Searching for massive clusters in weak lensing surveys

Takashi Hamana, Masahiro Takada, Naoki Yoshida

Mon. Not. R. Astron. Soc. **350,** 893–913 (2004) doi:10.1111/j.1365-2966.2004.07691.x

Searching for massive clusters in weak lensing surveys

Takashi Hamana,[1,2]★ Masahiro Takada[3] and Naoki Yoshida[2,4]

[1]*Institut d'Astrophysique de Paris, 98bis Boulevard Arago, F 75014 Paris, France*
[2]*National Astronomical Observatory of Japan, Mitaka, Tokyo 181-8588, Japan*
[3]*Department of Physics and Astronomy, University of Pennsylvania, 209 S. 33rd Street, Philadelphia, PA 19104, USA*
[4]*Harvard-Smithsonian Centre for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*

Accepted 2004 January 30. Received 2004 January 30; in original form 2003 October 24

ABSTRACT

We explore the ability of weak lensing surveys to locate massive clusters. We use both analytic models of dark matter haloes and mock weak lensing surveys generated from a large cosmological N-body simulation. The analytic models describe the average properties of weak lensing haloes and predict the number counts, enabling us to compute an effective survey selection function. We argue that the detectability of massive haloes depends not only on the halo mass but also strongly on the redshift where the halo is located. We test the model prediction for the peak number counts in weak lensing mass maps against mock numerical data, and find that the noise resulting from intrinsic galaxy ellipticities causes a systematic effect which increases the peak counts. We develop a correction scheme for the systematic effect in an empirical manner, and show that, after correction, the model prediction agrees well with the mock data. The mock data is also used to examine the completeness and efficiency of the weak lensing halo search by fully taking into account the noise and the projection effect by large-scale structures. We show that the detection threshold of $\mathrm{S/N} = 4 \sim 5$ gives an optimal balance between completeness and efficiency. Our results suggest that, for a weak lensing survey with a galaxy number density of $n_{\mathrm{g}} = 30\,\mathrm{arcmin}^{-2}$ with a mean redshift of $z = 1$, the mean number of haloes which are expected to cause lensing signals above $\mathrm{S/N} = 4$ is $N_{\mathrm{halo}}(\mathrm{S/N} > 4) = 37$ per $10\,\mathrm{deg}^2$, whereas 23 of the haloes are actually detected with $\mathrm{S/N} > 4$, giving the effective completeness as good as 63 per cent. Alternatively, the mean number of peaks in the same area is $N_{\mathrm{peak}} = 62$ for a detection threshold of $\mathrm{S/N} = 4$. Among the 62 peaks, 23 are caused by haloes with the expected peak height $\mathrm{S/N} > 4$, 13 result from haloes with $3 < \mathrm{S/N} < 4$ and the remaining 26 peaks are either the false peaks caused by the noise or haloes with a lower expected peak height. Therefore the contamination rate is 44 per cent (this could be an overestimation). Weak lensing surveys thus provide a reasonably efficient way to search for massive clusters.

Key words: gravitational lensing – galaxies: haloes – cosmology: theory – dark matter – large-scale structure of Universe.

1 INTRODUCTION

Recently, it has become feasible to locate massive clusters directly as density enhancements using weak gravitational lensing. Miyazaki et al. (2002) indeed discovered many clusters in a weak lensing halo survey over a 2.1-deg^2 field, proving the ability of weak lensing to identify massive clusters (see also Wittman et al. 2001). Unlike conventional optical or X-ray selected cluster catalogues, weak lensing cluster catalogues are free from a bias towards luminous objects because this type of cluster finding is not based on flux enhancement but on *projected mass* enhancement. This is one great advantage of gravitational lensing. However, like every other cluster survey technique (i.e. optical, X-ray and Sunyaev–Zel'dovich survey) lensing does have disadvantages; for example, the projection effect of unrelated structures in the same line of sight (Reblinsky & Bartelmann 1999; Metzler, White & Loken 2001). As galaxy clusters are some of the most important cosmological probes, it is of fundamental importance to construct large unbiased samples. Given the advantages and disadvantages involved in every survey technique, it is desirable to combine, in a complementary manner, several techniques to construct an unbiased catalogue. Such catalogues provide a valuable data set to investigate, for example, the physical state of the intracluster medium. To do this, it is important to correctly understand the selection function and completeness of each survey technique.

★E-mail: hamana@iap.fr

It has often been argued that weak lensing provides a truly mass-selected sample of clusters. This is not strictly true; weak lensing measures the two-dimensional tidal field which is a line of sight projection of the three-dimensional tidal field weighted by the distance ratio, $D_{\rm l}D_{\rm ls}/D_{\rm s}$ (see Section 2 for definitions and details; see (Mellier 1999; Bartelmann & Schneider 2001) for reviews of weak lensing). In simpler terms, the amplitude of the lensing signal from a cluster is not solely determined by its mass, but also depends on the redshift and the shape of the gravitational potential caused by the cluster. In addition, both foreground and background large-scale structures contribute to the lensing signal, and this projection effect usually adds a noise to the lensing signal from a halo. Further, since the weak lensing signal is measured from tiny coherent distortions in galaxy shapes, intrinsic ellipticities of galaxies introduce an irremovable noise. Thus, the detectability of the weak lensing halo depends on all these factors.

The primary purpose of the present paper is to explore the ability of weak lensing surveys to locate massive clusters. We especially address the following three points:

(i) to examine the selection function of weak lensing cluster surveys;

(ii) to develop a theoretical model that describes the weak lensing halo counts; and

(iii) to examine how the detailed structure of clusters, projection effect and noise affect weak lensing cluster surveys under a typical observational condition from a ground-based 4–10 m class telescope.

The last point is especially important to understand a bias that weak lensing halo catalogues may have. To address these things, we use both simple analytic descriptions of dark matter haloes and mock numerical weak lensing survey data. The former offers a useful way to compute expected lensing properties of massive haloes. We use them to compute the selection function. We also develop a theoretical model for the number counts of the weak lensing haloes based on the analytic descriptions. Alternatively, the latter allow us to examine *all* the factors listed above in a direct manner. Mock numerical catalogues are generated using weak lensing ray-tracing simulations. We extensively use the mock data and the halo catalogues directly produced from the simulation outputs to examine the *completeness* and *efficiency* of the weak lensing halo survey. The mock data are also used to test the model prediction of the weak lensing halo counts.

White, Van Waerbeke & Mackey (2002) and Padmanabhan, Seljak & Pen (2003) addressed the ability of weak lensing surveys focusing on the projection effect. In those studies, the completeness of weak lensing surveys is defined as the fraction of detected haloes with mass above a certain value relative to haloes that lie in a given volume. This definition compromises two completely different elements: (i) the selection effect; and (ii) the effects from individuality of the halo mass distribution, the projection and noise. For our purpose, these elements needs to be examined separately. We first compute the selection function using the analytical descriptions of dark matter haloes which enables us to define 'potentially detectable haloes'. Here, 'potentially' refers to an ideal case in the absence of noise. Then, we shall adopt the different definition of the *completeness* that is the fraction of actually detected weak lensing haloes relative to the potentially detectable haloes.

Our results have important implications to real observations. When making an observational strategy, one would like to compute the expected number of detectable clusters using, for example, a simplified model, and also needs to estimate the expected completeness and false detection rate. Our theoretical model of the weak lensing halo counts may be useful for the former case, and our results from the mock data are directly applicable to derive the latter estimates. In addition, from an observational point of view, we adopt a directly observable quantity, a peak height of weak lensing mass map divided by the noise rms, that is the signal-to-noise ratio, S/N, of the detection, as the fundamental estimate of the lensing signal.

The rest of this paper is organized as follows. In Section 2, we summarize lensing properties of massive haloes assuming the universal density profile proposed by (Navarro, Frenk & White 1996, 1997). Using the expected properties, we compute the selection function of the weak lensing cluster survey. Then, we combine the detectability of weak lensing haloes with the halo mass function based on the Press–Schechter approach (Press & Schechter 1974; we adopt the modified fitting function by Sheth & Tormen 1999) to compute the weak lensing halo counts and the selection functions with respect to the halo mass and redshift. This simple analytical approach should offer a theoretical basis to explore properties of weak lensing haloes. In Section 3, we describe our numerical experiment of taking weak lensing surveys and our weak lensing ray-tracing technique. Statistical properties of halo catalogues are obtained and compared with the theoretical predictions derived in Section 2. Using the mock weak lensing survey data, in Section 4 we examine the number counts of peaks in the weak lensing convergence map, and the theoretical prediction developed in Section 2 is tested. In Section 5, we investigate the relationships between haloes in the halo catalogue and peaks identified in the lensing convergence map. In particular, we examine the completeness and efficiency of the weak lensing halo search paying special attention to points (1)–(3) above. A summary and discussion are given in Section 6. In Appendix A, we present some lensing convergence maps in which a missing halo or a false peak exists for illustrative examples of irregular systems. In Appendix B, we develop a correction scheme for the theoretical prediction of weak lensing peak counts. In Appendix C, the relationships between the halo shape (and orientation) and the strength of the lensing signal are examined.

Throughout this paper, we work with the standard flat Λ-cold dark matter (CDM) cosmological model with the density parameter $\Omega_{\rm m}=0.3$, the cosmological constant $\Omega_\Lambda=0.7$, and the Hubble constant $H_0=100\,h$ km s^{-1} Mpc^{-1} with $h=0.7$. We adopt the fitting function of the CDM power spectrum by Bardeen et al. (1986) with the normalization of $\sigma_8=0.9$. We take the observational parameters of a typical weak lensing survey from a present-day ground-based 4–10 m telescope. Specifically, we consider an imaging observation with a limiting magnitude fainter than $R=25.5$ mag in a subarcsecond seeing condition; this provides the galaxy number density $n_{\rm g}\gtrsim 30$ arcmin^{-2} with the mean redshift $z\simeq 1$.

2 PREDICTIONS FROM STANDARD MODELS

2.1 Lensing properties of an NFW halo

The universal density profile of dark matter haloes proposed by (Navarro et al. 1996, 1997 NFW hereafter) is given by

$$\rho_{\rm NFW}(x)=\frac{\rho_{\rm s}}{x(1+x)^2},\quad x=\frac{r}{r_{\rm s}}, \tag{1}$$

where $r_{\rm s}$ is the scaleradius. It is convenient to introduce the concentration parameter $c_{\rm NFW}=r_{\rm vir}/r_{\rm s}$ where $r_{\rm vir}$ is the virial

radius. Bullock et al. (2001) found from N-body simulations that the concentration parameter is related to the halo mass by

$$c_{\rm NFW}(M,z)=\frac{c_*}{1+z}\left(\frac{M}{10^{14}h^{-1}{\rm M}_\odot}\right)^{-0.13}, \tag{2}$$

where we set $c_*=8$ for the ΛCDM model. It is important to notice that there is a relatively large scatter in this relationship (Jing 2000). The mass enclosed within a sphere of radius $r_{\rm vir}$ gives the virial mass by definition, i.e.

$$M_{\rm vir}=\frac{4\pi\rho_{\rm s}r_{\rm vir}^3}{c_{\rm NFW}^3}\left[\log(1+c_{\rm NFW})-\frac{c_{\rm NFW}}{1+c_{\rm NFW}}\right]. \tag{3}$$

The virial mass is also defined by the spherical top-hat collapse model as

$$M_{\rm vir}=\frac{4\pi}{3}\delta_{\rm vir}(z)\bar{\rho}_0 r_{\rm vir}^3, \tag{4}$$

where $\delta_{\rm vir}$ is the threshold overdensity for spherical collapse (see Nakamura & Suto 1997; Henry 2000 for useful fitting functions). From equations (3) and (4) one can express $\rho_{\rm s}$ in terms of $\delta_{\rm vir}(z)$ and $c_{\rm NFW}$. Introducing $\delta_{\rm s}=\rho_{\rm s}/\bar{\rho}_0-1$, one finds

$$\delta_{\rm s}=\frac{\delta_{\rm vir}}{3}\frac{c_{\rm NFW}^3}{\log(1+c_{\rm NFW})-c_{\rm NFW}/(1+c_{\rm NFW})}. \tag{5}$$

Although the NFW profile formally extends to infinity, an actual halo has a finite extent. We consider a truncated NFW halo, the density profile of which is truncated at the virial radius. The surface mass density of the truncated NFW halo is given by the line-of-sight integration of the density profile (Takada & Jain 2003), i.e.

$$\Sigma(y)=\int_{-\sqrt{c_{\rm NFW}^2-y^2}}^{\sqrt{c_{\rm NFW}^2-y^2}}{\rm d}z\,\rho_{\rm NFW}(\boldsymbol{y},z)=2\rho_{\rm s}r_{\rm s}f(y),\quad y=\frac{r}{r_{\rm s}}, \tag{6}$$

with

$$f(y)=\begin{cases}-\frac{\sqrt{c_{\rm NFW}^2-y^2}}{(1-y^2)(1+c_{\rm NFW})}+\frac{1}{(1-y^2)^{3/2}}\,\arccosh\,\frac{y^2+c_{\rm NFW}}{y(1+c_{\rm NFW})}, \\ \quad {\rm for} y<1 \\ \frac{\sqrt{c_{\rm NFW}^2-1}}{3(1+c_{\rm NFW})}\left(1+\frac{1}{1+c_{\rm NFW}}\right), \quad {\rm for}\ y=1, \\ -\frac{\sqrt{c_{\rm NFW}^2-y^2}}{(1-y^2)(1+c_{\rm NFW})}-\frac{1}{(y^2-1)^{3/2}}\,\arccos\,\frac{y^2+c_{\rm NFW}}{y(1+c_{\rm NFW})}, \\ {\rm for} 1<y\leqslant c, \\ 0, \quad {\rm for} y>c.\end{cases} \tag{7}$$

Note that the above expression differs from that by Bartelmann (1996) and Wright & Brainerd (2000) which is for a non-truncated NFW halo and gives a larger surface mass density than ours. The lensing convergence profile is given by

$$\kappa(y)=\frac{\Sigma(y)}{\Sigma_{\rm cr}}=\kappa_{\rm s}f(y), \tag{8}$$

with the critical surface mass density

$$\Sigma_{\rm cr}=\frac{c^2}{4\pi G}\frac{D_{\rm s}}{D_{\rm l}D_{\rm ls}} \tag{9}$$

and

$$\kappa_{\rm s}=3\Omega_{\rm m}\delta_{\rm s}r_{\rm s}\left(\frac{H_0}{c}\right)^2\frac{\chi_{\rm l}(\chi_{\rm s}-\chi_{\rm l})}{a(z_{\rm l})\chi_{\rm s}}, \tag{10}$$

where a is the scalefactor normalized at the present time, D denotes the angular diameter distance, χ denotes the comoving radial distance (which is equal to the comoving angular diameter distance for a flat universe), and the subscripts l and s stand for lens and source, respectively. It is worth noting that κ scales with $c_{\rm NFW}$ approximately as $\kappa\propto c_{\rm NFW}$ for $1\lesssim c_{\rm NFW}\lesssim 5$ and $\kappa\propto c_{\rm NFW}^{1.5}$ for $5\lesssim c_{\rm NFW}\lesssim 30$. This clearly proves that the convergence amplitude is sensitive to the halo concentration.

With the number density of source galaxies and the intrinsic ellipticity rms that we are interested in, a raw (unsmoothed) convergence map is very noisy. To get a lensing mass map with a reasonably small noise amplitude, we smooth the convergence map with a Gaussian window function, i.e.

$$W_{\rm G}(x)=\frac{1}{\pi\theta_{\rm G}^2}\exp\left(-\frac{x^2}{\theta_{\rm G}^2}\right). \tag{11}$$

The Gaussian smoothing is indeed adopted in actual weak lensing surveys (e.g. Miyazaki et al. 2002). As shown in the study of Van Waerbeke (2000), if the intrinsic ellipticities are uncorrelated between different source galaxies, the statistical properties of the resulting noise field can be described by Gaussian random field theory (Bardeen et al. 1986; Bond & Efstathiou 1987), on scales where the discreteness effect of the source galaxies can be ignored. The Gaussian field is specified by the variance of the noise field that is specified by the number of galaxies contained within a smoothing aperture (Kaiser & Squires 1993; Van Waerbeke 2000)

$$\sigma_{\rm noise}^2=\frac{\sigma_\epsilon^2}{2}\frac{1}{2\pi\theta_{\rm G}^2 n_{\rm g}}, \tag{12}$$

where σ_ϵ is the rms amplitude of the intrinsic ellipticity distribution and $n_{\rm g}$ is the number density of the source galaxies. Throughout this paper we adopt $\sigma_\epsilon=0.4$ and $n_{\rm g}=30$ arcmin^{-2}, which are typical values for a survey with a limiting magnitude that is fainter than $R=25.5$ mag in a subarcsecond seeing condition. With this choice, the noise rms is $\sigma_{\rm noise}=0.02\times(1\ {\rm arcmin}/\theta_{\rm G})$. It is worth noting that $\sigma_{\rm noise}$ is a direct observable in The sense that it can be estimated from a reconstructed weak lensing mass map obtained after randomizing the orientation of the images of the source galaxy.

We also need to take into account the smoothing effect on the lensing signal from a halo. The peak amplitude of the convergence profile of haloes is modified by smoothing into

$$\kappa_{\rm G}(\theta_{\rm G})=\int{\rm d}^2\phi\,W_{\rm G}(\phi;\theta_{\rm G})\,\kappa(\phi), \tag{13}$$

where the centre of the smoothing kernel is set to the halo centre. Hereafter we refer to this peak amplitude as 'the lensing signal resulting from a halo'. Fig. 1 plots the lensing signal as a function of the smoothing scale. The dotted line shows a scaling of the noise rms with the smoothing scale ($\sigma_{\rm noise}\propto\theta_{\rm G}^{-1}$). The comparison between the solid and dotted curves thus shows how S/N varies with the smoothing scale. Clearly, an optimal S/N can be attained for $\theta_{\rm G}=(1-2)\times\theta_{\rm s}$. Fig. 2 shows the contour plot of $\theta_{\rm s}$ in the halo mass and redshift plane, where the mean mass–concentration relationship given by equation (2) is employed. Note that throughout this paper, the source redshift is fixed at $z_{\rm s}=1$, which is a typical value for the mean redshift of source galaxies for an actual survey with limiting magnitude $R\simeq 25.5$. By combining Figs 1 and 2, it can be seen that $\theta_{\rm G}\sim 1$ arcmin would be the optimal choice for an efficient survey of massive haloes with $M_{\rm halo}\gtrsim 10^{14}\,h^{-1}\,{\rm M}_\odot$ at intermediate redshifts.

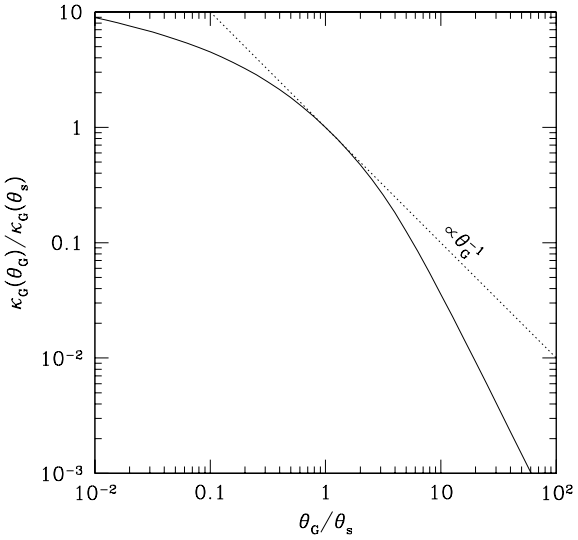

Figure 1. The solid line shows the lensing signal from an NFW halo smoothed with the Gaussian kernel (equation 13) against the smoothing length. For the concentration parameter, $c_{\rm NFW} = 5$ is adopted here. The dotted line shows scaling of the noise rms with the smoothing scale ($\sigma_{\rm noise} \propto \theta_{\rm G}^{-1}$, see equation 12).

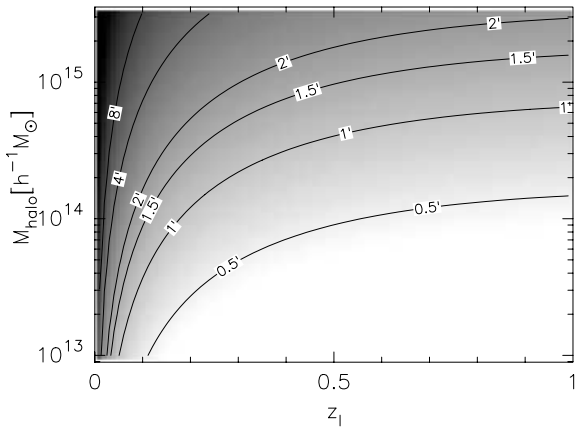

Figure 2. The grey scale with contour lines show the angular scaleradius $\theta_{\rm s}$ (in arcmin), in the halo mass ($M_{\rm halo}$) and redshift ($z_{\rm l}$) plane. The source redshift is fixed at $z_{\rm s} = 1$.

We define S/N for a weak lensing halo detection by[1]

$$\nu = \frac{\kappa_{\rm G}}{\sigma_{\rm noise}}. \tag{14}$$

Fig. 3 shows the contour plot of the S/N value computed assuming the universal NFW profile in the halo mass and redshift plane. It is evident from this figure that the lensing signal not only depends on the mass but also on the redshift. This redshift dependence enters mainly through the lensing efficiency function $D_{\rm l}D_{\rm ls}/D_{\rm s}$ (see

[1] With respect to notation throughout this paper, we denote the S/N of weak lensing signals by ν and distinguish its type with a subscript; $\nu_{\rm NFW}$ stands for the S/N computed assuming a universal NFW profile, while $\nu_{\rm lens}$ and $\nu_{\rm noisy}$ stands for the S/N measured concerning noise-free and noisy numerical weak lensing mock data, respectively (see Sections 3 and 4). We will also use $\nu_{\rm th}$ and $\nu_{\rm min}$ to denote a threshold detection value and the minimum value (to be set; see Sections 3 and 4), respectively.

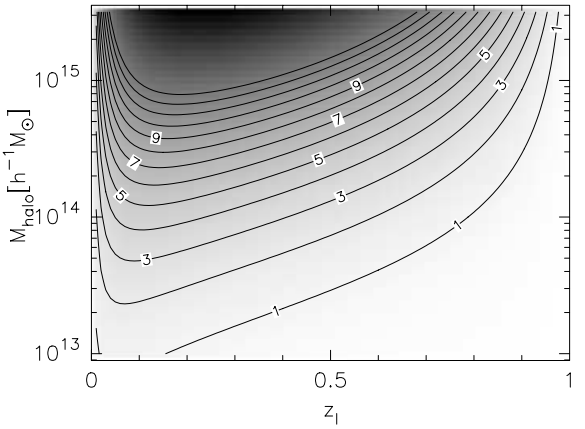

Figure 3. The grey scale with contour lines shows the S/N value for weak lensing halo detection defined by equation (14). The S/N value is computed assuming a universal NFW profile. We adopt, throughout this paper, $\sigma_\epsilon = 0.4$, $n_{\rm g} = 30$ arcmin^{-1} and $\theta_{\rm G} = 1$ arcmin.

equations 9 and 10). This makes the mass sensitivity of weak lensing surveys strongly redshift dependent. In other words, the lower mass limit of the detectable haloes depends strongly on the redshift. This figure also implies that the *effective survey volume* of a weak lensing halo search depends on the minimum halo mass that one is locating.

It should be noted that we adopt the *mean* mass–concentration relationship to compute the model prediction for the S/N value, even though the relationship has large scatters (Bullock et al. 2001; Jing 2000). As the lensing signal, $\kappa_{\rm G}$, depends on the concentration parameter by $\kappa_{\rm G} \propto c_{\rm NFW}^{1-1.5}$, the scatter likely translates to scatter in the $\kappa_{\rm G}$ value or equivalently in the S/N. It should also be noted that the lensing signal for a realistic halo is affected by substructures and asphericity of the mass distribution seen in N-body simulations. We will carefully investigate these effects, using mock data of a weak lensing survey.

2.2 Model prediction of weak lensing halo counts

Haloes that induce a sufficiently high-lensing signal are likely to be identified as peaks in the convergence map, even if the map is contaminated by noise caused by intrinsic ellipticities. Hence, it is convenient to work with the number counts of peaks above a certain threshold value: $N(\nu > \nu_{\rm th};\theta_{\rm G})$. To compute the number counts, we assume that a high peak is generated by the lensing signal from a single massive halo and that the number of false peaks owing to noise (i.e. contamination) is much smaller than that of the real peaks, as will be shown in detail later. With our fiducial choice of the smoothing scale $\theta_{\rm G} = 1$ arcmin, this assumption is reasonably valid for sufficiently high peaks (say $\nu > 4$), and thus we focus on such high peaks. In this case, the number density of peaks with height greater than the threshold $\nu_{\rm th}$ value per unit solid angle is given by (Kruse & Schneider 2000; Bartelmann, King & Schneider 2001)

$$N(\nu > \nu_{\rm th};\theta_{\rm G}, z_{\rm s}) = \int {\rm d}\chi \frac{{\rm d}V}{{\rm d}\chi} \int {\rm d}M \frac{{\rm d}n_{\rm halo}[M, z(\chi)]}{{\rm d}M} \times \mathcal{H}[\nu(M, z) - \nu_{\rm th}], \tag{15}$$

where χ is the radial comoving distance, ${\rm d}V/{\rm d}\chi$ is the comoving volume element per unit solid angle (${\rm d}V/{\rm d}\chi = \chi^2$ for a flat cosmology), ${\rm d}n_{\rm halo}/{\rm d}M$ is the halo mass function, for which we adopt the fitting function by Sheth & Tormen (1999), and $\mathcal{H}(x)$ is the

Table 1. Summary of the weak lensing peak counts computed using equation (15) for various smoothing scales. The observational parameters adopted are $\sigma_\epsilon = 0.4$, $n_g = 30$ arcmin^{-1} and a fixed source redshift of $z_s = 1$. The second and third columns are for the counts with $\nu > 5$ and $\nu > 4$, respectively.

θ_G (arcmin)	$N(\nu > 5)$ (deg^{-2})	$N(\nu > 4)$ (deg^{-2})
0.5	1.47	3.38
1	2.08	4.05
2	1.76	3.15
3	1.30	2.27
4	0.97	1.67

Heaviside step function ($\mathcal{H}(x) = 1$ for $x > 0$ and 0 otherwise). The differential counts are obtained by

$$n_{\rm peak}(\nu) = \left| \frac{{\rm d}N(\nu > \nu_{\rm th}; \theta_G, z_s)}{{\rm d}\nu} \right|. \quad (16)$$

Table 1 summarizes the predicted weak lensing peak counts for various smoothing scales. It is shown that $\theta_G = 1$ arcmin gives the largest counts, for $\sigma_\epsilon = 0.4$ and $n_g = 30$ arcmin^{-2}. In the following discussion, we will adopt this smoothing scale.

In Fig. 4, we plot the selection function of a weak lensing halo search with respect to the halo mass and redshift. The grey scale with the solid line contours shows the number density of haloes per 1 deg^2 against a given redshift interval of dz and a given logarithmic mass interval of dlog M as follows:

$$\frac{{\rm d}^2 N_{\rm halo}}{{\rm d}z {\rm d}\log M} \equiv \frac{{\rm d}V}{{\rm d}z} \frac{{\rm d}n_{\rm halo}}{{\rm d}\log M}. \quad (17)$$

Note that this is the integrand of equation (15). It is clear that, for a given halo mass, the number density rapidly increases with redshift at low redshifts ($z \lesssim 0.2$), and then gradually decreases toward higher redshifts. This is because the volume element, ${\rm d}V \propto \chi^2 {\rm d}\chi$, increases with redshift, whereas the halo number density decreases. The dashed contours show the expected peak height, $\nu_{\rm NFW}$, for the universal NFW halo at intervals of $\Delta\nu_{\rm NFW} = 1$. The comparison between the solid and dashed contours clarifies that, if a threshold peak height of $\nu_{\rm th}$=4–5 is employed, the weak lensing halo search is most sensitive to haloes that are located at the redshift between $0.1 \lesssim z \lesssim 0.7$ and have masses greater than $10^{14}\ h^{-1}\ {\rm M}_\odot$.

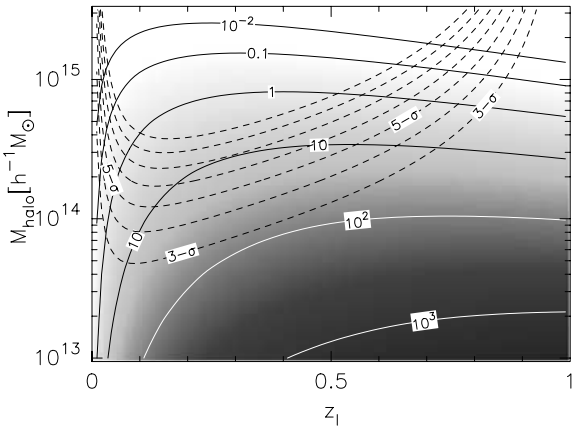

Figure 4. The grey scale with solid contour lines shows the number density of haloes per 1-deg^2 area, a given redshift interval dz and a given logarithmic mass interval d log M: $d^2 N_{\rm halo}/{\rm d}z/{\rm d}\log M \equiv {\rm d}V/{\rm d}z\, {\rm d}n_{\rm halo}/{\rm d}\log M$. This is the integrand of equation (15). The dashed contours show the *expected* peak height $\nu_{\rm NFW}$ from 3 to 9 at an interval of $\Delta\nu_{\rm NFW} = 1$. Integrating this number over the peak height above a threshold value $\nu_{\rm th}$ gives the expected number of peaks per 1 deg^2 as given by equation (15).

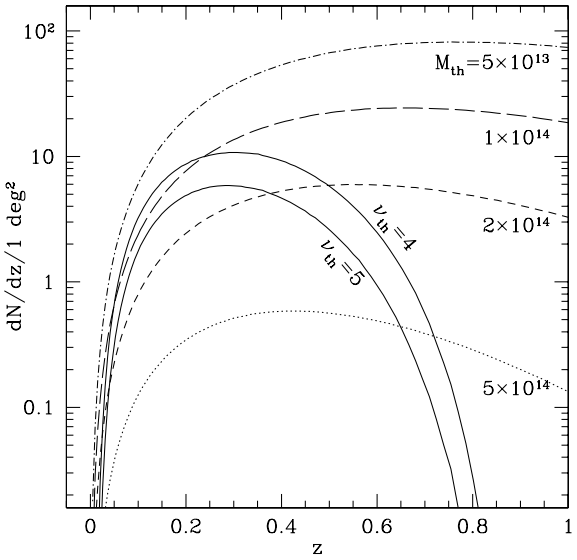

Figure 5. The solid lines show the redshift distribution of haloes with a peak height above a threshold value, $\nu_{\rm th}$, per 1 deg^2, ${\rm d}N(\nu_{\rm NFW} > \nu_{\rm th})/{\rm d}z/1$ deg^2, computed using equation (15). The upper and lower solid lines are for $\nu_{\rm th} = 4$ and 5, respectively. Note that we employed $\theta_G = 1$ arcmin and $z_s = 1$. The broken lines show the same redshift distributions but for haloes with masses above a given minimum mass, ${\rm d}N_{\rm halo}(M > M_{\rm th})/{\rm d}z/1$ deg^2.

Let us look further into the redshift and mass dependence of the weak lensing selected halo counts. The solid curves in Fig. 5 show the redshift distribution of haloes selected by the weak lensing signal. These curves can be compared with the distribution of haloes above a certain mass threshold (broken lines). A remarkable difference between these two is evidently seen at redshifts of $z > 0.5$; the distribution of weak lensing-selected haloes drops very rapidly, while the mass-selected ones decrease very gradually. This is again because the lensing efficiency peaks around $z = 0.4$ for source galaxies of $z = 1$. Fig. 6 shows essentially the same information but from a different angle. The left-hand panel shows the mass function of haloes with peak height exceeding a threshold value (the cases of $\nu_{\rm th} = 4$ and 5 are displayed). The lower mass limit in the mass function is determined by the detection threshold, while the decline in the high-mass side (which has an approximate exponential slope) is caused by one property of the halo mass function (i.e. the exponential shape of its high-mass end, see Kruse & Schneider 2000). The right-hand panels of Fig. 6 shows the mass distribution of haloes for different redshift intervals as labelled in each plot. The solid (dashed) lines are for haloes with $\nu_{\rm NFW} > 5$ ($\nu_{\rm NFW} > 4$), while the dotted lines are for all the haloes in the corresponding volume. The upper two plots clearly show that, up to $z = 0.4$, almost all the haloes with $M > 2 \times 10^{14}\ h^{-1}\ {\rm M}_\odot$ indeed produce high peaks with $\nu_{\rm NFW} > 5$. However, for higher redshifts, only very massive haloes can produce high peaks.

To summarize the results shown in Figs 4–6 , we argue that weak lensing is not an efficient technique for finding high-redshift haloes (say $z > 0.6$), except for very massive haloes of $M > 5 \times 10^{14}\ h^{-1}$ ${\rm M}_\odot$, if a mean source redshift is around $z_s = 1$. It should be noted that the capability of the weak lensing halo survey is dependent on observational conditions. As a survey is deeper, a halo search enables us to detect less massive haloes and haloes at higher redshift.

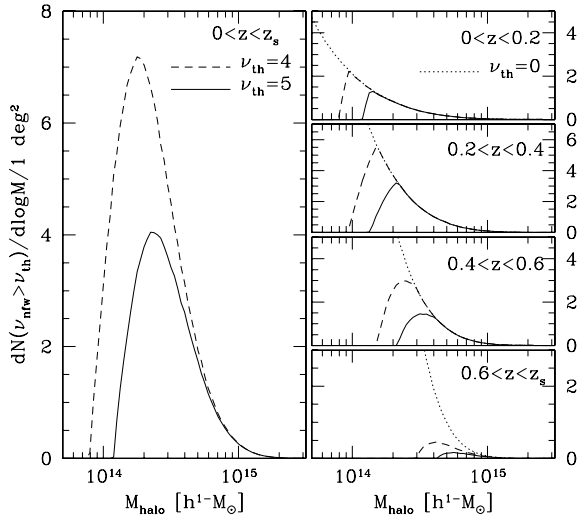

Figure 6. Left-hand panel: the mass function of haloes with $\nu_{\rm NFW} > \nu_{\rm th}$ computed using equation (15). The threshold values are $\nu_{\rm th} = 4$ and 5 for the dashed and solid lines, respectively, The source redshift is taken to be $z_{\rm s} = 1$. Right-hand panels: same as the left-hand panel but divided into four ranges of halo redshift as denoted in each plot. The dashed lines show the mass function of haloes within the corresponding volume (which is equivalent to setting $\nu_{\rm th} > 0$, that is, no limit on the weak lensing signal is imposed). Note that the scales of the y-axis are labelled on the right-hand side of the plots.

Table 2. Summary of the angular number density (per 1 deg^2) of haloes with masses more massive than a given minimum halo mass (the first column) defined in equation (17) (see Section 2 for details). The second and third columns are for the maximum redshifts of $z < 1$ and $z < 0.7$, respectively.

Halo mass ($h^{-1}{\rm M}_\odot$)	$N_{\rm halo}$ (deg^{-2})	
	$z < 1$	$z < 0.7$
$M > 5 \times 10^{13}$	55	30
$M > 1 \times 10^{14}$	17	10
$M > 2 \times 10^{14}$	4.1	2.7
$M > 5 \times 10^{14}$	0.36	0.28

Table 2 gives the number density of haloes with mass above a given minimum mass, which is obtained by integrating the halo mass function up to a given maximum redshift value. Comparison with Table 1 provides a rough estimate of how a weak lensing survey misses mass-selected haloes in the given survey volume. A large discrepancy between the halo counts and the weak lensing halo counts is seen, which arises from the redshift dependence of the lensing selection function as shown in Figs 5 and 6.

3 NUMERICAL SIMULATION AND MOCK CATALOGUES

The theoretical model prediction discussed in the previous section provides us with a basic selection function of the weak lensing cluster survey. However, since our model is based on the simple averaged descriptions of dark matter haloes, it does not take into account the individuality of haloes which causes a substantial scatter in the lensing signal and may even cause a bias. Furthermore, there are two additional sources of scatter; the projection effect and intrinsic ellipticity noise. In this section, we will examine these effects in detail using mock numerical experiments.

3.1 *N*-body simulation

We used N-body data from the Very Large N-body Simulation (VLS) carried out by the Virgo Consortium (Jenkins et al. 2001, and see also Yoshida, Sheth & Diaferio 2001 for simulation details). The simulation was performed using a parallel P^{3}M code (Macfarland et al. 1998) with a force softening length of $l_{\rm soft} \sim 30\,h^{-1}$ kpc. The simulation employed 512^3 CDM particles in a cubic box of $479\,h^{-1}$ Mpc on a side. It uses a flat cosmological model with $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$ and $h = 0.7$. The initial matter power spectrum was computed using CMBFAST (Seljak & Zaldarriaga 1996) assuming a baryonic matter density of $\Omega_{\rm b} = 0.04$. Normalization of the power spectrum is given by $\sigma_8 = 0.9$. The particle mass ($m_{\rm part} = 6.86 \times 10^{10}\,h^{-1}\,{\rm M}_\odot$) of the simulation is sufficiently small so as to practically guarantee no discreteness effect on dark matter clustering on scales down to the softening length in the redshift range of interest here (Hamana, Yoshida & Suto 2002).

3.2 Weak lensing ray-tracing simulation

The multiple-lens plane ray-tracing algorithm we used is detailed in Hamana & Mellier (2001; see also Bartelmann & Schneider 1992; Jain, Seljak & White 2000; Vale & White 2003 for the theoretical basics and technical issues); here we thus only describe the aspects specific to the VLS N-body data. To generate a density field between $z = 0$ and $z \simeq 1$, we use a stack of six snapshot outputs from two runs of the N-body simulation, which differ only in the realization of the initial fluctuation field. Each cubic box is divided into four subboxes of $479^2 \times 119.75h^{-3}{\rm Mpc}^3$ with the shorter box side being aligned with the line-of-sight direction. N-body particles in each subbox are projected on to the plane perpendicular to the shorter box side and thus to the line-of-sight direction. In this way, the particle distribution between the observer and $z \simeq 1$ is projected on to 19 lens planes separated equally by $119.75h^{-1}$Mpc. Note that to minimize the difference in redshift between a lens plane and an output of N-body data, only one half of the outputs (i.e. two subboxes) from the $z = 0$ N-body box was used.

The particle distribution on each plane is converted into the surface density field on a 4096^2 regular grid using the triangular-shaped cloud (TSC) assignment scheme (Hockney & Eastwood 1988). A 4096^2 grid is chosen to maintain the resolution provided by the N-body simulation as well as to remove at the same time the shot noise caused by the discreteness in the N-body simulation. The computation follows the procedure described in Hamana & Mellier (2001) and Jain et al. (2000).

Having produced surface density fields on all lens planes, we trace 512^2 rays backwards from the observer's point to the source plane using the multiple-lens plane algorithm (e.g. Schneider, Ehlers & Falco 1992). The initial ray directions are set on 512^2 grids with a grid spacing of 0.24 arcmin, and thus the total area covered by the rays is $2.048^2{\rm deg}^2$. We made 50 realizations of the underlying density field by randomly shifting the simulation boxes in the direction perpendicular to the line of sight using the periodic boundary conditions of the simulation.[2] Note that the lens planes coming from the same box are shifted in the same way to maintain the clustering of matter in one box.

In the following analyses, we use the lensing convergence map of the source plane closest to $z = 1$. This is, we call, the *noise-free lensing map*.

[2] The weak lensing ray-tracing simulation data used in this paper are available from T. Hamana on request, hamanatk@cc.nao.ac.jp.

3.3 Adding noise

To simulate a more realistic lensing survey, we add noises caused by intrinsic ellipticities of source galaxies to the simulated lensing map. The actual observations are further affected by other systematics such as imperfections in the telescope and/or the electric devices of the detector. Examining these effects is beyond the scope of the present paper, and thus we do not consider such observational effects.

As extensively shown in Van Waerbeke (2000), the noise field caused by intrinsic ellipticities is well approximated by the Gaussian random field on the angular scales of interest, if the ellipticities are uncorrelated between different sources (see also Jain & Van Waerbeke 2000). The Gaussian noise added to our κ map has variance as follows:

$$\sigma_{\rm pix}^2 = \frac{\sigma_\epsilon^2}{2}\frac{1}{n_{\rm g}\theta_{\rm pix}^2}, \tag{18}$$

where $\theta_{\rm pix} = 0.24$ arcmin is the pixel size of our κ maps. The noise rms on the pixels is $\sigma_{\rm pix} = 0.22$ which is much larger than the rms of the lensing convergence field.

3.4 Halo catalogues on the light cone

We identify dark matter haloes in five low-redshift outputs of the N-body simulation using the standard friends-of-friends algorithm with a linking parameter of $b = 0.164$ (in units of mean particle separation). Haloes with mass greater than $1 \times 10^{13}\, h^{-1}\, {\rm M}_\odot$, which corresponds to the mass of 146 simulation particles, are used in the following analyses. We define the halo centre to be the centre-of-mass position of the member particles.

To evaluate the halo shape, we compute the inertia tensor of the mass distribution, i.e.

$$\begin{aligned} I_{ij} &= \int {\rm d}\boldsymbol{x}^3 (x_i - \bar{x}_i)(x_j - \bar{x}_j)\rho(\boldsymbol{x}) \\ &= m_{\rm p} \sum_{k=1}^{N_{\rm p}} (x_{i,k} - \bar{x}_i)(x_{j,k} - \bar{x}_j), \end{aligned} \tag{19}$$

where $i, j = 1$–3, $\bar{x}_i$ denotes the centre-of-mass position and $N_{\rm p}$ is the number of member particles of the halo. From the inertia tensor we compute the eigenvectors and convert them to the axial ratios a/c and a/b ($a \leqslant b \leqslant c$). We also compute the trace of the inertia tensor,

$$Q \equiv \frac{1}{3}(I_{11} + I_{22} + I_{33}). \tag{20}$$

This is compared with the expected value for the NFW halo, i.e.

$$\begin{aligned} Q_{\rm NFW} &= \frac{4\pi}{3}\int_0^{r_{\rm vir}} {\rm d}r r^4 \rho_{\rm NFW}(r) \\ &= M_{\rm vir} r_{\rm vir}^2 q(c_{\rm NFW}), \end{aligned} \tag{21}$$

with

$$q(c_{\rm NFW}) = \frac{c_{\rm NFW}^3 - 3c_{\rm NFW}^2 - 6c_{\rm NFW} + 6(1 + c_{\rm NFW})\log(1 + c_{\rm NFW})}{6c_{\rm NFW}^2[(1 + c_{\rm NFW})\log(1 + c_{\rm NFW}) - c_{\rm NFW}]}. \tag{22}$$

Note that $q(c_{\rm NFW})$ can be approximated by $q(c_{\rm NFW}) \simeq 0.22 c_{\rm NFW}^{-0.26}$ for $3 \lesssim c_{\rm NFW} \lesssim 30$. The ratio $Q/Q_{\rm NFW}$ gives an estimate for the concentration of the mass distribution. We also define the following quantity as an estimate of the line-of-sight elongation of the halo mass distribution:

$$R = \frac{2I_{33}}{I_{11} + I_{22}}, \tag{23}$$

where x_3 is taken by the direction parallel to the line of sight. Although this is rather a crude estimation, $R > 1$ or $R < 1$ generally implies haloes elongated along or perpendicular to the line of sight.

Halo catalogues on the light cone are generated by stacking the simulation outputs in the same manner as in the ray-tracing experiments. We use the 50 realizations that have the same underlying matter distribution as used in the ray-tracing data. This combined data set allows us to directly relate the distribution of dark matter haloes to the weak lensing convergence map.

In summary, each halo in the mock catalogue has data on the mass, estimates of the halo shape Q and R defined by equations (20) and (23), respectively, the redshift computed from the radial distance to a halo, and the angular position converted from the spatial position. In addition, for each halo, we compute the concentration parameter and the virial radius using equations (2) and (4), from the halo mass and the redshift of the N-body data output time (i.e. $z = 0$, 0.1, 0.3, 0.5 and 0.8). These values allow us to evaluate the *expected* lensing signal using equation (13) with equation (8), which is correct only if the halo has the NFW profile. Similarly, we also compute $Q_{\rm NFW}$ using equation (21). Moreover, we introduce a 'flag' parameter for each halo, to quantify the following possible problem that arises from our use of the multiple-lens plane ray-tracing algorithm. As a result of the projection of N-body particles into four lens planes, the member particles of one halo located close to the boundary could be separated into different planes. Such a split halo would generate an artificially deformed (or even separated) lensing signal in a κ map. We assign flag $= 1$ for such haloes whose virial radius is larger than the distance between the halo centre and the closest plane, flag $= 2$ for haloes whose scaleradius is larger than the distance to the closest plane, and flag $= 0$ otherwise. The fraction of flagged haloes is very small; 0.88 per cent for flag $= 1$ and 0.19 per cent for flag $= 2$, thus it does not have significant influence on our analyses, but we carefully check if the flagged haloes cause any artificial eccentric signals in the lensing convergence map.

The *expected* lensing signal of each halo is converted into the peak height in the same way as in equation (14), which we denote as $\nu_{\rm NFW}$. The total numbers of haloes with $\nu_{\rm NFW} > 5$ and $\nu_{\rm NFW} > 4$ in our 50 realizations are 459 and 871, respectively; this is sufficiently large to allow for the statistical analyses of weak lensing haloes.

The redshift and mass distributions of haloes in the halo catalogue agree well with the predictions given using the model in Section 2.2. Fig. 7 plots the accumulate number counts of haloes against the *expected* peak height, $\nu_{\rm NFW}$. The solid curve is the theoretical prediction of equation (15) and agrees well with the simulation result. This agreement is not surprising, because ingredients of the theoretical model (the universal profile and halo mass function) are calibrated by N-body simulations.

The counts of haloes with $\nu_{\rm NFW} > 5$ in 2.048^2 deg^2 are found to be $N(\nu_{\rm NFW} > 5) = 9.18 \pm 3.63$, where the error represents the rms among 50 realizations. This is in good agreement with the theoretical prediction of 8.71. However, the rms is somewhat large and slightly larger than the Poisson fluctuation. Fig. 8 shows the distribution of numbers of haloes with $\nu_{\rm NFW} > 5$ in each realization, quite a wide spread is evidently seen. This is mainly caused by very strong

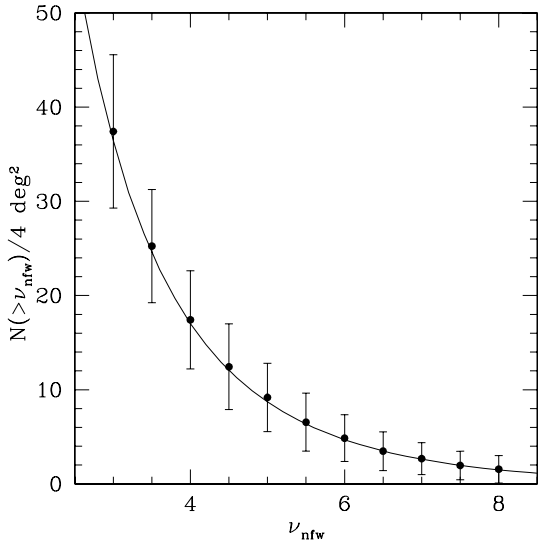

Figure 7. The number counts of the lensing peaks above a given threshold $\nu_{\rm NFW}$ per 4-deg^2 area. The solid curve shows the model prediction computed from equation (15), while the filled circles with error bars denote the mean and rms among 50 realizations of mock simulation data. Note that the peak height of a halo is estimated from the halo mass and redshift in the manner described in Section 3.4.

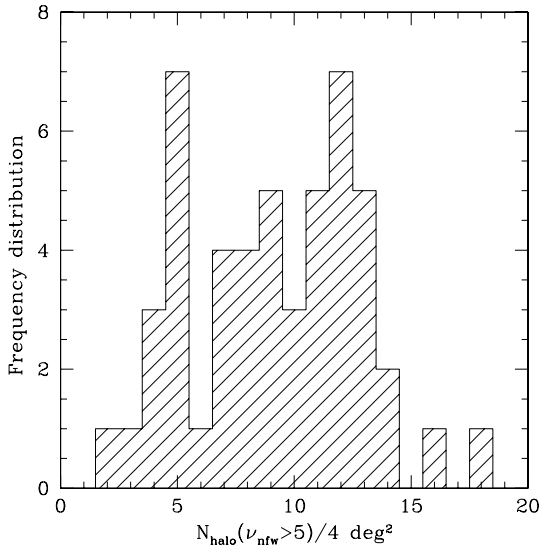

Figure 8. The frequency distribution of numbers of haloes with expected peak height (see Section 3.4 for a definition) $\nu_{\rm NFW} > 5$ from 50 realizations of 4 deg^2.

clustering of massive haloes. Accordingly, given a limited survey area, the halo counts may vary significantly from survey to survey.

The lower panel of Fig. 9 shows the probability distribution functions (PDFs) of $Q/Q_{\rm NFW}$ computed from massive haloes with $M > 3 \times 10^{13}\ h^{-1}\ {\rm M}_\odot$ ($Q_{\rm NFW}$ is computed using the mean mass–concentration relationship in equation 2). The PDF peaks at $Q/Q_{\rm NFW} \sim 1$ as expected, but has a rather broad distribution. No significant difference in the PDFs between the five different redshift boxes was found. The upper panel shows the scatter plot in the $Q/Q_{\rm NFW} - a/c$ plane for haloes with $M > 3 \times 10^{13}\ h^{-1}\ {\rm M}_\odot$ from an output of $z = 0.1$. It is clear that the mass concentration defined by $Q/Q_{\rm NFW}$ correlates with the axial ratio, which implies that an

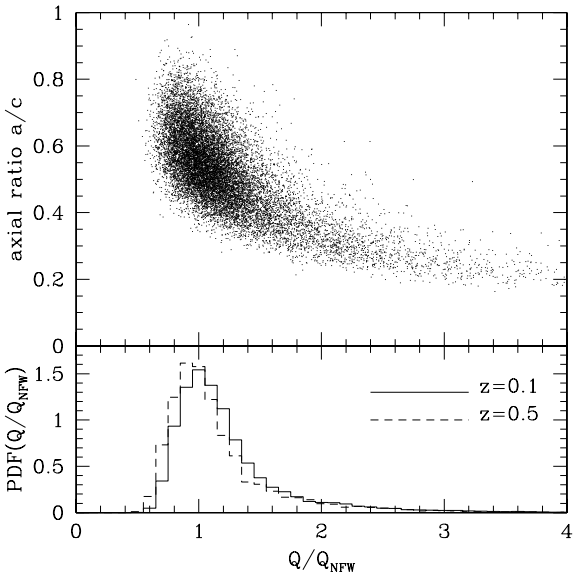

Figure 9. Upper panel: the scatter plot showing the relationship between $Q/Q_{\rm NFW}$ and the axial ratio a/c for haloes with $M > 3 \times 10^{13}\ h^{-1}\ {\rm M}_\odot$ in the $z = 0.1$ output. Here $Q_{\rm NFW}$ is defined by equation (21) and is computed using equation (4). Lower panel: the PDF of $Q/Q_{\rm NFW}$ of haloes with $M > 3 \times 10^{13}\ h^{-1}\ {\rm M}_\odot$, the solid and dashed lines for $z = 0.1$ and $z = 0.5$ outputs, respectively.

extended mass distribution is likely to be associated with eccentric mass distributions such as haloes having substructures or an ongoing merger. We note that the PDF of the axial ratio peaks at $a/c \sim 0.5$ and thus most of the haloes significantly deviate from spherical symmetry. This broad distribution leads to significant scatter in the relationship between the expected and observed lensing signals as seen below.

Fig. 10 shows the PDF of the shape parameter, R, defined by equation (23) for massive haloes with $M > 3 \times 10^{13}\ h^{-1}\ {\rm M}_\odot$. A broad distribution is evidently seen and we checked that the distribution

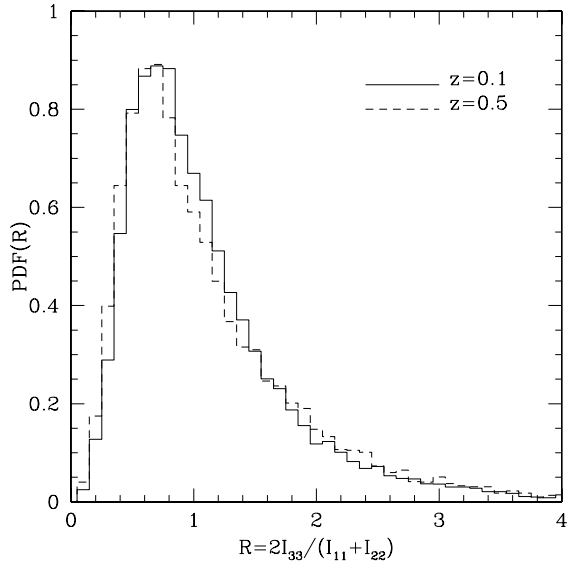

Figure 10. The PDF of the R defined by equation (23) for haloes with $M > 3 \times 10^{13}\ h^{-1}\ {\rm M}_\odot$ at $z = 0.1$ (solid line) and $z = 0.5$ (dashed line) outputs. Note that the median of the distributions is ∼0.92, and 55 per cent of all the samples has $R < 1$.

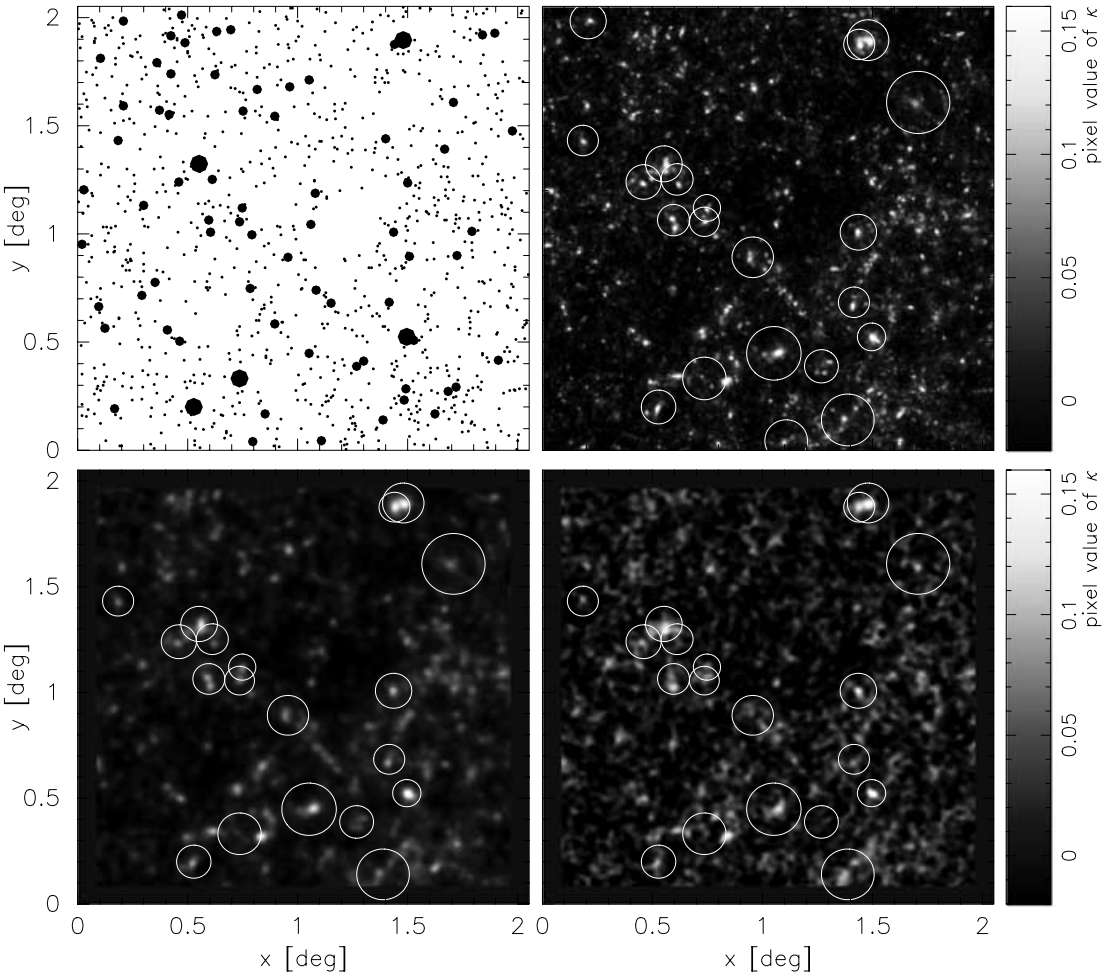

Figure 11. Upper left: the celestial distribution of haloes identified from one realization of mock light-cone data generated stacking N-body outputs. The large filled circles denote haloes with $M_{\rm halo}[h^{-1}\,{\rm M}_\odot] > 3\times10^{14}$, the small filled circles are for $3\times10^{14} > M_{\rm halo}[h^{-1}\,{\rm M}_\odot] > 8\times10^{13}$ and the dots are for $8\times10^{13} > M_{\rm halo}[h^{-1}\,{\rm M}_\odot] > 1\times10^{13}$. Only the haloes within the redshift interval of between 0.1 and 0.7 are displayed. Upper right: the grey map shows the lensing convergence map on 512^2 pixels with a pixel size of 0.24 arcmin obtained from a ray-tracing experiment through the same mass distribution as used in the upper-left panel. The circles show the positions of haloes with $\nu_{\rm NFW} > 4$, the centre of which is set at the position of each halo and the radius of which shows the projected virial radius (see Section 3.4 for details). Bottom left: same as the top-left panel but the convergence map is smoothed by the Gaussian window function with $\theta_{\rm G} = 1$ arcmin. The edge regions of the field with a 5-arcmin width are masked owing to incomplete smoothing caused by the lack of κ data on the outer region. The grey scale of this plot is the same as that of the bottom-right panel. Bottom right: same as the bottom-left panel but the Gaussian noise is added before smoothing is applied. From comparison with the other plots, it is clear that high peaks resulting from massive haloes are still identifiable.

does not largely change over the redshift range under consideration ($z < 0.8$). It is important to note that the median of the distribution is less than unity ($\sim$0.92), thus more than half of the massive haloes (specifically 55 per cent for haloes with $M > 3\times10^{13}\,h^{-1}\,{\rm M}_\odot$ irrespective of the redshift) have $R < 1$. This is a natural consequence of the random orientation of the haloes' major axis; among the spatial three directions, it is more likely that the major axis of a halo is elongated along either of two directions perpendicular to the line of sight rather than along it. The skewed distribution leads to systematic bias in the relationship between the expected and observed lensing signals (see Section 5.2 and Appendix C for details).

3.5 Visual impressions

Fig. 11 shows the lensing convergence maps and the celestial distribution of the haloes, which are taken from one of the 50 realizations. In the top-right-hand panel, the grey scale shows the κ value with 512^2 pixels with a pixel size of 0.24 arcmin. The bottom-left-hand panel shows the same convergence map as in the top-right-hand panel but after being smoothed by the Gaussian filter function of $\theta_{\rm G} = 1$ arcmin. Adding the Gaussian noise yields the noisy map shown in the bottom-right-hand panel. While the noise indeed alters the appearance of the convergence map, the very high peaks remain identifiable.

The top-left-hand panel shows the distribution of all the haloes identified by the friends-of-friends (FOF) scheme in the original N-body simulations from which the convergence maps in the other plots are generated. Note that the point size scales with the halo masses. A good correspondence with respect to position between massive haloes and high lensing signals on the κ map is seen. This correspondence is more clearly seen if haloes are selected by means of the expected $\nu_{\rm NFW}$ value: in the convergence maps (the top-right-hand and two bottom panels) the positions of haloes with $\nu_{\rm NFW} > 4$ are marked by the circles with their centres being placed at the centre-of-mass of the halo, and their radii show their angular virial radius (see Section 3.4). The improvement in the correspondence

reflects the fact that the lensing signal is determined by combined contributions from the halo mass, the redshift and the halo concentration, rather by the mass alone.

4 RESULTS ON THE PEAK COUNTS

In this section, we study the number counts of high peaks in the lensing convergence map. To do this, we use three kinds of κ maps: (i) the noise-free κ maps, (ii) the noisy κ maps where the Gaussian noise was added in the manner described in Section 3; and (iii) the pure-noise maps without the cosmological lensing signals. All these maps are smoothed with a Gaussian window function with $\theta_{\rm G} = 1$ arcmin (equation 11). We identify positive (negative) curvature peaks in the maps as pixels that have higher (lower) values of κ than all the surrounding eight pixels (Jain & Van Waerbeke 2000, and see also Miyazaki et al. 2002 for the first application to actual observational data). The peak height is defined as the ratio of the convergence value at the peak to the noise rms, as given by equation (14). For all the three cases, $\sigma_{\rm noise} = 0.02$ ($\sigma_\epsilon = 0.4$ and $n_{\rm g} = 30$ arcmin^{-2}) is adopted.

4.1 The peak PDF

The PDF of peaks in the convergence map is studied in detail in Jain & Van Waerbeke (2000). Here we briefly summarize the characteristics of the peak PDF that are relevant for the subsequent discussion.

Fig. 12 shows the peak PDFs in the noise-free κ (dashed histogram), the noisy κ map (solid histogram) and the pure-noise map (dotted histogram). The upper panel shows the sum of the positive and negative curvature peaks, while the lower panel shows the contribution from only the positive peaks. In the absence of noise, the peak PDF reflects the characteristics of non-linear gravitational clustering. Most of the negative peaks are a result of negative curvature peaks (troughs). There is a cut-off in the negative peak height because the mass density contrast cannot be lower than -1. One can also find a tail toward the high positive peak height which arises from the lensing signals caused by massive haloes (Kruse & Schneider 2000). Another noticeable feature seen from the lower panel is that some of the *positive* curvature peaks have a *negative* peak height, which occurs when the peak caused by a halo is superposed on a void region in the same line of sight. If the noise field is Gaussian, the PDF from the pure-noise map can be derived from Gaussian random field theory (Bardeen et al. 1986; Bond & Efstathiou 1987; Van Waerbeke 2000). The symmetric bimodal feature of the peak PDF is owing to the superposition of two contributions from peaks with positive and negative curvatures. While adding the noise to the cosmological κ map alters the peak PDF drastically, there remain some characteristic features that contain useful cosmological information. First, the asymmetric double peaks are caused by the skewed feature of the cosmological κ PDF as stated above. Second, there are excesses over the noise PDF on both the high positive ($\nu > 3$) and negative peaks ($\nu < -3$). These features were indeed detected in the recent weak lensing survey by Miyazaki et al. (2002).

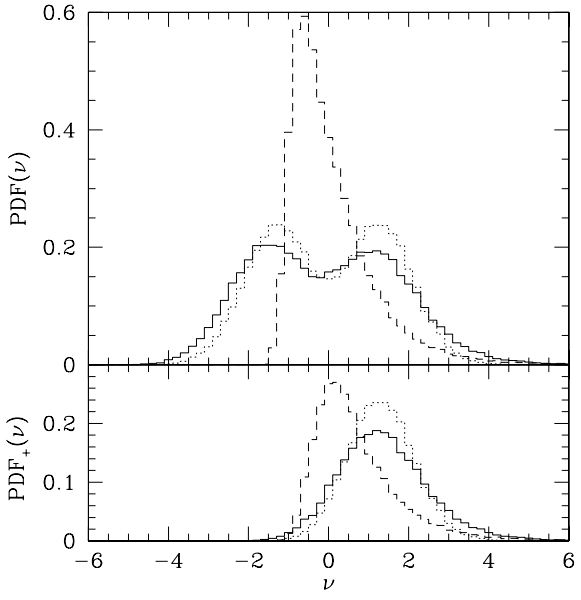

Figure 12. The simulation results for the PDFs of the peaks in the noise-free κ map (dashed histogram), in the noisy κ map (solid histogram) and in the pure-noise map (dotted histogram). Shown is the cumulative results from the 50 realizations. The upper panel shows the sum of the positive and negative curvature peaks, while the lower panel shows the contribution from the positive peaks only. In the total 50×2.048^2-deg^2' area, there are 59 908, 107 165 and 114 250 peaks for the noise-free κ map, the noisy κ map and the pure-noise map, respectively. The corresponding numbers for positive curvature peaks are 31 043, 53 123, and 57 141.

It is worth pointing out that the pure-noise PDF can be directly estimated from actual observed data. First, we can generate a pure-noise map by applying mass reconstruction to a randomized catalogue where the orientation of the source galaxy images is randomized (to remove coherent image distortions from lensing signals). Then we can measure the peak PDF from the pure-noise map, as performed in Miyazaki et al. (2002). This allows one to estimate the contamination rate (the false peak rate) directly from the data.

4.2 High-peak counts

In the following, we restrict our discussion to the number counts of high peaks that are mainly a result of massive haloes. Fig. 13 shows the number density of (positive curvature) peaks per 1 deg^2. The

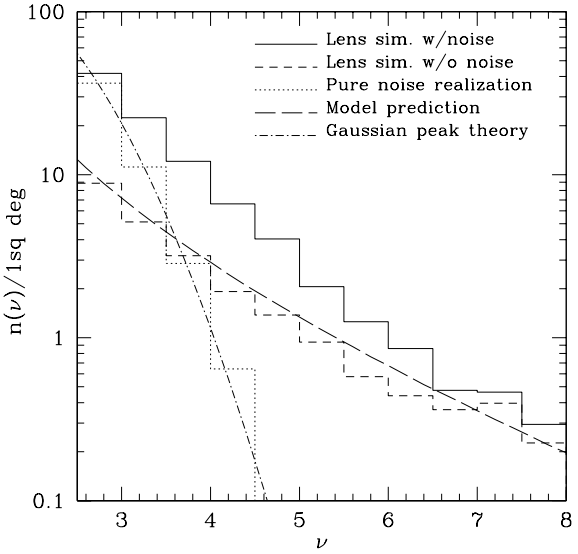

Figure 13. Counts of peaks per 1 deg^2. The histograms are from the numerical experiments; the noisy κ maps (solid), the noise-free κ maps (dashed) and the pure-noise realizations (dotted). The long-dashed line shows the prediction using the model (equation 16 with equation 15), and the dot-dashed line shows the prediction using Gaussian random field theory (e.g. Bond & Efstathiou 1987; Van Waerbeke 2000).

histograms show the average values computed from the 50 realizations of the numerical experiments; the noisy κ maps (solid), the noise-free κ maps (dashed) and the pure-noise maps (dotted). The long-dashed line shows the predictions of the model from equation (16), while the dot-dashed line shows the prediction for the pure-noise counts (Van Waerbeke 2000). When comparing the results for the lensing maps with and without noise, it is clear that the noise significantly increases the peak counts over the range of peak heights under consideration. Even high peaks with $\nu > 5$ are similarly boosted, while there is no such event in the pure-noise counts, as shown by the dotted curve. There are two main reasons for this enhancement as follows (see Section 5.3 and Appendix B for further details). First, the addition of noise leads to modifications in the peak height as well in the peak position. Second, the noise generates spurious peaks in the noisy maps. In particular, the latter effect is likely to show up in a biased way depending on the peak height. As the convergence field around a given high peak has greater amplitude than the average field, adding the noise to such a region is more likely to induce spurious high peaks around the real peak. We develop a correction scheme to the model peak counts in Appendix B. The corrected counts are in reasonable agreement with the simulation result for high peaks, $\nu > 5$.

5 RESULTS ON THE CORRESPONDENCE BETWEEN HALOES AND PEAKS

In this section, we examine the correspondence between massive haloes and peaks in the weak lensing convergence maps, using the results of our ray-tracing experiments and the halo catalogues created from the *N*-body simulation outputs. In doing so, we follow a realistic procedure which is to be used in weak lensing surveys: we focus on peaks above a certain threshold in the lensing map as being strong candidates for haloes. The usefulness of the weak lensing halo search can be quantified in terms of *completeness* and *efficiency*. As will be shown later in detail, choosing a threshold of $\nu_{th} \simeq 4$–5 gives an optimal balance between efficiency and completeness. We will also carefully examine if there is a correlation between the shape of haloes and their detectability, because this could produce a bias in halo catalogues constructed from the weak lensing method.

To make realistic weak lensing survey catalogues, we impose a criterion on the field selection of simulations so that no nearby massive halo is included in the target fields. Specifically, we discard a simulated field if there is a halo (or possibly more than one) with $M_{\rm halo} > 10^{14}\,h^{-1}\,{\rm M}_{\odot}$ at $z < 0.1$. Similar criteria, such as no known nearby X-ray luminous cluster of galaxies or no nearby optically identified massive clusters, are usually imposed on the field selection of an actual observation, because a nearby massive cluster would occupy a large portion of a survey field and makes a distant halo search difficult and inefficient. Of the simulated fields 31 out of 50 pass this criterion. In addition, we discard one field in which there is a massive halo with $\nu_{\rm NFW} > 5$ having flag = 2 (see Section 3.4 for a description of flag).[3] We thus analyse the remaining 30 fields. We note that we do not use data within 5 arcmin of the edge of each simulated field owing to the incomplete smoothing caused by the lack of data at the outer region. The total field size is then 106 deg^2.

[3] Including this field in our analysis does not change our conclusions, but we decided to discard it for having conclusive results.

5.1 Matching haloes with peaks in lensing mass maps

We shall now examine the relationships between the peak height in weak-lensing maps and the properties of haloes through the *expected* peak height for the NFW haloes ($\nu_{\rm NFW}$, see Section 2.1). We restrict our discussion to haloes with $\nu_{\rm NFW} > 3$ only. The halo catalogues contain 856 such haloes in total. Alternatively, the ray-tracing simulations directly enable us to find peaks in the κ maps. We consider two kinds of the κ maps; those with and without the noise caused by intrinsic ellipticities. The peaks in the two maps are identified in the manner described in Section 4. The noise-free maps contain 669 peaks with $\nu_{\rm lens} > 3$, while the noisy maps contain 2417 peaks with $\nu_{\rm noisy} > 3$. Thus, the noise significantly enhances the peak counts. It is also worth noting that the 669 peaks in the noise-free maps are significantly smaller than the *expected* number of peaks from the halo catalogues. We will carefully investigate this discrepancy in the subsequent sections.

We consider haloes and peaks with only $\nu > 3$ to avoid mismatches between peaks and physically unrelated haloes. With the large number density of low-mass haloes it should be noted that including such haloes rapidly increases the chance of mismatching. This is especially the case when one attempts to match peaks in the noisy κ maps with haloes because false peaks caused by noise can be accidentally matched with physically unrelated haloes. The criterion $\nu > 3$ is chosen based on the fact that most such (high) peaks are expected to be caused by real haloes, even when the lensing map is contaminated by noise (see Fig. 13). In addition, the criterion value is reasonably lower than the observational threshold value (generally defined by $\nu \gtrsim 4$). This guarantees most haloes of interest to us are included in the final halo catalogue.

We carry out the peak–halo matching in two directions: whether a peak has a corresponding halo in the halo catalogue; and whether a halo has a counter-peak in the simulated κ map. Note that the peak position does not necessarily agree with the centre-of-mass of the halo exactly and this is more likely the case for haloes that have substructures or are in the merger process. For this reason, we search for a matched pair candidate within a radius of 12 pixel (2.88 arcmin) from the peak position or from the halo centre. This maximum angular separation of 2.88 arcmin is chosen so that it is larger than the smoothing radius of $\theta_{\rm G} = 1$ arcmin, but still smaller than the angular virial radius of a massive halo at $0.1 \lesssim z \lesssim 0.7$. If there are more than one pair candidate within the radius, we take the closest one as the primary candidate and consider the others as the secondary candidate.

We classify the peak–halo matching into the following five classes.

(i) Class (i) is that both directions agree; both a halo and a peak are the primary pair candidate.

(ii) Class (ii) is defined as haloes that have no paired peak within a radius of 12 pixels from the halo centre. We will call such haloes as '*missing haloes*'.

(iii) Class (iii) is defined as peaks with no paired halo within a radius of 12 pixels from the peak position.We will often refer to such peaks as 'false peaks'.

(iv) Class (iv) denotes a halo that has a matched peak candidate, but is not the primary matched halo of that peak.

(v) Class (v), which is the opposite to the class (iv), is that a peak has a matched halo candidate but is not the primary matched peak of that halo.

The results for the noise-free and noisy maps are summarized in Table 3.

Table 3. Classification of matching relationships between haloes and peaks. The double-lined arrows denote the primary match, while the single-lined arrows denote the secondary match. The last two columns summarize the number of haloes/peaks in each class, the third (fourth) column is for matching between the halo catalogue and the peak list from the noise-free (noisy) κ map. Note that haloes with $\nu_{\rm NFW} > 3$ (856 haloes in total) and peaks with $\nu_{\rm lens} > 3$ (669 peaks) and $\nu_{\rm noisy} > 3$ (2417 peaks) are considered.

Class	Matching relationship	Number of objects	
		$\nu_{\rm NFW} - \nu_{\rm lens}$	$\nu_{\rm NFW} - \nu_{\rm noisy}$
i	halo$\rightleftarrows$ peak	475	624
ii	halo with no paired peak	367	222
iii	peak with no paired halo	177	1637
iv	halo$\rightleftarrows$ peak	14	10
v	halo$\rightleftarrows$ peak	17	156

The typical properties of haloes or peaks in classes (ii)–(v) are explained as follows. Class (ii) haloes are those which have a peak height lower than the criterion value ($\nu = 3$) although the expected peak height, $\nu_{\rm NFW}$, exceeds it. This is mainly caused by departures from the NFW profile or/and the projection effect. In the class (iii) peaks there are two kinds of peaks. The first are peaks which have corresponding haloes but have expected peak heights, $\nu_{\rm NFW}$, lower than the criterion value, even though the measured peak heights in a κ map exceed it. This is again caused by departures from the NFW profile or/and the projection effect. The other type are peaks which do not have corresponding haloes but are generated by the noise. The former accounts for most of the class (iii) peaks in noise-free κ maps, while the latter accounts for a large fraction of those in noisy κ maps. Class (iv) haloes are caused by the projection effect; if two or more haloes are located at different redshifts along almost the same line of sight, after smoothing the κ map only one peak appears, because the smoothing merges those into one structure. Finally, in class (v) peaks there are two kinds of peaks. The first are peaks caused by substructures within a halo. This accounts for most of the class (v) peaks in noise-free κ maps. The other are false peaks caused by the noise generated around the primary peak owing to the massive halo. Because the convergence field around the high peak is biased and thus larger than the average field, adding the noise into such a region is more likely to generate false peaks with a relatively high peak height. This accounts for most of the class (v) peaks in noisy κ maps.

In the rest of this section, we consider only objects from classes (i)–(iii), and do not deal with objects from classes (iv) and (v), because objects from classes (iv) and (v) are secondary matched pairs associated with nearby primary matched pairs. This roughly mimics the actual observation procedure in which the secondary objects are investigated in detail in the course of a follow-up observation of the primary matched object.

5.2 Haloes versus peaks in noise-free κ maps

Let us first look into the correspondence between high peaks identified in the noise-free κ maps ($\nu_{\rm lens}$) and properties of haloes through the *expected* peak height ($\nu_{\rm NFW}$). The main purpose of this subsection is to examine the amplitude of the irreducible scatter in the relationship induced by both the individuality of the mass distribution of a halo and the projection effect.

Fig. 14 shows the distribution of the objects of classes (i)–(iii) in the $\nu_{\rm NFW} - \nu_{\rm lens}$ plane. As class (ii) haloes do not have a matched peak with height $\nu_{\rm lens} > 3$ in the κ map (although the haloes have a high expected peak height of $\nu_{\rm NFW} > 3$), such haloes are placed in a narrow locus around $\nu_{\rm lens} = 2$ for illustrative purposes. For similar reasons, the class (iii) peaks are placed around $\nu_{\rm NFW} = 2$. The distribution of the class (i) haloes shows large scatter around the ideal relationship of $\nu_{\rm lens} = \nu_{\rm NFW}$ (shown by the solid line in the upper panel) over the range of ν under consideration. It is worth noting that the distribution for a fixed $\nu_{\rm NFW}$ value is slightly biased toward lower $\nu_{\rm lens}$ values. The mean and rms of the differences, $\nu_{\rm lens} - \nu_{\rm NFW}$, are -0.24 and 1.2, respectively. As no noise is added to the κ map in this case, the scatter and bias must be ascribed to two effects as follows. (i) The mass distribution within a halo is not spherical nor completely universal, causing deviations of the peak heights from the expected values computed in our simple model. (ii) A chance projection of structures along the same line of sight alters the lensing signal caused by the halo. For the former, although we have so far used the average relationship for the halo concentration, as given by equation (2), it was shown in Jing (2000) that the halo concentrations among haloes with a given mass obey the log-normal distribution with variance $\sigma(\ln c_{\rm NFW}) \approx 0.2$. This roughly corresponds to a 20 per cent rms scatter in the $c_{\rm NFW}$ relationship. Given the approximate relationship $\kappa \propto c_{\rm NFW}^{1-1.5}$ for the peak height caused by a massive halo (see Section 2), the rms scatter in the $c_{\rm NFW}$ distribution translates into $20 - 30$ *percent* rms scatters in the lensing signal $\nu_{\rm NFW}$ and therefore equivalently in the $\nu_{\rm lens} - \nu_{\rm NFW}$ relationship. For example, for haloes with $\nu_{\rm NFW} = 5$, a rms of $\sigma_\nu \sim 1$ results. Second, lensing projection caused by physically unrelated structures along the same line of sight of a halo also contributes to the scatter in $\nu_{\rm lens}$. Following Hoekstra (2001), this scatter can be estimated from the cosmic weak lensing variance under the assumption that the halo and the large-scale structures are uncorrelated. For the Gaussian smoothing scale $\theta_{\rm G} = 1'$, the amplitude of the (linear) convergence variance is $\sigma_\kappa(1') = 0.01$ for the ΛCDM model, corresponding to the rms deviation of $\sigma_\nu = 0.5$ caused by the noise, $\sigma_{\rm noise}(1') = 0.02$ (see equation 12). Hence, these two effects can basically explain the amplitude of the scatter in the $\nu_{\rm lens} - \nu_{\rm NFW}$ relationship in Fig. 14.

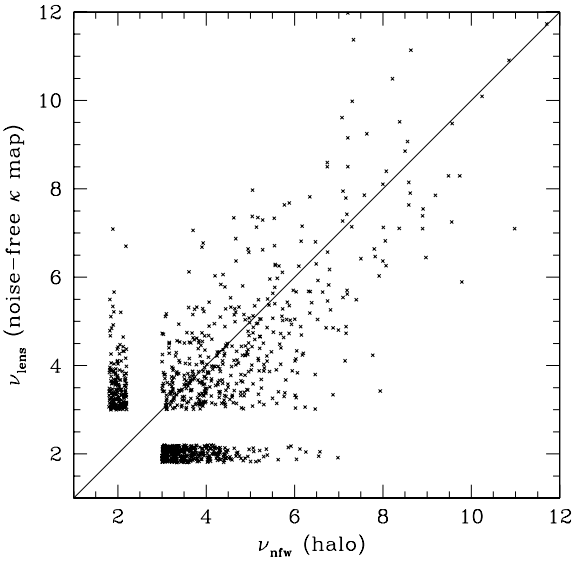

Figure 14. The correspondences between haloes and peaks identified in the noise-free κ maps in the $\nu_{\rm NFW} - \nu_{\rm lens}$ plane. Objects from classes (i)–(iii) are shown. The class (ii) objects are placed within a narrow track around $\nu_{\rm lens} \sim 2$, regardless of their true value, and the class (iii) objects are also placed at $\nu_{\rm NFW} \sim 2$.

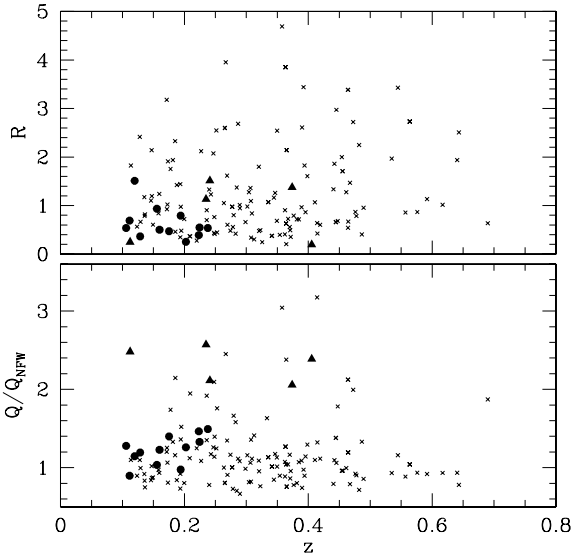

Figure 15. The scatter plot for the distribution of the haloes in the $z - Q/Q_{\rm NFW}$ plane (lower panel) and in the $z - R$ plane (upper panel). The filled circle and triangle symbols are for the class (ii) haloes with $\nu_{\rm NFW} > 5$. For comparison, class (i) haloes with $5 < \nu_{\rm NFW} < 7.5$ are also plotted by the crosses. It is apparent from the lower panel that the class (ii) haloes can be separated into two subclasses: those with $Q/Q_{\rm NFW} > 2$ are plotted by the filled triangles, while the rest are plotted by the filled circles.

We also argue that the systematic bias in the $\nu_{\rm lens} - \nu_{\rm NFW}$ relationship can be explained by the following two effects. The first is the fact that more than half of the haloes have $R < 1$ as discussed in Section 3.4 and Fig. 10. As R positively correlates with $\nu_{\rm lens}/\nu_{\rm NFW}$ (see Fig. C2 in Appendix C), it causes the bias. The second is the projection with a void region in the same line of sight which reduces the lensing signal (see Metzler et al. 2001; White et al. 2002; Padmanabhan et al. 2003, for detailed study on the projection effect).

Let us turn to the class (ii) haloes in Fig. 14. Among the 367 class (ii) haloes, there are 17 haloes that are supposed to produce peaks with $\nu > 5$ in the κ maps if the haloes have the NFW profile (i.e. $\nu_{\rm NFW} > 5$ in our notation). Why are these haloes missing? Fig. 15 explains this in terms of deviation of the mass distribution within a halo from the spherically symmetric NFW profile. We again use the parameters Q and R to quantify the mass distribution within a halo (see equations 20 and 23, respectively): $Q/Q_{\rm NFW} < 1(>1)$ characterizes a halo that has centrally more (less) concentrated mass distribution than expected from the NFW profile, while $R > 1(<1)$ describes that the mass distribution is elongated along (perpendicular) the line of sight ($R = 1$ for a spherically symmetric halo). Therefore, a halo with $Q/Q_{\rm NFW} > 1$ and/or $R < 1$ probably leads to a smaller lensing signal than that expected from the NFW profile (see Appendix C for more detailed analysis of the relationships between the halo shape and the amplitude of the lensing signal). The upper panel of Fig. 15 shows the distribution of the class (ii) haloes in the $z - Q/Q_{\rm NFW}$ plane, while the lower panel shows the distribution in the $z - R$ plane. For comparison, the figure also shows the distributions for the class (i) haloes with $5 < \nu_{\rm NFW} < 7.5$ by the cross symbols. From the lower panel, it is clear that class (ii) haloes can be divided into two distinct subgroups; those which have $Q/Q_{\rm NFW} > 2$ (plotted by the filled triangles) or $Q/Q_{\rm NFW} < 2$ (plotted by the filled circles). We indeed find that haloes with $Q/Q_{\rm NFW} > 2$ show clear signatures of substructures or ongoing merging, resulting in an extended mass distribution (the κ map of such haloes is shown in Fig. A1, for example). The remaining objects with $Q/Q_{\rm NFW} < 2$ tend to be located at a low redshift ($z < 0.25$), where the lensing efficiency is relatively small for the source redshift, $z_{\rm s} = 1$, under consideration, and have small R values, mostly $R < 1$. The examples are shown in Figs A2 and A3. These haloes typically have prominent substructures and the whole mass distribution within the halo is indeed elongated along the perpendicular direction to the line of sight. With the exception of such irregular systems, almost all massive haloes with $\nu_{\rm NFW} > 5$ generate a high peak with $\nu_{\rm lens} > 3$. It is important to note that the distribution of the class (ii) haloes in Fig. 15 is not clearly distinguished from that of the class (i) haloes, reflecting the fact that the lensing signal is also affected by the projection effect.

We turn next to the class (iii) objects in Fig. 14 which are peaks with $\nu_{\rm lens} \geqslant 3$ in the lensing map, but do not have corresponding haloes in the halo catalogue (contains only haloes with $\nu_{\rm NFW} > 3$). Among the class (iii) objects there are eight events with $\nu_{\rm lens} > 5$ which we call 'false peaks'. We carefully looked into each case in the κ map and searched for a possibly related halo in the halo catalogue. We found that the five events are associated with the substructures of massive haloes or the secondary local peaks of merger systems (note that the peaks are separated from the primary peak of the halo centre by more than our searching distance, 12pixels = 2.88 arcmin). Fig. A4 shows the κ map of one such case. The remaining three peaks are caused by a chance projection of two moderately massive haloes at different redshifts (Fig. A5 shows one example). Therefore, with the except of the three chance projection systems, the *false peaks* with $\nu_{\rm lens} > 5$ are all associated with a real massive halo as the secondary peak.

To quantify the ability of a weak lensing survey to locate massive haloes, we consider its *completeness* and *efficiency*. We define the *completeness* by the fraction of haloes, which are detected as high peaks above a given threshold, $\nu_{\rm th}$, in the κ map, relative to all the haloes with the expected lensing signal, $\nu_{\rm NFW}$, greater than a given minimum value $\nu_{\rm min}$:

$$f_{\rm c}(\nu_{\rm lens} > \nu_{\rm th} \mid \nu_{\rm NFW} > \nu_{\rm min}) = \frac{N_{\rm i}(\nu_{\rm lens} > \nu_{\rm th} \mid \nu_{\rm NFW} > \nu_{\rm min})}{N_{\rm i+ii}(\nu_{\rm NFW} > \nu_{\rm min})}, \quad (24)$$

where $N_{\rm i+ii}(\nu_{\rm NFW} > \nu_{\rm min})$ denotes the number of haloes with $\nu_{\rm NFW} > \nu_{\rm min}$ which, under our classification, corresponds to the number of such haloes in classes (i) and (ii), and $N_{\rm i}(\nu_{\rm lens} > \nu_{\rm th} | \nu_{\rm NFW} > \nu_{\rm min})$ denotes the haloes with $\nu_{\rm NFW} > \nu_{\rm min}$ matched with the peaks with $\nu_{\rm lens} > \nu_{\rm th}$ (these are thus class (i) objects). Similarly, we define the *efficiency* by the fraction of peaks, which are associated with haloes with the expected peak height exceeding a given minimum value $\nu_{\rm min}$, relative to all the peaks with height greater than a given threshold $\nu_{\rm th}$

$$f_{\rm e}(\nu_{\rm NFW} > \nu_{\rm min} \mid \nu_{\rm lens} > \nu_{\rm th}) = \frac{N_{\rm i}(\nu_{\rm NFW} > \nu_{\rm min} \mid \nu_{\rm lens} > \nu_{\rm th})}{N_{\rm i+iii}(\nu_{\rm lens} > \nu_{\rm th})}, \quad (25)$$

where $N_{\rm i+iii}(\nu_{\rm lens} > \nu_{\rm th})$ denotes the number of peaks with $\nu_{\rm lens} > \nu_{\rm th}$ which corresponds to the number of such peaks in classes (i) and (iii).

The results are plotted in the left-hand panels of Fig. 16 as a function of $\nu_{\rm th}$. The upper panels show the completeness ($f_{\rm c}$) and the lower panels show the efficiency ($f_{\rm e}$). As expected, these two estimates behave inversely, that is, when $f_{\rm e}$ goes up, $f_{\rm c}$ goes down. Therefore the optimal choice of the $\nu_{\rm th}$ value should be determined by a balance between these two estimates, which we discuss in the latter section with taking into account the noise effects.

The completeness plot shows that only about half of the haloes with $\nu_{\rm NFW} > 5$ produce peaks with $\nu_{\rm lens} > 5$ in the noise-free κ maps, and that 77 per cent of the haloes have $\nu_{\rm lens} > 4$. As expected, the completeness deteriorates if $\nu_{\rm th} \gtrsim \nu_{\rm min}$. It is also shown that a good level of completeness (say $f_{\rm c} \gtrsim 0.8$) is attainable if one sets

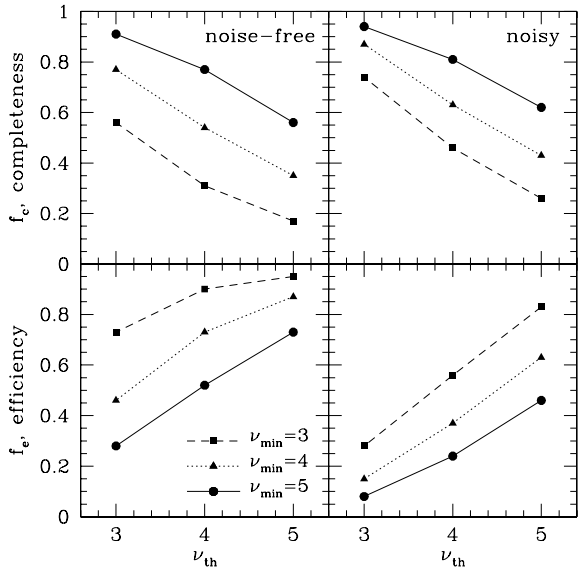

Figure 16. Upper panels: the completeness, $f_c(\nu > \nu_{th}|\nu_{NFW} > \nu_{min})$, defined by equation (24), as a function of the threshold peak height ν_{th}. Three cases for the minimum expected peak value, $\nu_{min} = 3$, 4 and 5 are shown by the filled squares, triangles and circles, respectively. Lower panels: the efficiency, $f_e(\nu_{NFW} > \nu_{min}|\nu > \nu_{th})$, defined by equation (25), as a function of the threshold peak height ν_{th}. The left panels are for the noise-free case ($\nu = \nu_{lens}$), while the right panels are for the noisy case ($\nu = \nu_{noisy}$).

$\nu_{th} \lesssim \nu_{min} - 1$. This can be understood from the fact that the rms scatter caused by the individuality of haloes and the projection effect is $\sigma_\nu \sim 1.2$. Note that some fraction of the haloes with ν_{lens} values that are much smaller than the expected peak heights ν_{NFW} are irregular systems or unrelaxed systems, and therefore the completeness is likely to be improved for samples of regular, well-relaxed haloes. It should be also noted that noise increases the completeness (see Section 5.4 and the upper right-hand panel of Fig. 16).

Let us turn to the efficiency issue. As no noise is added to the κ maps, the efficiency given here should be regarded as an estimate for an ideal lensing survey. Clearly, the efficiency is very high. If one sets the detection threshold to $\nu_{th} = 5$, 73 per cent of all the peaks correspond to real haloes with $\nu_{NFW} > 5$, and 95 per cent for $\nu_{NFW} > 3$. Even for $\nu_{th} = 4$ the efficiency is still high, 90 per cent for $\nu_{NFW} > 3$. Therefore, in the absence of noise, the contamination rate (the fraction of false peaks) is very small for a detection threshold of $\nu_{th} > 4$.

Fig. 17 shows the scatter plot of the haloes of classes (i) and (ii) in the redshift–halo mass plane. The contours show tracks of the expected peak height ν_{NFW} from 3 to 9 at an interval of $\Delta\nu_{NFW} = 1$. The class (i) haloes matched with peaks in the *noise-free* κ maps are plotted in the upper-left-hand ($\nu_{lens} > 5$), upper-right-hand ($5 > \nu_{lens} > 4$) and lower-left-hand panels ($4 > \nu_{lens} > 3$), respectively. The lower-right-hand panel shows the scatter plot for the class (ii) haloes, namely '*missing haloes*'. Importantly, haloes within a small range of the peak height are spread over a wide range of masses as a result of the projection effect and the individuality of the halo mass distribution.

5.3 Effects of the noise on the lensing peaks

Before examining any correspondences between peaks identified in the noisy κ maps and haloes in the halo catalogue, we quantitatively examine how the noise alters the peaks caused by haloes and generates false peaks in the lensing map.

Matching between the noise-free peaks and the noisy peaks was done in the same manner described in Section 5.1. In what follows, the peaks classified in classes (i)–(iii) (their analogy to $\nu_{lens} - \nu_{noisy}$ correspondence) are considered. Again, we only deal with peaks with $\nu > 3$.

Fig. 18 shows the relationships between ν_{lens} and ν_{noisy} plotted in the same manner as in Fig. 14. We emphasize that the statistical properties of the $\nu_{lens} - \nu_{noisy}$ relationship do not follow a Gaussian distribution, even though Gaussian noise is added. The reason is as follows. The noise not only alters the peak height from massive haloes but also changes the peak position. For example, if a large positive noise is accidentally added on a neighbouring pixel around the original peak position, the pixel could be identified as a peak that is even higher than the original peak. Therefore, the distribution of the ν_{noisy} values is not simply expressed as the convolution of that of ν_{lens} values with Gaussian noise, but the $\nu_{lens} - \nu_{noisy}$ relationship becomes asymmetric and is biased toward large ν_{noisy} values. This explains the asymmetric shown in Fig. 18 $\nu_{lens} - \nu_{noisy}$ relationship clearly. The mean and rms of the differences, $\nu_{noisy} - \nu_{lens}$, among the class (i) objects with $\nu_{lens} > 4$ (this restriction is imposed to avoid an incomplete sampling of low ν_{noisy} pairs) are 0.31 and 0.90, respectively. Therefore the peak heights of massive haloes identified in a noisy κ map are, on average, boosted by the noise. This bias partly accounts for the excess in the peak number counts over the simple theoretical prediction (equation 16) as discussed in Section 4.2.

The class (ii) objects here correspond to the *missing peaks*, the peaks that have $\nu_{lens} > 3$ but are absent in the corresponding noisy κ map or have $\nu_{noisy} < 3$. With the exception of one peak with $\nu_{lens} = 5.1$, the missing peaks sit in the range of $\nu_{lens} \leqslant 4$. These peaks are likely to be erased or their heights decreased by the addition of negative noise. The one exception (the peak with $\nu_{lens} = 5.1$ at the missing peak locus) is a local peak associated with a substructure of a massive halo which disappears in the noisy κ map. We may, therefore, safely conclude that more than 80 per cent of the peaks with $\nu_{lens} \geqslant 3$ in the noise-free map remain in the noisy κ map.

Next, we consider the *false peaks*, the peaks that are identified in the noisy κ map with $\nu_{noisy} > 3$ but does not correspond to peaks in the noise-free κ map or have $\nu_{lens} < 3$, which are placed around $\nu_{lens} \sim 2$ in Fig. 18. A large fraction of the false peaks, especially for $\nu_{noisy} < 4$, are the false peaks generated by the noise (see Section 4 for details), and the small fraction (especially those with relatively high ν_{noisy}) is associated with real peaks on which a large noise is accidentally added. There are actually 14 objects with $\nu_{noisy} > 5$ in the *false peaks* locus. All of those have a corresponding peak in the noise-free κ map with low peak height ($1.5 < \nu_{lens} < 2.9$) on which a relatively large noise of $2 \sim 3.5\sigma$ is accidentally added. Therefore, all peaks with $\nu_{noisy} > 5$ identified in the noisy κ map are associated with real peaks in the noise-free κ map. Alternatively, most of the class (iii) peaks with lower heights are false peaks generated by the noise. In fact, one can find the rapid increasing number of the false peaks toward low peak heights, especially for $\nu_{noisy} < 4$, as shown in Section 4. It immediately follows from this that to reduce the contamination rate, a high threshold value, $\nu_{th} \gtrsim 4$ at least, is required.

5.4 Haloes versus peaks in noisy κ maps

In this subsection, we examine correspondences between haloes in the halo catalogue and peaks in the κ map including the noise caused

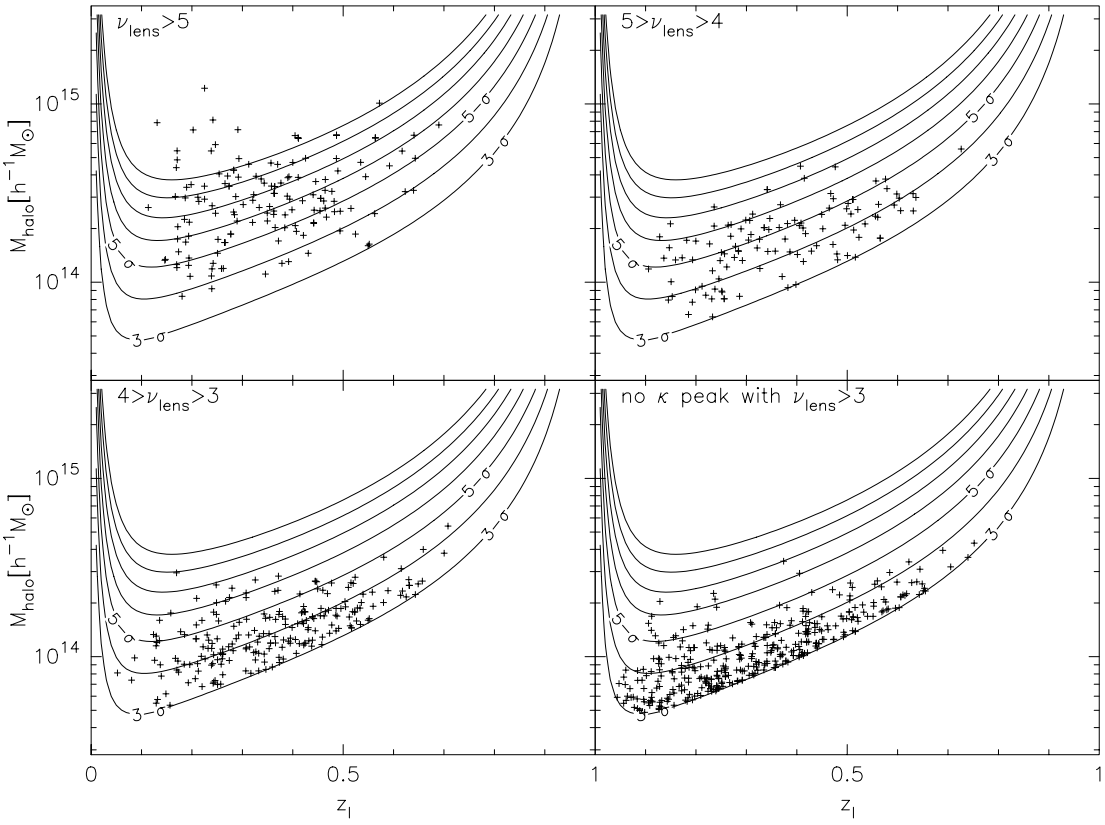

Figure 17. The scatter plot for the haloes in the redshift–halo mass plane. The contours show the *expected* peak height $\nu_{\rm NFW}$ from 3 to 9 at an interval of $\Delta\nu_{\rm NFW} = 1$. The class (i) haloes matched with peaks identified in *noise-free* κ maps are plotted in the upper-left ($\nu_{\rm lens} > 5$), upper-right ($5 > \nu_{\rm lens} > 4$) and lower-left panels ($4 > \nu_{\rm lens} > 3$). The lower-right panel shows the scatter plot for the class (ii) '*missing haloes*'.

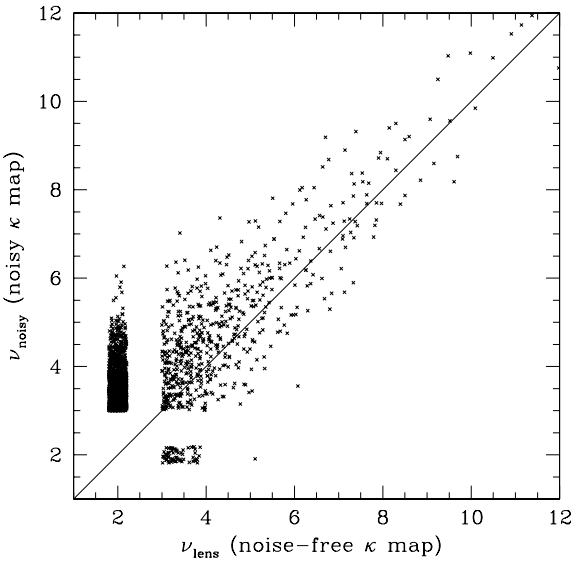

Figure 18. The correspondences between the peak heights identified in the noise-free and noisy κ maps. The peaks, which have $\nu_{\rm lens} > 3$ but are missed in the noisy κ map or have $\nu_{\rm noisy} < 3$, are placed within a narrow locus at $\nu_{\rm noisy} \sim 2$, regardless of their true $\nu_{\rm noisy}$ value. Similarly, the peaks, which have $\nu_{\rm noisy} > 3$ but no paired peak with $\nu_{\rm lens} > 3$, are placed at $\nu_{\rm lens} \sim 2$.

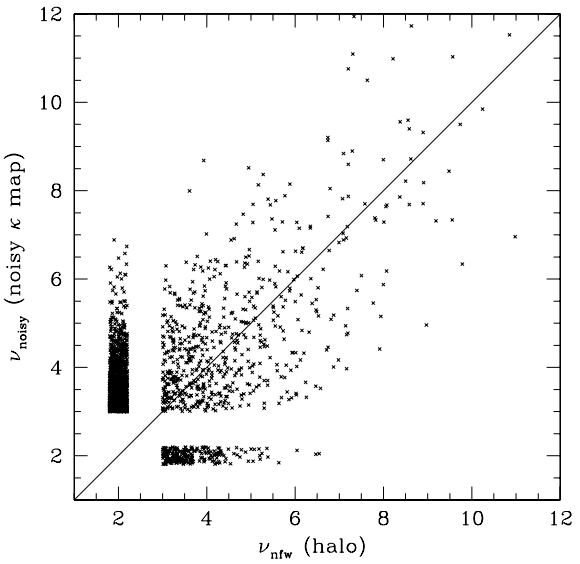

Figure 19. Same as for Fig. 14 but for the correspondence between haloes and peaks identified in the noisy κ maps.

by intrinsic ellipticities. Fig. 19 shows the distribution of the class (i–iii) objects in the $\nu_{\rm NFW} - \nu_{\rm noisy}$ plane plotted in the same manner as in Fig. 14. The comparison with Fig. 14 reveals that the noise not only increases the scatter in the $\nu_{\rm noisy} - \nu_{\rm NFW}$ relationship but also causes systematic bias in that the peak heights in the noisy map are statistically boosted, as seen in the previous subsection. Interestingly, this bias mean the *completeness* of the weak-lensing halo search can be improved. The upper-right-hand and lower-right-hand panels in Fig. 16 show the completeness and efficiency, defined in the same manner as equations (24) and (25), for the noisy maps.

About 81 (63) per cent of massive haloes with $\nu_{\rm NFW} > 5$ (>4) are identified as peaks with $\nu_{\rm noisy} \geqslant 4$ in the noisy maps. This is one of the most important results of this paper, indicating that a high completeness of the weak lensing halo search could be attained in a realistic data.

There are 13 *missing* haloes that have $\nu_{\rm NFW} > 5$ but $\nu_{\rm noisy} < 3$ or have no corresponding peak in the noisy κ map. We look into each case in the κ map and find that 11 haloes out of 13 are irregular systems such as unrelaxed systems with substructures or ongoing merger. Therefore, the *completeness* of the weak lensing halo search for regular, well-relaxed systems would be higher than the estimate plotted in Fig. 16. The remaining two *missing* haloes, which have $\nu_{\rm NFW} = 5.3$ and 5.4, are haloes which accidentally meet a large negative noise so that their κ values are below $\nu_{\rm noisy} = 3$. The fraction of such haloes over all haloes with $\nu_{\rm NFW} > 5$ is 0.8 per cent (two out of 250 haloes), which is very small indeed. Hence, we can safely conclude that almost all haloes with $\nu_{\rm NFW} > 5$ generate peaks with $\nu_{\rm noisy} > 3$ even in the presence of noise.

Let us turn to the class (iii) objects. As the noise generates false peaks in the κ map, the *efficiency* deteriorates and becomes worse than in the noise-free case (see Fig. 16). The fraction of contamination by false peaks is larger with a lower peak height because the noise generates relatively lower peaks, mostly $\nu_{\rm noisy} < 4$. Therefore, if $\nu_{\rm th} = 3$ is employed as the detection threshold, only 28 per cent of all peaks are associated with haloes with $\nu_{\rm NFW} > 3$. Note that as we are only considering haloes with $\nu_{\rm NFW} > 3$, some fraction of *missing peaks* could be associated with haloes with $\nu_{\rm NFW} < 3$. Alternatively, the efficiencies for higher detection thresholds are reasonably high: for $\nu_{\rm th} = 5$ ($\nu_{\rm th} = 4$), 85 per cent (56 per cent) of all peaks with $\nu_{\rm noisy} > \nu_{\rm th}$ are associated with haloes with $\nu_{\rm NFW} > 3$.

From Fig. 16, it is clear that the improved *efficiency* achieved by setting a higher detection threshold is a 'trade-off' for poorer *completeness*. We also note that the peak counts vary sensitively with the threshold. For example, the counts with $N(\nu > 4)$ are about twice as large as $N(\nu > 5)$ (see Section 4). Hence, the detection threshold should be determined so that the *efficiency* and *completeness* are optimized for specific survey strategies. The results we have so far shown indicate that $\nu_{\rm noisy} \simeq 4$–5 is reasonably optimal for current typical weak lensing surveys. To be more specific, the results from the simulated κ maps in the ΛCDM model indicate that the mean number of haloes with $\nu_{\rm NFW} > 4$ is $N_{\rm halo}(\nu_{\rm NFW} > 4) = 37$ per 10 deg^2, and on average 63 per cent of such haloes (23 out of 37) are detected with $\nu_{\rm noisy} > 4$. The completeness is better for higher $\nu_{\rm NFW}$ haloes: for haloes with $\nu_{\rm NFW} > 5$ the mean number is $N_{\rm halo}(\nu_{\rm NFW} > 5) = 19$ per 10 deg^2, where 81 per cent of the haloes (15 out of 19) are detected with $\nu_{\rm noisy} > 4$, and 12 haloes have a high-lensing signal of $\nu_{\rm noisy} > 5$. With respects to efficiency, if one sets the detection threshold at $\nu_{\rm th} = 4$, the mean number of peaks is $N_{\rm peak}(\nu_{\rm noisy} > 4) = 62$ per 10 deg^2. Among the 62 peaks, 23 are accountable to real haloes with $\nu_{\rm NFW} > 4$, 13 peaks are haloes with $3 < \nu_{\rm NFW} < 4$ and the remaining 26 peaks are likely to be either haloes with $\nu_{\rm NFW} < 3$ or false peaks from the noise. Alternatively, if one adopts a higher detection threshold of $\nu_{\rm th} = 5$, the mean number of peaks is reduced to $N_{\rm peak}(\nu_{\rm noisy} > 5) = 25$ per 10 deg^2, However, completeness is improved significantly: in fact it is found that 19 (3) peaks out of 25 are signals from real haloes with $\nu_{\rm NFW} > 4$ ($3 < \nu_{\rm NFW} < 4$).

Fig. 20 shows the scatter plot for the class (i) and class (ii) haloes in the redshift–halo mass plane as in Fig. 17. From comparison with Fig. 17, one finds that more haloes pass a certain detection threshold than in the noise-free map, owing to the systematic bias caused by

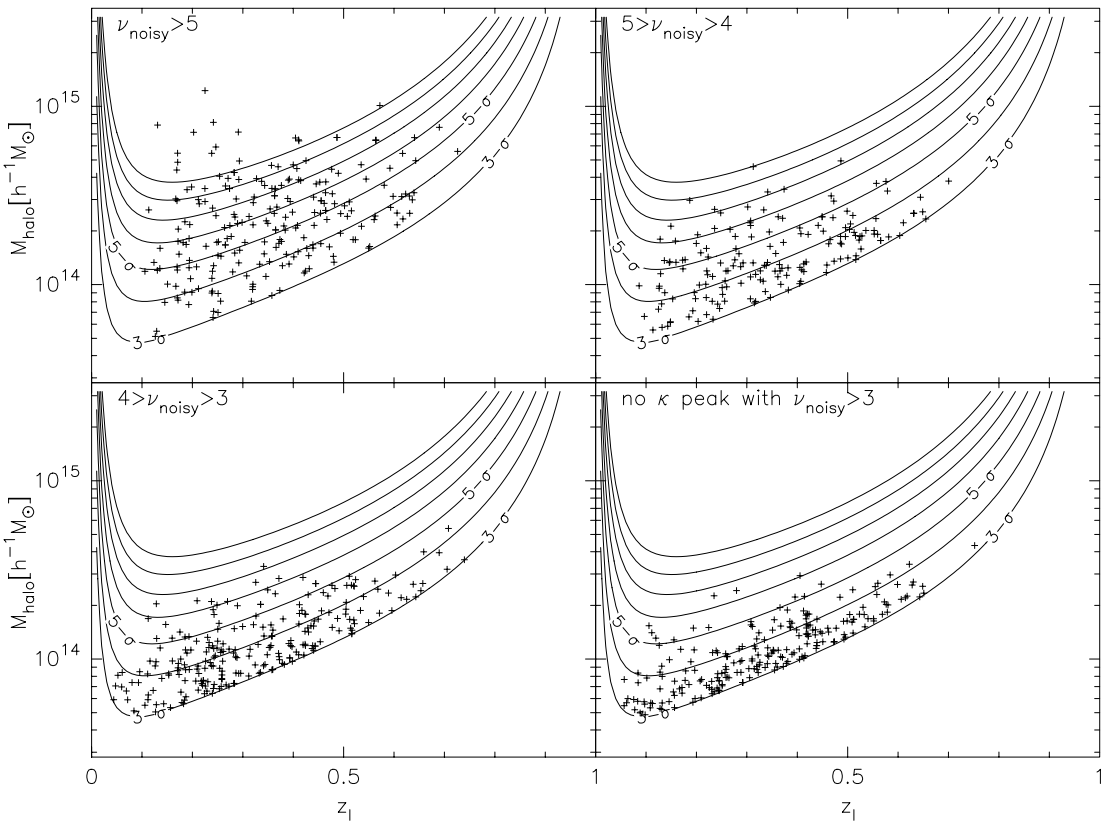

Figure 20. Same as for Fig. 17 but for haloes matched (not matched, for lower-right-hand panel) with peaks identified in *noisy* κ maps.

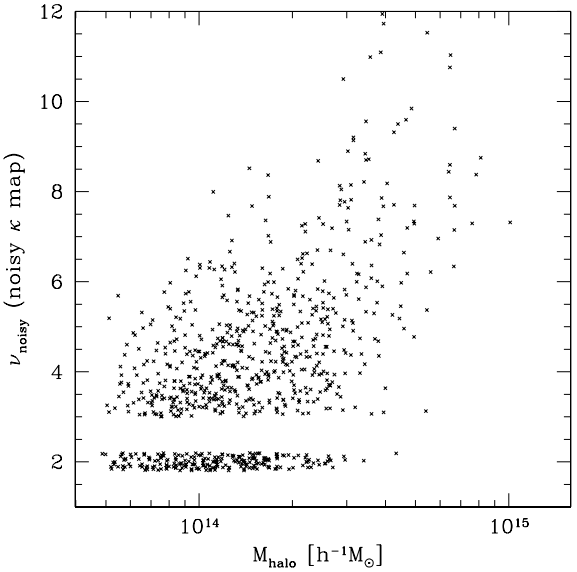

Figure 21. Same as for Fig. 19 but the halo mass is used instead of the *expected* peak height of $\nu_{\rm NFW}$. Note that only haloes with $\nu_{\rm NFW} > 3$ are considered.

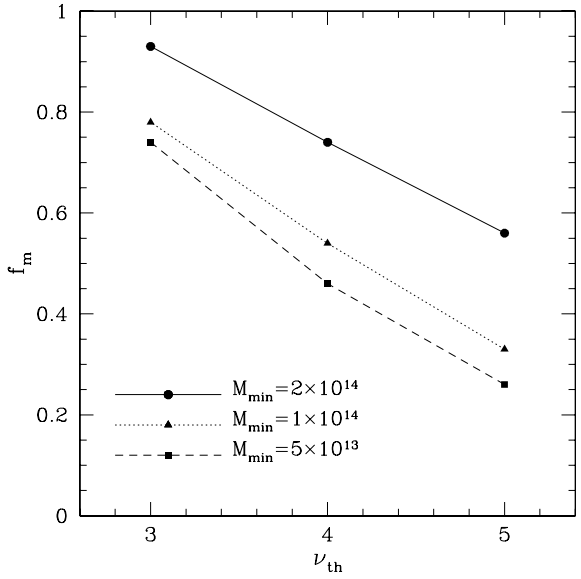

Figure 22. The number fraction of haloes identified by peaks with heights higher than $\nu_{\rm th}$ relative to all the haloes with mass greater than a given minimum mass, $M_{\rm min}$, $f_{\rm m}(\nu_{\rm noisy} > \nu_{\rm th}|M_{\rm halo} > M_{\rm min})$ (see equation 26). Three cases of $M_{\rm min} = 5 \times 10^{13}$, 1×10^{14} and $2 \times 10^{14}\ h^{-1}\ {\rm M}_\odot$ are plotted. Note that only haloes with $\nu_{\rm NFW} > 3$ are considered.

the noise. Another important point revealed in this figure is that the peak height in the noisy κ map is a poor estimator of the halo mass. For example, masses of the haloes detected with $4 < \nu_{\rm noisy} < 5$ span over an order of magnitude. Even in a narrow redshift range such as $0.3 < z < 0.4$, there still exits a significant scatter in the halo mass. Note that we do not plot haloes with $\nu_{\rm NFW} < 3$, as this could cause a larger scatter toward small mass ranges. It should also be noted that in the distribution of the missing haloes (lower-right-hand panel) no strong tendency with redshift is seen, indicating that completeness would not vary significantly with redshift. This is largely a result of our definition of completeness, that is, we define completeness so that the effect of the selection function, which very strongly depends on redshift, is decoupled.

So far we have examined the relationship between haloes and peaks in a κ map in terms of peak height as opposed to halo mass because the observable lensing is determined by a combination of halo mass, redshift and noise level, rather than by halo mass alone. Nevertheless, it is interesting to see what mass range of haloes produce the peaks in the noisy κ maps, which is shown in Fig. 21 as in Fig. 19. Again only haloes with $\nu_{\rm NFW} > 3$ are considered. As can be seen, the peaks of a given height are caused by haloes with a wide range of masses. We define the number fraction in a similar way to how completeness is defined by equation (24) but replacing the minimum peak height with the minimum halo mass as follows:

$$f_{\rm m}(\nu_{\rm noisy} > \nu_{\rm th} \,|\, M_{\rm halo} > M_{\rm min}) = \frac{N_{\rm i}(\nu_{\rm noisy} > \nu_{\rm th} \,|\, M_{\rm halo} > M_{\rm min})}{N_{\rm i+ii}(M_{\rm halo} > M_{\rm min})}, \tag{26}$$

where $N_{\rm i+ii}(M_{\rm halo} > M_{\rm min})$ is the number of [classes (i) and (ii)] haloes with mass greater than $M_{\rm min}$, and $N_{\rm i}(\nu_{\rm noisy} > \nu_{\rm th}|M_{\rm halo} > M_{\rm min})$ is the number of haloes with mass greater than $M_{\rm min}$ having peak height greater than $\nu_{\rm th}$. Fig. 22 shows this estimate as a function of $\nu_{\rm th}$. It is apparent that $f_{\rm m}$ declines toward a higher detection threshold and for lower minimum mass. For a detection threshold of $\nu_{\rm th} = 4$, more than 70 per cent of haloes with $M_{\rm halo} > 2 \times 10^{14}\ h^{-1}\ {\rm M}_\odot$ are identified, while about half of haloes with $M_{\rm halo} > 5 \times 10^{13}\ h^{-1}\ {\rm M}_\odot$ are identified

6 SUMMARY AND DISCUSSION

We have investigated various aspects of weak lensing cluster surveys employing both analytic descriptions of dark matter haloes and the mock data of weak lensing surveys generated from numerical simulations. For the latter, we combined weak lensing ray tracing through mass distribution and dark matter halo catalogues. Our major findings are summarized as follows.

(1) In Section 2, we examined the expected properties of weak lensing haloes using the analytic descriptions of the dark matter haloes, including the universal density profile (Navarro et al. 1996; Navarro et al. 1997) and the Press–Schechter halo mass function (Press & Schechter 1974; Sheth & Tormen 1999). We found that Gaussian smoothing with $\theta_{\rm G} \simeq 1$ arcmin gives the largest expected weak lensing halo counts (which may, however, depend on the source redshift and noise properties). We computed the selection function of the weak lensing cluster survey and examined, in detail, the mass and redshift distribution of the weak lensing haloes. It was shown that the detectability of haloes depends not only on the mass but also strongly on the redshift.

(2) In Section 3.4, we compared the model prediction of the weak lensing halo counts developed by Kruse & Schneider (2000) and Bartelmann et al. (2001) with the halo counts in the mock catalogue, and found a good agreement. We also found a large scatter in the numbers of weak lensing haloes within a field of 4 deg^2, which is larger than the Poisson fluctuation. This can be explained by the strong clustering of massive haloes.

(3) In Section 4, we tested the predictions of the model of the peak counts against the mock weak lensing data. It was found that a systematic bias is induced by the noise caused by intrinsic galaxy ellipticities. This bias increases the peak counts, and hence the prediction using the model underestimates the counts. We developed a correction scheme in an empirical manner in Appendix B, and showed that the improved model reproduces the peak counts reasonably well.

(4) In Section 5.2, using the mock weak lensing data combined with the halo catalogue, we examined the matching between haloes and peaks identified in the noise-free κ map. This was done to clarify the influence of halo individuality and the projection effect on a weak lensing halo search. We showed that these effects cause not only a large scatter in the $\nu_{\rm lens} - \nu_{\rm NFW}$ relationship but also a systematic bias. The mean and rms of the differences $\nu_{\rm lens} - \nu_{\rm NFW}$ are -0.24 and 1.2, respectively. The level of scatter caused by the effects is comparable to that from the noise. We argue that the negative mean values are caused by a population of unrelaxed, less centrally concentrated haloes and the fact that more than half of the haloes are elongated in the direction perpendicular to the line of sight. In addition, the projection of large under-dense regions may partly account for the scatter. It was also shown that the chance projection of massive haloes in the same line of sight is very rare.

(5) In Section 5.3, we used the noise-free κ map and the noisy κ map to clarify how the peak distribution is affected by noise and how false peaks are generated. We showed that the $\nu_{\rm lens} - \nu_{\rm noisy}$ distribution is biased toward large $\nu_{\rm noisy}$ values. The mean and rms of the differences, $\nu_{\rm noisy} - \nu_{\rm lens}$, among the matched peaks with $\nu_{\rm lens} > 4$ and $\nu_{\rm noisy} > 4$ are 0.31 and 0.90, respectively. Thus the noise not only generates the scatter but also systematically 'boosts' the peak heights. We found that almost all the peaks identified in the noise-free κ map with $\nu_{\rm lens} > 4$ are identified in the noisy κ map, thus the noise rarely erases the high peaks. All the peaks with $\nu_{\rm noisy} > 5$ identified in the noisy κ map are associated with real peaks in the noise-free κ map, whereas most of the false peaks (resulting from the noise) have a relatively lower peak height, $\nu_{\rm noisy} < 4$.

(6) In Section 5.4, we examined the correspondence between the haloes and peaks identified in the noisy κ map. In particular, we studied the *efficiency* and *completeness* of the weak lensing halo survey. We found that the detection threshold of $\nu_{\rm th} \simeq 4$–5 gives an optimal balance between efficiency and completeness. It was shown that about 81 (62) per cent of massive haloes with $\nu_{\rm NFW} > 5$ (>4) are identified as high peaks with $\nu_{\rm th} = 4$. This suggests that the completeness of the weak lensing cluster survey is reasonably high. Concerning the efficiency, we found that for $\nu_{\rm th} = 4$ (5), 56 per cent (82 per cent) of all the peaks are real signals from haloes with $\nu_{\rm NFW} > 3$, while 37 per cent (63 per cent) are signals from haloes with a higher expected peak height of $\nu_{\rm NFW} > 4$. Therefore, it is possible to attain relatively high levels of efficiency *and* completeness by selecting a moderately high detection threshold.

(7) In Section 5.4, we conclude that, for a detection threshold of $\nu_{\rm th} = 4$, about half of the haloes with $M_{\rm halo} > 5 \times 10^{13}\, h^{-1}\, {\rm M}_\odot$ and more than 70 per cent of haloes with $M_{\rm halo} > 2 \times 10^{14}\, h^{-1}\, {\rm M}_\odot$ indeed produce the lensing signals with $\nu_{\rm noisy} \geqslant 4$.

The weak lensing cluster search technique explored in this paper can be directly applied to data obtained by ongoing/future wide field surveys. Extending the pilot 2.1-deg^2 survey (Miyazaki et al. 2002), a wide field weak lensing survey (Suprime33, PI: S. Miyazaki) is being carried out using the wide field prime focus camera on the Subaru telescope, Suprime-Cam (Miyazaki et al., private communication). Suprime33 will cover 33 deg^2 with a limiting magnitude of $R = 25.5$ (which provides $n_{\rm g} \gtrsim 30$ and $\langle z_{\rm s} \rangle \simeq 1$). For the survey exposure time of 1800 s and the field-of-view of the Suprime-Cam, 0.25 deg^2, the effective survey cost is 2.5 h per 1 deg^2 including overheads. The expected number of clusters that should be located by Suprime33 is 70 (120) at a threshold of $\nu_{\rm th} = 5$ ($\nu_{\rm th} = 4$). It can thus be seen that weak lensing surveys exploiting a wide field camera on a large telescope offer a reasonably efficient way of locating massive clusters. The CFHT Legacy Survey[4] will significantly enlarge the survey area. Its 'wide' survey will observe 170 deg^2 in total with a limiting magnitude of $i' = 25.5$ (providing $n_{\rm g} \sim 30$ and $\langle z_{\rm s} \rangle \simeq 1$). The expected number of clusters is 360 (600) for $\nu_{\rm th} = 5$ ($\nu_{\rm th} = 4$), and thus it will provide invaluable data for studies on large-scale structures of the Universe. Although herein we have primarily considered weak lensing surveys that are feasible with current ground-based telescopes, future wide field surveys using a space telescope will enable to us accurately measure the shape of distant, very small galaxy images, which will significantly improve the capabilities of weak lensing cluster surveys and allow the detection of lower mass and/or higher redshift clusters.

Finally, we note that it is, in principle, possible to enhance the detection S/N by choosing a suitable smoothing function (White et al. 2002; Padmanabhan et al. 2003). Padmanabhan et al. (2003) proposed the function of $W(\theta) = (1 + \theta/\theta_{\rm c})^{-2}$ with $\theta_{\rm c} \sim 1$ arcmin, which is motivated by the asymptotic behaviour of the outer part of the projected NFW profile. Such optimized filters may improve the S/N if the matter distribution is indeed close to the NFW profile. Alternatively, its applicability to clusters whose mass distribution deviates from the NFW profile remains unclear. It needs to be explored how much improvement can be obtained both in efficiency and completeness. Optimizing the smoothing function clearly warrants further studies.

ACKNOWLEDGMENTS

We are grateful to Bhuvnesh Jain for many valuable discussions. We would like to thank Satoshi Miyazaki for useful discussions. We thank the anonymous referee for detailed and constructive comments on the earlier manuscript which improved various aspects of the paper. TH and NY thank University of Pennsylvania for the warm hospitality during their visit where this work was initiated. TH and NY acknowledge support from the Japan Society for Promotion of Science (JSPS) Research Fellowships. The *N*-body simulations used in this work were carried out by the Virgo Consortium at the computer centre at the Max-Planck-Institut, Garching (http://www.mpa-garching.mpg.de/NumCos). The numerical computations presented in this paper were partly carried out at ADAC (the Astronomical Data Analysis Centre) of the National Astronomical Observatory, Japan.

REFERENCES

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Bartelmann M., 1996, A&A, 313, 443
Bartelmann M., Schneider P., 1992, A&A, 259, 413
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Bartelmann M., King L. J., Schneider P., 2001, A&A, 378, 361
Bond J. R., Efstathiou G., 1987, MNRAS, 226, 655
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
Hamana T., Mellier Y., 2001, MNRAS, 169, 176
Hamana T., Yoshida N., Suto Y., 2002, ApJ, 568, 455
Henry J. P., 2000, ApJ, 534, 565
Hockney R. W., Eastwood J. W., 1988, Computer Simulation Using Particles. Adam Hilger, Bristol
Hoekstra H., 2001, A&A, 370, 743
Jain B., Van Waerbeke L., 2000, ApJ, 530, L1
Jain B., Seljak U., White S. D. M., 2000, ApJ, 530, 547

[4] http://www.cfht.hawaii.edu/Science/CFHLS/

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
Jing Y. P., 2000, ApJ, 535, 30
Kaiser N., Squires G., 1993, ApJ, 404, 441
Kruse G., Schneider P., 2000, MNRAS, 318, 321
Macfarland T., Couchman H. M. P., Pearce F. R., Pichlmeier J., 1998, New Astron., 3, 687
Mellier Y., 1999, ARA&A, 37, 127
Metzler C. A., White M., Loken C., 2001, ApJ, 547, 560
Miyazaki S. et al., 2002, ApJ, 580, L97
Nakamura T. T., Suto Y., 1997, Prog. Theor. Phys., 97, 49
Navarro J., Frenk C., White S. D. M., 1996, ApJ, 462, 563
Navarro J., Frenk C., White S. D. M., 1997, ApJ, 490, 493
Padmanabhan N., Seljak U., Pen U. L., 2003, New Astron., 8, 581
Reblinsky K., Bartelmann M., 1999, A&A, 345, 1
Press W. H., Schechter P., 1974, ApJ, 187, 425
Schneider P., Ehlers J., Falco E. E., 1992, Gravitational Lenses. Springer Verlag, Heidelberg
Seljak U., Zaldarriaga M., 1996, ApJ, 469, 437
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Takada M., Jain B., 2003, MNRAS, 340, 580
Vale C., White M., 2003, ApJ, 592, 699
Van Waerbeke L., 2000, MNRAS, 313, 524
White M., Van Waerbeke L., Mackey J., 2002, ApJ, 575, 640
Wittman D., Tyson J. A., Margoniner V. E., Cohen J. G., Dell'Antonio I. P., 2001, ApJ, 557, L89
Wright C. O., Brainerd T. G., 2000, ApJ, 534, 34
Yoshida N., Sheth R., Diaferio A., 2001, MNRAS, 328, 669

APPENDIX A: THE VISUAL APPEARANCE OF SYSTEMS WITH IRREGULAR SIGNALS

Here we present some κ maps, in which *missing haloes* (class ii) or *false peaks* (class iii) with high signals ($\nu > 5$) exists, for examples of irregular systems. Figs A1–A5 show the (noise-free) convergence maps smoothed with the Gaussian window function with $\theta_G = 1$ arcmin. The plus symbols denote the centre-of-mass position of the haloes concerned and the circles show their virial radius. The crosses denote the positions of the peaks.

Figs A1–A3 show three examples of *missing haloes* with the expected peak value of $\nu_{NFW} > 5$. Such high ν_{NFW} value *missing haloes* mostly have an unrelaxed appearance and are probably ongoing merger systems. Figs A4 and A5 show two examples of *false peaks* with $\nu_{lens} > 5$. Note that *false peaks* are rare in noise-free κ maps. Most *false peaks* with $\nu_{lens} > 5$ are associated with substructures of massive haloes. The occurrence of chance projections is very rare; we found three cases with $\nu_{lens} > 5$ in our 30 maps.

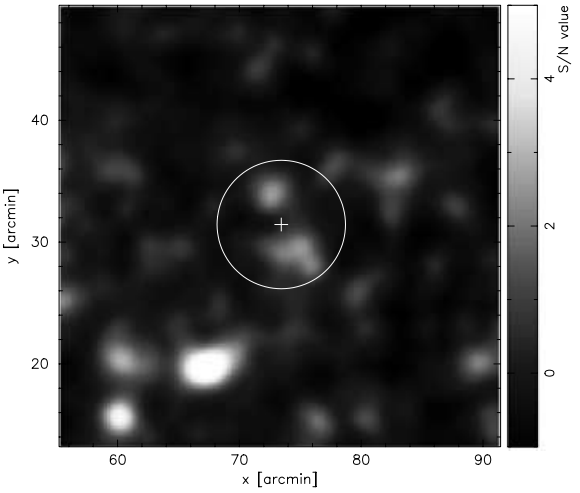

Figure A1. *Missing halo*, an example of a halo with an extended matter distribution (possibly in the process of merging). The plus symbol denotes the centre-of-mass position of the halo, and the circle shows the virial radius. The map apparently shows a very eccentric density distribution. The halo has $M_{halo} = 3.4 \times 10^{14}\ h^{-1}\ M_\odot$ at $z = 0.37$, and the shape parameters are $Q/Q_{NFW} = 2.1$, $R = 1.4$ and $c/a = 0.23$. The expected peak value is $\nu_{NFW} = 7.0$, but the ν value measured in the lensing map is less than 3.

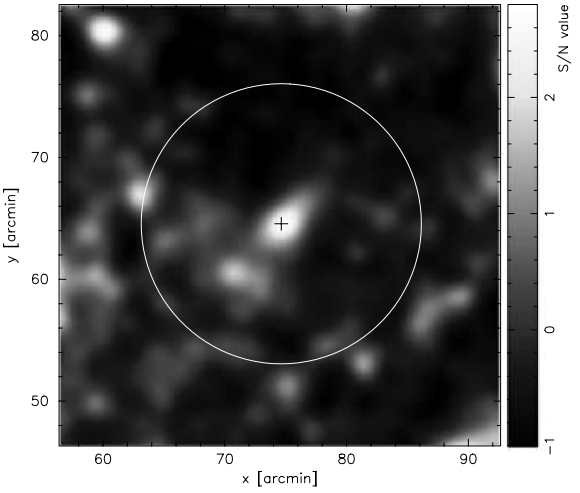

Figure A2. *Missing halo*, an example of low-z unrelaxed system. The halo ($M_{halo} = 2.0 \times 10^{14}\ h^{-1}\ M_\odot$, $z = 0.13$ and the shape parameters are $Q/Q_{NFW} = 1.2$, $R = 0.36$ and $c/a = 0.41$) does not seem to have reached a relaxed stage, but is probably an ongoing merger system. The expected peak value is $\nu_{NFW} = 6.6$ but the actual ν value is less than 3.

APPENDIX B: A SIMPLE CORRECTION SCHEME TO THE PEAK COUNTS

As shown in Section 4 the theoretical model of the peak counts given by equation (16) underestimates the counts obtained from the realistic noisy κ map by about 20 per cent. We here discuss the reasons for this change, and develop a correction scheme to the peak counts in an empirical manner. Here we focus on the high peaks of $\nu > 5$.

The sources that modify the peak counts are as follows.

(1) The biased relationship between ν_{NFW} and ν_{lens}. In Section 5.2, it is found that the mean of the differences $\nu_{NFW} - \nu_{lens}$ over class (i) objects is -0.24. This means that the peak height computed assuming the NFW profile tends to overestimate the real peak height. The reason for this is discussed in Section 5.2.

(2) Spurious peaks generated by the noise. It is found in Section 5 that spurious peaks yield about a 10 *per cent* fraction of the total number of high peaks ($\nu > 5$).

(3) The scatter in the relationship between ν_{noisy} and ν_{lens} resulting from the noise. As will be shown, this scatter increases the peak counts.

(4) The biased relationship between ν_{noisy} and ν_{lens}. In Section 5.3, it is found that the noise produces not only the scatter but also the systematic bias, and that the mean of the differences $\nu_{noisy} - \nu_{lens}$ over class (i) objects is 0.31. The peak heights in the noisy κ map are thus statistically boosted by the noise. The reason for this is discussed in Section 5.3.

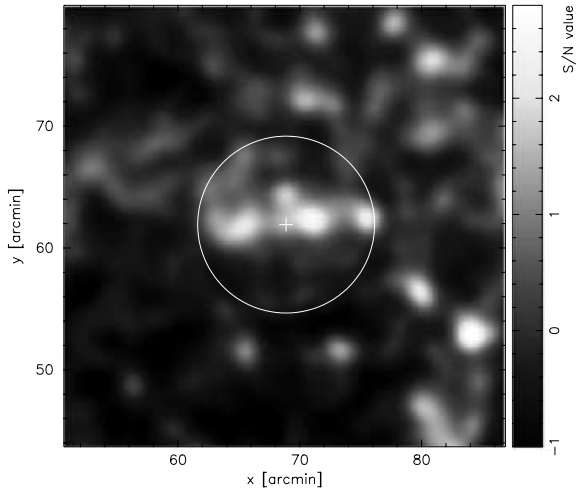

Figure A3. *Missing halo*, an example of an irregular system. The map apparently shows a very eccentric density distribution. It is likely that the halo ($M_{\rm halo} = 1.9 \times 10^{14}\ h^{-1}\ {\rm M}_\odot$, $z = 0.20$ and the shape parameters are $Q/Q_{\rm NFW} = 1.3$, $R = 0.25$ and $c/a = 0.36$) is an ongoing merger system. The expected peak value is $\nu_{\rm NFW} = 6.1$ but the actual ν value is less than 3.

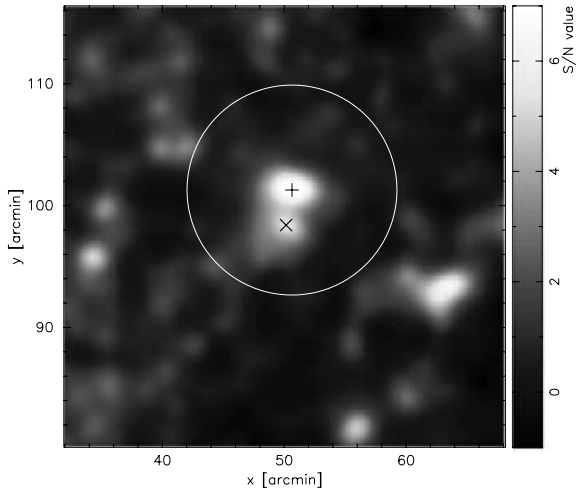

Figure A4. An example of a *local peak* associated with the substructure of a massive halo. The cross denotes the position of the peak with $\nu_{\rm lens} = 5.3$. This is the local peak associated with the substructure of the massive halo ($M_{\rm halo} = 3.5 \times 10^{14}\ h^{-1}\ {\rm M}_\odot$, $z = 0.20$, $\nu_{\rm NFW} = 8.8$ and the shape parameters are $Q/Q_{\rm NFW} = 0.94$, $R = 0.61$ and $c/a = 0.66$). The plus symbol denotes the centre-of-mass position of the halo and the circle shows the virial radius.

First, we develop the correction scheme for point (3). To do this, we use a simplified method, employing the assumption that very high peaks are neither removed nor generated by the noise but their peak heights are altered by its addition. In other words, we ignore, for the moment, any spurious peaks resulting from the noise and other sources. Under this above assumption, the peak counts in the noisy κ map are given by combining the noise-free PDF (equation 16) with the Gaussian noise PDF as follows:

$$n_{\rm noisy}(\nu) = \int {\rm d}\nu' \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\nu'^2}{2}\right) n_{\rm peak}(\nu - \nu'). \tag{B1}$$

We further assume that within a small range of ν values the peak counts in the original lensing map can be approximated by the exponential form $n_{\rm peak}(\nu) = n_* \exp(p\nu)$, with a constant exponential index of p. As can be seen in Fig. B1, this would be a reasonable approximation over the range of $3 \lesssim \nu \lesssim 8$. This approximation allows us to compute the integration in equation (B1) analytically, i.e.

$$\begin{aligned} n_{\rm noisy}(\nu) &\simeq \int {\rm d}\nu' \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\nu'^2}{2}\right) n_* \exp\left(p(\nu - \nu')\right) \\ &= \exp\left(\frac{p^2}{2}\right) n_{\rm peak}(\nu). \end{aligned} \tag{B2}$$

Thus the scatter resulting from the noise increases the counts by a factor of $\exp(p^2/2)$.

Turn now to points (1) and (4) in the above list. From the biases found in the mock data in the $\nu_{\rm NFW} - \nu_{\rm lens}$ and $\nu_{\rm lens} - \nu_{\rm noisy}$ relationships, we derive an approximate mean relationship of $\nu_{\rm noisy} - \nu_{\rm NFW} = 0.07$. We take this bias by simply shifting the counts as $n_{\rm noisy}(\nu) \to n_{\rm noisy}(\nu + 0.07)$.

Finally, point (2) is corrected by simply increasing the peak counts by 10 per cent.

In Fig. B1, the corrected model prediction is plotted and compared with the results from numerical experiments. To compute the corrected model prediction, $n_{\rm noisy}(\nu)$, we first computed the uncorrected counts, $n_{\rm peak}(\nu)$, at a small interval of $\delta\nu$ using equation (16). The local exponential slope of the counts is then evaluated by a finite differential scheme, $p(\nu_i) = [\log n(\nu_{i+1}) - \log n(\nu_i)]/\delta\nu$. We compute $n_{\rm noisy}(\nu)$ using equation (B2) with $p(\nu_i)$. Then, the correction $n_{\rm noisy}(\nu) \to n_{\rm noisy}(\nu + 0.07)$ was made for $p(\nu_i)$ at every ν_i to correct points (1) and (4). Finally the counts are increased by 10 per cent to correct point (2). The corrected prediction agrees with the noisy counts reasonably well for at least $\nu > 5$, where the above approximations are valid. On lower peak heights, i.e. $\nu < 4$, the prediction is systematically smaller than the measured values, where false peaks resulting from noise make a larger contribution. Note that the amplitude of effects (2) and (4) may depend on the noise properties and on the choice of the smoothing scheme.

It is clearly seen in Fig. B1 that the corrected peak counts has a slope that is very similar to that of the uncorrected counts. One

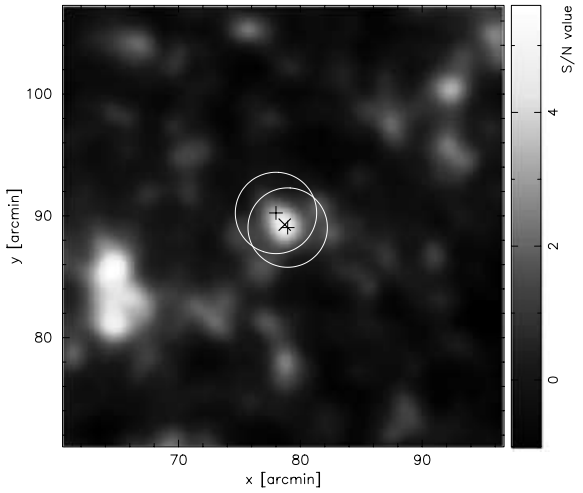

Figure A5. A very rare example of a *chance projection* of two moderately massive haloes. The cross denotes the position of the peak with $\nu_{\rm lens} = 5.1$. This peak is not generated by a single halo but is generated by the chance projection of two moderately massive haloes ($M_{\rm halo} = 7.5 \times 10^{13}\ h^{-1}\ {\rm M}_\odot$, $z = 0.35$ and $M_{\rm halo} = 5.4 \times 10^{13}\ h^{-1}\ {\rm M}_\odot$, $z = 0.68$). Neither of the two haloes on its own can generate a lensing signal that is larger than $\nu > 3$.

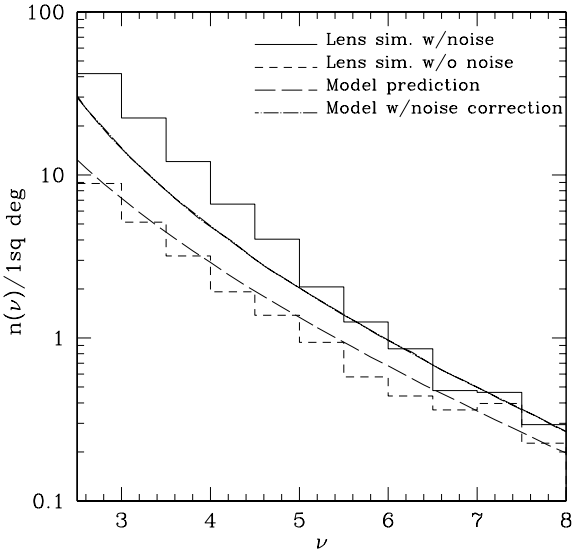

Figure B1. The number count of the lensing peaks per 1 deg^2. The histograms are from the numerical experiments from the noisy κ maps (solid) and the noise-free κ maps (dashed). The long-dashed line shows the uncorrected model prediction computed from equation (16), while the solid line shows the corrected model prediction (see Appendix B and Section 2.2 for details).

may thus make a simple correction to the model prediction obtained from equation (16) by multiplying by a factor of 1.2, which gives an alternative approximate way of computing the corrected peak counts with a reasonable level of accuracy.

APPENDIX C: RELATIONSHIP BETWEEN THE HALO SHAPE AND κ AMPLITUDE

Here we examine the effect of individuality in the halo mass distribution on the amplitude of the κ peak. Figs C1 and C2 show how the ratio $\nu_{\rm lens}/\nu_{\rm NFW}$ varies with the halo shape parameters Q and R (see equations 20 and 23 for definitions), where we only consider class (i) haloes with $\nu_{\rm NFW} > 4$ and $\nu_{\rm lens} > 4$. The filled circles show the average over haloes sitting in the range denoted by the x-direction bar, where each interval is defined so that it contains 35 haloes, and the error bar in the y-direction denotes the rms among them.

The correlations between the shape parameters and amplitude of the lensing signal are observed in both the figures, although the statistical significance is not very high. The anti-correlation in $\nu_{\rm lens} - Q$ is because a higher concentration of mass distribution ($Q < 1$) compared to the average NFW profile gives a greater lensing signal. The lensing signal is also amplified if the mass distribution within a halo is elongated along the line of sight ($R > 1$). However, we also draw attention to the rather large scatters in these relationships that may arise from the following two sources. First, the shape parameters we employ characterize a rather global mass distribution within a halo and cannot describe detailed structures. Second (and probably more important) is the projection effect resulting from different structures along the same line of sight of the halo.

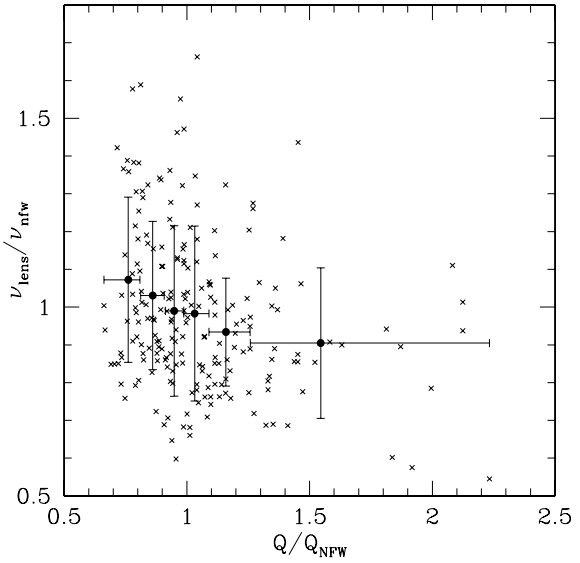

Figure C1. Shown is the scatter plot of the distribution of the haloes in the $Q/Q_{\rm NFW} - \nu_{\rm lens}/\nu_{\rm NFW}$ plane. Just class (i) haloes with $\nu_{\rm NFW} > 4$ and $\nu_{\rm lens} > 4$ are plotted. The filled circles show the average value over 35 haloes within bins whose ranges are denoted by the x-error bars, while the y-error bars denote rms among the 35 haloes.

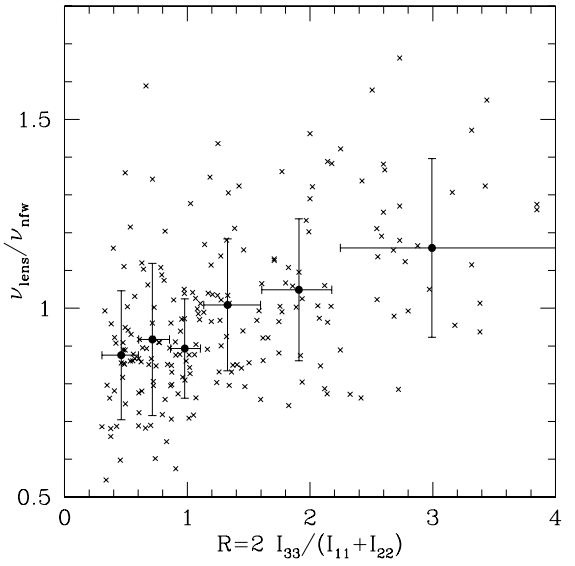

Figure C2. Same as for Fig. C1 but for the $R - \nu_{\rm lens}/\nu_{\rm NFW}$ plane.

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