An analytic approach to number counts of weak-lensing peak detections

Matteo Maturi*, Christian Angrick, Matthias Bartelmann, and Francesco Pace

Zentrum für Astronomie der Universität Heidelberg, Institut für Theoretische Astrophysik, Albert-Ueberle-Str. 2, 69120 Heidelberg, Germany

Astronomy & Astrophysics, submitted

ABSTRACT

We develop and apply an analytic method to predict peak counts in weak-lensing surveys. It is based on the theory of Gaussian random fields and suitable to quantify the level of spurious detections caused by chance projections of large-scale structures as well as the shape and shot noise contributed by the background galaxies. We compare our method to peak counts obtained from numerical ray-tracing simulations and find good agreement at the expected level. The number of peak detections depends substantially on the shape and size of the filter applied to the gravitational shear field. Our main results are that weak-lensing peak counts are dominated by spurious detections up to signal-to-noise ratios of 3–5 and that most filters yield only a few detections per square degree above this level, while a filter optimised for suppressing large-scale structure noise returns up to an order of magnitude more.

Key words. Cosmology: theory – large-scale structure of Universe – Galaxies: clusters: general – Gravitational lensing

1. Introduction

Wide-area surveys for weak gravitational lensing can be and have been used for counting peaks in the shear signal, which are commonly interpreted as the signature of sufficiently massive dark-matter halos. However, such detections are clearly contaminated by spurious detections caused by the chance superposition of large-scale structures, and even by the shape- and shot-noise contributions from the background galaxies used to sample the foreground shear field. As a function of the peak height, what is the contribution of genuine halos to these detections, and how much do the large-scale structure and the other sources of noise contribute?

Given the power of lensing-peak number counts as a cosmological probe [Marian et al. 2009; Kratochvil et al. 2009; Dietrich & Hartlap 2009], we address this question here after developing a suitable analytic approach based on peak counts in Gaussian random fields, as laid out by Bardeen et al. (1986). It is reasonable to do so even though at least the high peaks are caused by halos in the non-Gaussian tail of the density fluctuations because the noise and large-scale structures contribute to the foreground shear field. As a function of the peak height, what is the contribution of genuine halos to these detections, and how much do the large-scale structure and the other sources of noise contribute?

We study three such filters here, with the optimal filter among them. Results will differ substantially, arguing for a careful filter choice if halo detections are the main goal of the application. We compare our analytic results to a numerical simulation and show that they agree at the expected level. We begin in § 2 with a brief summary of gravitational lensing as needed here and describe filtering methods in § 3. We present our analytic method in § 4 and compare it to numerical simulations in § 5, where we also show our main results. Conclusions are summarised in § 6. In Appendix A, we show predictions of peak counts and the noise levels in them for several planned and ongoing weak-lensing surveys.

2. Gravitational lensing

Isolated lenses are characterised by their lensing potential

\[ \psi(\theta) \equiv \frac{2}{c^2} \frac{D_{ds}}{D_d D_s} \int \Phi(D_d \theta, z) \, dz, \tag{1} \]

where \( \Phi \) is the Newtonian gravitational potential and \( D_{ds}, D_d, D_s \) are the angular-diameter distances between the observer and the source, the observer and the lens, and the lens and the source, respectively. The potential \( \psi \) relates the angular positions \( \beta \) of the source and \( \theta \) of its image on the observer’s sky through the lens equation

\[ \beta = \theta - \nabla \psi. \tag{2} \]

Since sources such as distant background galaxies are much smaller than the typical scale on which the lens properties change and the angles involved are small, it is possible to linearise Eq. (2) such that the induced image distortion is expressed by the Jacobian

\[ A = (1 - \kappa) \begin{pmatrix} 1 - g_1 & g_2 \\ g_2 & 1 + g_1 \end{pmatrix}, \tag{3} \]
where $\kappa = \nabla^2 \psi / 2$ is the convergence responsible for the isotropic magnification of an image relative to its source, and $g(\theta) = \gamma(\theta) / [1 - n(\theta)]$ is the reduced shear quantifying its distortion. Here, $\gamma_1 = (\psi_{11} - \psi_{22}) / 2$ and $\gamma_2 = \psi_{12}$ are the two components of the complex shear. Since the angular size of the source is unknown, only the reduced shear can be estimated starting from the observed ellipticity of the background sources,

$$
\epsilon = \frac{\epsilon_s + g}{1 + g^* \epsilon_s},
$$

(4)

where $\epsilon_s$ is the intrinsic ellipticity of the source and the asterisk denotes complex conjugation.

3. Measuring weak gravitational lensing

3.1. Weak lensing estimator

In absence of intrinsic alignments between background galaxies due to possible tidal interactions (Heavens & Peacock 1998; Schneider & Bridle 2009), the intrinsic source ellipticities in Eq. (3) average to zero in a sufficiently large source sample. An appropriate and convenient measure for the lensing signal is the weighted average over the tangential component of the shear $\gamma_t$ relative to the position $\theta$ on the sky,

$$
\bar{A}(\theta) = \int d^2 \theta' \gamma_t(\theta', \theta)Q(|\theta' - \theta|). 
$$

(5)

The filter function $Q$ determines the statistical properties of the estimator $\bar{A}$. We shall consider three filter functions here which will be described in §3.2.

Data on gravitational lensing by a mass concentration can be modeled by a signal $s(\theta) = A \tau(\theta)$ described by its amplitude $A$ and its radial profile $\tau$, and a noise component $n(\theta)$ with zero mean, e.g.

$$
\gamma_t(\theta) = A \tau(\theta) + n(\theta) 
$$

(6)

for the tangential shear. The variance of the estimator $\bar{A}$ in (5) is

$$
\sigma^2_{\bar{A}} = \int \frac{kdk}{2\pi} P(k) \hat{W}(k) \hat{Q}(k)^2,
$$

(7)

where $\hat{W}(k)$ is the frequency response of the survey depending on its geometrical properties, $\hat{Q}(k)$ is the Fourier transform of the filter $Q$, and $P(k)$ is the power spectrum of the noise component.

3.2. Weak lensing filters

Different filter profiles have been proposed in the literature depending on their specific application in weak-lensing. We adopt three of them here which have been used so far to identify halo candidates through weak lensing.

(1) The polynomial filter described by Schneider et al. (1998)

$$
Q_{\text{poly}}(x) = \frac{6x^2}{\pi r^2_a} \left(1 - x^2\right) H(1 - x),
$$

(8)

where the projected angular distance from the filter centre, $x = r / r_a$, is expressed in units of the filter scale radius, $r_a$, and $H$ is the Heaviside step function. This filter was originally proposed for cosmic-shear analysis but several authors have used it also for dark matter halo searches (see for e.g. Erben et al. 2000; Schirmer et al. 2004).

(2) A filter optimised for halos with NFW density profile, approximating their shear signal with a hyperbolic tangent (Schirmer et al. 2004),

$$
Q_{\text{tanh}}(x) = \left(1 + e^{a-bx} + e^{c-x-d}\right)^{-1} \tanh \left(\frac{x}{x_c}\right),
$$

(9)

where the two exponentials in parentheses are cut-offs imposed at small and large radii ($a = 6$, $b = 150$, $c = 50$, and $d = 47$) and $x_c$ is a parameter defining the filter-profile slope. A good choice for the latter is $x_c = 0.1$ as empirically shown by Hetterscheidt et al. (2005).

(3) The optimal linear filter introduced by Maturi et al. (2005), which, together with the optimisation with respect to the expected halo-lensing signal, optimally suppresses the contamination due to the line-of-sight projection of large-scale structures (LSS),

$$
\hat{Q}_{\text{opt}}(k) = a \frac{\tau(k)}{P(k)} \quad \text{with} \quad a^{-1} = \int d^2k \left|\hat{\tau}(k)\right|^2 \frac{P(k)}{P(k)}.
$$

(10)

Here, $\hat{\tau}(k)$ is the Fourier transform of the expected shear profile of the halo and $P(k) = P_g + P_{\text{LSS}}(k)$ is the complete noise power spectrum including the LSS through $P_{\text{LSS}}$ and the noise contributions from the intrinsic source ellipticities and the shot noise by $P_g = \sigma_g^2 / (2n_g)$, given their angular number density $n_g$ and the intrinsic ellipticity dispersion $\sigma_g$. This filter depends on parameters determined by physical quantities such as the halo mass and redshift, the galaxy number density and the intrinsic ellipticity dispersion and not on an arbitrarily chosen scale which has to be determined empirically through costly numerical simulations (e.g. Hennawi & Spergel 2005). An application of this filter to the GaBoDS survey (Schirmer et al. 2003) was presented in Maturi et al. (2007), while a detailed comparison of these filters was performed by Pace et al. (2007) by means of numerical ray-tracing simulations. They found that the optimal linear filter given by Eq. (10) returns the halo sample with the largest completeness (100% for masses $M \geq 3 \times 10^{14} h^{-1} M_\odot$ and $\sim 50\%$ for masses $M \sim 2 \times 10^{14} h^{-1} M_\odot$ for sources at $z_s = 1$) and the smallest number of spurious detections caused by the LSS ($\leq 10\%$ for a signal-to-noise threshold of $S/N \sim 5$).

3.3. Weak lensing estimator and convergence

In order to simplify comparisons with numerical simulations, we convert the estimator $\bar{A}$ from Eq. (5) to an estimator of the convergence,

$$
\bar{A}(\theta) = \int d^2 \theta' U(|\theta' - \theta|),
$$

(11)

where $U$ is related to $Q$ by

$$
Q(\theta) = \frac{2}{\theta^2} \int_0^\theta d\theta' U(\theta') - U(\theta)
$$

(12)

(Schneider 1994) if the weight function $U(\theta)$ is defined to be compensated, i.e.

$$
\int d\theta' U(\theta') = 0.
$$

(13)

Equation (12) has the form of a Volterra integral equation of the first kind which can be solved with respect to $U$ once $Q$ is specified. If $\lim_{x\to 0} Q(x)/x$ is finite, the solution is

$$
U(\theta) = -Q(\theta) - \int_0^\theta d\theta' \frac{2}{\theta^2} Q(\theta'),
$$

(14)
4. Predicting weak lensing peak counts

Our analytic predictions for the number counts of weak-lensing detections as a function of their signal-to-noise ratio are based on modelling the analysed lensing data, resulting from Eq. (11), as an isotropic and homogeneous Gaussian random field. This is an extremely good approximation for the noise and the LSS components, but not for the non-linear structures such as sufficiently massive halos, as we shall discuss in Sec. 5.3.
vanishing and one positive component (see the sketch for type-0
detections in the lower panel of Fig. [2],
\[ F(r_{up}) = \kappa_{th}, \quad \eta_1(r_{up}) = 0, \quad \eta_2(r_{up}) > 0. \] (20)

Since we assume \( \kappa \) to be a homogeneous and isotropic random
field, the orientation of the coordinate frame is arbitrary and ir-
relevant. The conditions expressed by Eq. (20) define the so-
called up-crossing criterion which allows to identify the detections
and to derive their statistical properties, such as their num-
counts, by associating their definition to the Gaussian random
field variables \( F, \eta_1 \) and \( \eta_2 \).

However, this criterion is prone to fail for low thresholds,
where detections tend to merge and the isocontours tend to devi-
ate from the assumed approximately circular shape. This causes
detection numbers to be overestimated at low cut-offs because each
“peninsula” and “bay” of their profile (see type-1 in Fig 2)
would be counted as one detection. We solve this problem by
dividing the up-crossing points into those with positive (red cir-
cles) and those with negative (blue squares) curvature, \( \zeta_{11} > 0 \)
and \( \zeta_{11} < 0 \) respectively. In fact, for each detection, their dif-
ference is one (type-1) providing the correct number count. The
only exception is for those detections containing one or more
“lagoons” (type-2) since each of them decreases the detection
count by one. But since this is not a frequent case and occurs
only at very low cut-off levels, we do not consider this case here.

4.3. The number density of detections

Once the relation between the detections and the Gaussian ran-
dom variables \( y = (\kappa_{th}, \eta_1, \eta_2, \zeta_{11}) \) and their constraints from
Eq. (20) together with \( \zeta_{11} > 0 \) or \( \zeta_{11} < 0 \) are defined, we
can describe their statistical properties through the multivariate
Gaussian probability distribution given by Eq. (13) with the co-
variance matrix
\[
M = \begin{pmatrix}
\sigma_0^2 & 0 & 0 & -\sigma_1^2/2 \\
0 & \sigma_2^2/2 & 0 & 0 \\
0 & 0 & \sigma_1^2/2 & 0 \\
-\sigma_1^2/2 & 0 & 0 & 3\sigma_2^2/8
\end{pmatrix}. \tag{21}
\]

This matrix differs from that derived by Bardeen et al. (1986)
and Angrick & Bartelmann (2009) because we are dealing with
a 2-dimensional rather than a 3-dimensional field. Here, the \( \sigma_j \)
are the moments of the power spectrum \( P(\kappa) \)
\[
\sigma_j^2 = \int \frac{k^{2j+1} dk}{2\pi} P(k) \hat{W}^2(k)|\hat{Q}(k)|^2, \tag{22}
\]
where \( \hat{W}(k) \) is the frequency response of the survey given by
its geometry (see Sec. 3.2) and \( \hat{Q}(k) \) is the Fourier transform of
the filter adopted for the weak lensing estimator (see Sec. 3.2).
The determinant of \( M \) is \( (3\sigma_0^2\sigma_1^2\sigma_2^2 - 2\sigma_1^4)/32 \) and Eq. (19)
can explicitly be written as
\[
Q = \frac{1}{2} \left( \frac{2\eta_2}{\sigma_1^2} + \frac{8\epsilon_{11}\sigma_1^2 + 8\zeta_{11}\sigma_0\sigma_2^2 + 3k^2\sigma_2^2}{3\sigma_0^2\sigma_2^2 - 2\sigma_1^4} \right). \tag{23}
\]
Both \( \kappa \) and \( \eta_1 \) can be expanded into Taylor series around the points \( r_{up} \) where the up-crossing conditions are fulfilled,
\[
\kappa(r) \approx \kappa_{th} + \sum_{i=1}^{2} \eta_i (r - r_{up}) i , \quad \eta_1(r) \approx \sum_{i=1}^{2} \zeta_{11} (r - r_{up}) i. \tag{24}
\]

so that the infinitesimal volume element \( dV = d\kappa d\eta_1 \) can be written as
\[ d\kappa d\eta_1 = |\det J| dr, \] where \( J \) is the Jacobian matrix,
\[
J = \begin{pmatrix}
\partial\kappa/\partial x_1 & \partial\kappa/\partial x_2 \\
\partial\eta_1/\partial x_1 & \partial\eta_1/\partial x_2 \\
\partial\eta_2/\partial x_1 & \partial\eta_2/\partial x_2 \\
\partial\zeta_{11}/\partial x_1 & \partial\zeta_{11}/\partial x_2
\end{pmatrix} = \begin{pmatrix}
\eta_1 & \eta_2 \\
\zeta_{11} & \zeta_{12}
\end{pmatrix}, \tag{25}
\]
and \( |\det J| = |n_x\zeta_{11}| \) since \( \eta_1 = 0 \). The number density of up-
crossing points at the threshold \( \kappa_{th} \) with \( \zeta_{11} > 0 \), and \( \zeta_{11} < 0 \), \( n_x \) and \( n^- \) respectively, can thus be evaluated as
\[
n_x(\kappa_{th}) = \pm \int_0^\infty d\kappa \int_0^\infty d\kappa d\zeta_{11} |P(\kappa = \kappa_{th}, \eta_1 = 0, \eta_2, \zeta_{11})| \tag{26}
\]
where \( P(\kappa, \eta_1, \eta_2, \zeta_{11}) \) is the multivariate Gaussian defined by
Eq. (13) with \( p = 4 \), the correlation matrix (21), and the
quadratic form (24). Both expressions can be integrated analyti-
cally and their difference, \( n_{det}(\kappa_{th}) = n^+(\kappa_{th}) - n^- (\kappa_{th}) \), as-

![Fig. 2. Weak lensing detection maps. The top four panels show the segmentation of a realistic weak-lensing S/N map for increasing thresholds: 0.1, 0.5, 1, and 2, respectively. The bottom panel sketches the three discussed detections types together with the points identified by the standard and the modified up-crossing criteria. Red circles and blue squares correspond to up-crossing points for which the second field derivatives are \( \zeta_{11} > 0 \) and \( \zeta_{11} < 0 \), respectively.](image-url)
plained in Sect. 4.2 returns the number density of detections

\[ n_{\text{det}}(\kappa_{\text{th}}) = \frac{1}{4 \sqrt{2\pi} \sigma_0^2} \left( \frac{\sigma_1}{\sigma_0} \right)^2 \kappa_{\text{th}} \exp \left( -\frac{\kappa_{\text{th}}^2}{2\sigma_0^2} \right). \]  

(27)

Note how the dependence on \( \sigma_2 \) drops out of the difference \( n^+ - n^- \), leading to a very simple result.

For completeness we report the number density estimate also for the classical up-crossing criterion, Eq. (20) only, where the constraint on the second derivative of the field, \( \zeta_{11} \), is not used,

\[ n_{\text{up}}(\kappa_{\text{th}}) = \frac{1}{4 \sqrt{2\pi} \sigma_0} \left( \frac{\sigma_1}{\sigma_0} \right)^2 \exp \left( -\frac{\kappa_{\text{th}}^2}{2\sigma_0^2} \right) \times \left[ \exp \left( \frac{\kappa_{\text{th}}^2}{\sigma_0^2} \right) \sigma_0 \gamma + \sqrt{\pi} \kappa_{\text{th}} \text{erf} \left( \frac{\kappa_{\text{th}} \sigma_0^2}{\sigma_0 \gamma} \right) \right], \]  

(28)

with \( \gamma := \sqrt{3\sigma_0^2 \sigma_2^2 - 2\sigma_1^4} \). This number density converges to the correct value \( n_{\text{det}} \) for \( \kappa_{\text{th}} \to \infty \), i.e. large thresholds, because \( \text{erf}(x) \to 1 \) and \( \exp(-x^2) \to 0 \). This reflects the fact that, for large thresholds, the detection shapes become fully convex and any issues with more complex shapes disappear.

5. Analytic predictions vs. numerical simulations

We now compare the number counts of detections predicted by our analytic approach with those resulting from the analysis of synthetic galaxy catalogues produced with numerical ray-tracing simulations.

5.1. Numerical simulations

We use a hydrodynamical, numerical N-body simulation carried out with the code GADGET-2 (Springel 2005). We briefly summarise its main characteristics here and refer to Borgani et al. (2004) for a more detailed discussion. The simulation represents a concordance \( \Lambda \)CDM model, with dark-energy, dark-matter and baryon density parameters \( \Omega_{\Lambda} = 0.7, \Omega_m = 0.3 \) and \( \Omega_b = 0.04 \), respectively. The Hubble constant is \( H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1} \), with \( h = 0.7 \), and the linear power spectrum of the matter-density fluctuations is normalised to \( \sigma_8 = 0.8 \). The simulated box is a cube with a side length of 192 h^{-1} \text{ Mpc}, containing 480^3 dark-matter particles with a mass of \( 6.6 \times 10^9 h^{-1} M_\odot \) each and an equal number of gas particles with \( 8.9 \times 10^9 h^{-1} M_\odot \) each. Thus, halos of mass \( 10^{13} h^{-1} M_\odot \) are resolved into several thousands of particles. The physics of the gas component includes radiative cooling, star formation and supernova feedback, assuming zero metallicity.

This simulation is used to construct backward light cones by stacking the output snapshots from \( z = 1 \) to \( z = 0 \). Since the snapshots contain the same cosmic structures at different evolutionary stages, they are randomly shifted and rotated to avoid repetitions of the same cosmic structures along one line-of-sight. The light cone is then sliced into thick planes, whose particles are subsequently projected with a triangular-shaped-cloud scheme (TSC, Hockney & Eastwood 1988) on lens planes perpendicular to the line-of-sight. We trace a bundle of 2048 \( \times \) 2048 light rays through the light cone which start propagating at the observer into directions on a regular grid of 4.9 degrees on each side. The effective resolution of this ray-tracing simulation is of 1’ (for further detail, see Pace et al. 2007).

The effective convergence and shear maps obtained from the ray-tracing simulations are used to lens a background source population according to Eq. (4). Galaxies are randomly distributed on the lens plane at \( z = 1 \) with a number density of \( n_g = 30 \text{ arcmin}^{-2} \) and have intrinsic random ellipticities drawn from the distribution

\[ p(\epsilon) = \frac{\exp \left( 1 - |\epsilon|^2 / \sigma_r^2 \right) \exp \left( 1 / \sigma_r^2 - 1 \right)}{\pi \sigma_r^4 \exp \left( 1 / \sigma_r^2 \right) - 1}, \]  

(29)

where \( \sigma_r = 0.25 \).

Fig. 3. Comparison between the power spectrum measured from the analysis of shear catalogues derived from numerical simulations (heavy line) and the prediction resulting from the Peacock & Dodds (1999) power spectrum convolved with the survey frequency response and the adopted filter. We show from top to bottom the results for the polynomial filter with \( r_s = 5/5 \), the hyperbolic-tangent filter with \( r_s = 10/5 \) and the optimal filter with the cluster model scale radius set to \( r_s = 1' \).
Our analytic predictions for the number density of detections accounts for the full survey geometry, through the frequency response \( \hat{W}(k) \) in Eq. (22). Any survey geometry can be considered but, for sake of simplicity, we evaluate our estimates for a square-shaped field-of-view without gaps. Under these assumptions, the frequency response, \( W(k) \), is the product of a high-pass filter suppressing the scales larger than the light cone’s side length \( L_f = 2\pi/k_f = 4.9 \) deg,

\[
\hat{W}_f^2(k) = \exp\left(-\frac{k^2}{k^2_f}\right),
\]

where \( k_f = 2\pi/1 \) arcmin\(^{-1}\) as discussed in §5.1.

A comparison of the original up-crossing criterion with the new blended up-crossing criterion presented here is shown in Fig. 4 together with the number counts obtained from the numerical simulations. Only the result for the optimal filter with \( r_s = 1' \) is shown for clarity. As expected, the two criteria agree very well for high signal-to-noise ratios since the detections are mostly of type-0, i.e. approximately circular, as shown in the left panel of Fig. 4. However, the merging of detections at lower signal-to-noise ratios is correctly taken into account only by our new criterion.

This is an interesting application for cosmology, where a robust prediction of number counts in the linear regime only can be directly compared to data or, turning the argument around, can be used to statistically correct the halo number counts by using the data only. In fact, the difference of the positive and negative detection counts is caused by non linear structures only and their Poisson fluctuations.

Our analytic predictions for all filters and both positive and negative detection counts resulting from the synthetic galaxies catalogue from the numerical simulation are shown in Fig. 5. The signal-to-noise ratio tail caused by the nonlinear structures is present only in the positive detection counts, as expected. The agreement with the negative detections is within the 1-\( \sigma \).
Fig. 5. Number of weak lensing peaks, shown as a function of the signal-to-noise ratio, predicted with the analytic method presented here for the Schneider et al. (1998), the Schirmer et al. (2004) and the Maturi et al. (2005) filters from top to bottom, and increasing filter radii from left to right. The number counts generated by the intrinsic galaxy noise alone, $P_g$, and the LSS alone, $P_{LSS}$, are also shown. Numbers refer to a survey of one square degree with a galaxy number density of $n_g = 30 \text{ arcmin}^{-2}$ and an intrinsic shear dispersion of $\sigma_g = 0.25$. The results are compared with the number counts of positive as well as negative peaks detected based on the synthetic galaxy catalogues from the numerical simulation.

We finally compare the contribution of the LSS and the noise to the total signal by treating them separately. Their number counts are plotted with dashed lines in Fig. 5. All filters show an unsurprisingly large number of detections caused by the noise up to signal-to-noise ratios of 3 and a number of detections caused by the LSS increasing with the filter scale except for the optimal filter, which always suppresses their contribution to a negligible level. Thus, the LSS contaminates halo catalogues selected by weak lensing up to signal-to-noise ratios of 4 – 5 if its contribution is ignored in the filter definition. Note that the total number of detections can be obtained only by counting the peaks from the total signal, i.e. LSS plus noise, and not by adding the peaks found in the two components separately, because the blending of peaks is different for the two cases.

6. Conclusion

We have developed an analytic method for predicting peak counts in weak-lensing surveys, based on the theory of Gaussian random fields (Bardeen et al. 1986). Peaks are typically detected in shear fields after convolving them with filters of different shapes and widths. We have taken these into account by first filtering the assumed Gaussian random field appropriately and then searching for suitably defined peaks. On the way, we have argued for a refinement of the up-crossing criterion for peak detection which avoids biased counts of detections with low signal-to-noise ratio, and implemented it in the analytic peak-count prediction. Peaks in the non-linear tail of the shear distribution are underrepresented in this approach because they are highly non-Gaussian, but our method is well applicable to the prediction of spurious counts, and therefore to the quantification of the background in attempts to measure number densities of dark-matter halos. We have compared our analytic prediction to peak counts in numerically simulated, synthetic shear catalogues and found agreement at the expected level.

Our main results can be summarised as follows:

– The shape and size of the filter applied to the shear field has a large influence on the contamination by spurious detections.
For the optimal filter, the contribution by large-scale structures is low on all filter scales, while they typically contribute substantially for other filters.

- Shape and shot noise due to the background galaxies used to measure the shear from are a large source of spurious peak counts for all filters, and the dominant source for the optimal filter.
- Taken together, large-scale structure and galaxy noise contribute the majority of detections up to signal-to-noise ratios between 3–5. Only above this level, detections due to real dark-matter halos begin dominating.
- The optimal filter allows the detection of \( \sim 30–40 \) halos per square degree at signal-to-noise ratios high enough for suppressing all noise contributions. For the other filters, this number is lower by almost an order of magnitude.

Our conclusions are thus surprisingly drastic: peak counts in weak-lensing surveys are almost exclusively caused by chance projections in the large-scale structure and by galaxy shape and shot noise unless only peaks with high signal-to-noise ratios are counted. With typical filters, only a few detections per square degree can be expected at that level, while the optimal filter returns up to an order of magnitude more.

Acknowledgements. This work was supported by the Transregional Collaborative Research Centre TRR 33 (MM, MB) and grant number BA 1369/12-1 of the Deutsche Forschungsgemeinschaft, the Heidelberg Graduate School of Fundamental Physics and the IMPRS for Astronomy & Cosmic Physics at the University of Heidelberg.

References
Angrick, C. & Bartelmann, M. 2009, A&A, 494, 461
Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Borgani, S., Murante, G., Springel, V., et al. 2004, MNRAS, 348, 1078
Dietrich, J. P. & Hartlap, J. 2009, ArXiv e-prints
Erben, T., van Waerbeke, L., Mellier, Y., et al. 2000, A&A, 355, 23
Heavens, A. & Peacock, J. 1988, MNRAS, 232, 339
Hennawi, J. F. & Spiegel, D. N. 2005, ApJ, 624, 59
Hetterscheidt, M., Erben, T., Schneider, P., et al. 2005, A&A, 442, 43
Hockney, R. & Eastwood, J. 1988, Computer simulation using particles (Bristol: Hilger, 1988)
Kaiser, N. & Squires, G. 1993, ApJ, 404, 441
Kratochvil, J. M., Haiman, Z., & May, M. 2009, ArXiv e-prints
Marian, L., Smith, R. E., & Bernstein, G. M. 2009, ApJ, 698, L33
Maturi, M., Meneghetti, M., Bartelmann, M., Dolag, K., & Moscardini, L. 2005, A&A, 442, 851
Maturi, M., Schirmer, M., Meneghetti, M., Bartelmann, M., & Moscardini, L. 2007, A&A, 462, 473
Pace, F., Maturi, M., Meneghetti, M., et al. 2007, A&A, 471, 731
Peacock, J. & Dodds, S. 1996, MNRAS, 280, L19
Polyanin, A. D. & Manzhirov, A. V. 1998, Handbook of Integral Equations, ed. B. Raton (CRC Press)
Schirmer, M., Erben, T., Schneider, P., et al. 2003, A&A, 407, 869
Schirmer, M., Erben, T., Schneider, P., Wolf, C., & Meisenheimer, K. 2004, A&A, 420, 75
Schneider, M. D. & Bridle, S. 2009, ArXiv e-prints
Schneider, P. 1996, MNRAS, 283, 837
Schneider, P., van Waerbeke, L., Jain, B., & Kruse, G. 1998, MNRAS, 296, 873
Springel, V. 2005, MNRAS, 364, 1105

Appendix A: Forecast for different weak lensing surveys

For convenience, we evaluate here the expected number density of peak counts for a collection of present and future weak-lensing surveys. To give typical values, we assumed for all of them a square-shaped field of view, a uniform galaxy number density and no gaps for two main reasons. First, their fields-of-view are typically very large and thus do not affect the frequencies relevant for our evaluation. Second, the masking of bright objects can be done in many different ways which cannot be considered in this paper in any detail. Finally we fixed the sampling scale, described by Eq. (32), to be 5 times smaller than the typical filter scale in order to avoid undersampling, i.e. such that the high frequency cut-off is imposed by the filters themselves. The results are shown in Tab. A together with the number counts obtained with a simple Gaussian filter, usually used together with the Kaiser & Squires shear inversion algorithm (Kaiser & Squires, 1993).
Table A.1. Expected number counts of peak detections per square degree for $S/N = 1, 3, 5$ and for several present and future weak-lensing surveys with different intrinsic ellipticity dispersion, $\sigma_e$, and galaxy number density, $n_g$, per arcmin$^2$. For each filter, we used three different scales, namely $Q_{\text{poly}}$, $Q_{\tanh}$, and $Q_{\text{opt}}$: scale-1=2'/75, scale-2=5'/5, scale-3=11'; $Q_{\text{tanh}}$: scale-1=5', scale-2=10', scale-3=20'; $Q_{\text{opt}}$: scale-1=$10^{14}$ $M_\odot$ and scale-2=5 $\times$ $10^{14}$ $M_\odot$. $Q_{\text{gauss}}$ (Gaussian FWHM): scale-1=1', scale-2=2', scale-3=5'.

| Pan-STARRS $\sigma_e = 0.3$, $n_g = 5$ | $Q_{\text{poly}}$ | $Q_{\tanh}$ | $Q_{\text{opt}}$ | $Q_{\text{gauss}}$ |
|------------------------------------------|-----------------|---------------|-----------------|-----------------|
| scale-1                                  | 207.7           | 8.127         | 0.002           | 186.3           |
| scale-2                                  | 61.8            | 3.214         | 0.002           | 62.69           |
| scale-3                                  | 14.02           | 1.518         | 0.003           | -               |

| DES $\sigma_e = 0.3$, $n_g = 10$ | $Q_{\text{poly}}$ | $Q_{\tanh}$ | $Q_{\text{opt}}$ | $Q_{\text{gauss}}$ |
|---------------------------------|-----------------|---------------|-----------------|-----------------|
| scale-1                        | 206.6           | 9.55          | 0.004           | 288.8           |
| scale-2                        | 50.09           | 4.178         | 0.005           | 95.6            |
| scale-3                        | 11.67           | 2.339         | 0.017           | -               |

| CFHTLS $\sigma_e = 0.3$, $n_g = 20$ | $Q_{\text{poly}}$ | $Q_{\tanh}$ | $Q_{\text{opt}}$ | $Q_{\text{gauss}}$ |
|-------------------------------------|-----------------|---------------|-----------------|-----------------|
| scale-1                             | 206.6           | 9.907         | 0.004           | 324             |
| scale-2                             | 49.76           | 4.545         | 0.007           | 104.5           |
| scale-3                             | 11.49           | 2.622         | 0.025           | -               |

| Subaru $\sigma_e = 0.3$, $n_g = 30$ | $Q_{\text{poly}}$ | $Q_{\tanh}$ | $Q_{\text{opt}}$ | $Q_{\text{gauss}}$ |
|------------------------------------|-----------------|---------------|-----------------|-----------------|
| scale-1                            | 198.5           | 16.22         | 0.020           | 219.5           |
| scale-2                            | 42.84           | 10.82         | 0.117           | 42.37           |
| scale-3                            | 9.406           | 6.321         | 0.528           | -               |

| EUCLID $\sigma_e = 0.3$, $n_g = 40$ | $Q_{\text{poly}}$ | $Q_{\tanh}$ | $Q_{\text{opt}}$ | $Q_{\text{gauss}}$ |
|------------------------------------|-----------------|---------------|-----------------|-----------------|
| scale-1                            | 194.3           | 20.01         | 0.039           | 206.3           |
| scale-2                            | 42.64           | 14.25         | 0.295           | 42.97           |
| scale-3                            | 9.653           | 7.642         | 1.104           | -               |

| LSST $\sigma_e = 0.22$, $n_g = 50$ | $Q_{\text{poly}}$ | $Q_{\tanh}$ | $Q_{\text{opt}}$ | $Q_{\text{gauss}}$ |
|-----------------------------------|-----------------|---------------|-----------------|-----------------|
| scale-1                           | 174.8           | 42.42         | 0.463           | 156.8           |
| scale-2                           | 24.59           | 56.13         | 1.333           | 120.6           |
| scale-3                           | 4.519           | 8.139         | 4.889           | -               |

| SNAP $\sigma_e = 0.3$, $n_g = 100$ | $Q_{\text{poly}}$ | $Q_{\tanh}$ | $Q_{\text{opt}}$ | $Q_{\text{gauss}}$ |
|----------------------------------|-----------------|---------------|-----------------|-----------------|
| scale-1                          | 1070            | 50.42         | 0.021           | 4110            |
| scale-2                          | 3256            | 42.15         | 0.012           | 269.2           |
| scale-3                          | 113.5           | 48.96         | 1.688           | -               |