa L., Scodeggio M., 2003, A&A, 409, 439
Olsen L. F., Benoist C., da Costa L., Hansen L., Jørgensen H. E., 2005a, A&A, 435, 781
Olsen L. F., Zucca E., Bardelli S., Benoist C., da Costa L., Jørgensen H. E., Biviano A., Ramella M., 2005b, A&A, 442, 841
Olsen L. F. et al., 2007, A&A, 461, 81 (O3)
Popesso P., Böhringer H., Romaniello M., Voges W., 2005, A&A, 433, 415
Postman M., Lubin L. M., Gunn J. E., Oke J. B., Hoessel J. G., Schneider D. P., Christensen J. A., 1996, AJ, 111, 615 (P96)
Ramella M. et al., 2000, A&A, 360, 861
Sealfon C., Verde L., Jimenez R., 2006, ApJ, 649, 118
Smith G. P., Kneib J., Smail I., Mazzotta P., Ebeling H., Czoske O., 2005, MNRAS, 359, 417
Soneira R. M., Peebles P. J. E., 1978, AJ, 83, 845
Springel V. et al., 2005, Nat, 435, 629 (S05)
Stark D. P., Ellis R. S., Richard J., Kneib J., Smith G. P., Santos M. R., 2007, ApJ, 663, 10
Thanjavur K., Willis J., Crampton D., 2009, ApJ, 706, 571
van Breukelen C., Clewley L., 2009, MNRAS, 395, 1845
van Waerbeke L., 2000, MNRAS, 313, 524
White M., Kochanek C. S., 2002, ApJ, 574, 24 (WK02)

This paper has been typeset from a TEX/LATEX file prepared by the author.