Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 7 November 2009 (MN LATEX style file v2.2)

Recovering the nonlinear density field from the galaxy distribution with a Poisson–Lognormal filter

Francisco-Shu Kitaura[1,2⋆], Jens Jasche[2] and R. Benton Metcalf[2]
[1] *SISSA, Scuola Internazionale Superiore di Studi Avanzati, via Beirut 2-4 34014 Trieste, Italy*
[2] *MPA, Max-Planck Institut für Astrophysik, Karl-Schwarzschildstr. 1, D-85748 Garching, Germany*

7 November 2009

ABSTRACT

We present a general expression for a lognormal filter given an arbitrary nonlinear galaxy bias. We derive this filter as the maximum a posteriori solution assuming a lognormal prior distribution for the matter field with a given mean field and modeling the observed galaxy distribution by a Poissonian process. We have performed a three–dimensional implementation of this filter with a very efficient Newton–Krylov inversion scheme. Furthermore, we have tested it with a dark matter N–body simulation assuming a unit galaxy bias relation and compared the results with previous density field estimators like the inverse weighting scheme and Wiener filtering. Our results show good agreement with the underlying dark matter field for overdensities even above $\delta \sim 1000$ which exceeds by one order of magnitude the regime in which the lognormal is expected to be valid. The reason is that for our filter the lognormal assumption enters as a prior distribution function, but the maximum a posteriori solution is also conditioned on the data. We find that the lognormal filter is superior to the previous filtering schemes in terms of higher correlation coefficients and smaller Euclidean distances to the underlying matter field. We also show how it is able to recover the positive tail of the matter density field distribution for a unit bias relation down to scales of about $\gtrsim 2$ Mpc/h.

Key words: (cosmology:) large-scale structure of Universe – galaxies: clusters: general – catalogues – galaxies: statistics

1 INTRODUCTION

The luminous matter we observe on the sky represents only a small fraction of the total matter in the Universe and yet with a careful treatment of the observational selection effects and the processes of galaxy formation we can hope to extract valuable information about the distribution of all matter from the distribution of luminous matter alone. The more precise the techniques for making this connection are the better we will be able to test our theories for the history of the Universe.

In 1934 Hubble found that the distribution of galaxy counts in cells on the sky is well fitted by a lognormal distribution (Hubble 1934). More recently, Wild et al. (2005) showed that this model is valid at least down to gridding scales of about 10 Mpc for galaxies in the 2DF catalogue. As galaxies are good tracers of matter on large cosmological scales the lognormal model should also apply to the matter field at least to some degree. Kitaura et al. (2009) showed recently that the matter field reconstruction using (least squares) Wiener filtering is very well fit by a lognormal distribution after smoothing the reconstruction with a Gaussian kernel of radius $r_{\rm S}$ for $10\,{\rm Mpc} \lesssim r_{\rm S} \lesssim 30\,{\rm Mpc}$.

From a physical point of view, one would expect the density field to be lognormally distributed after it has been smoothed on an appropriate scale. This follows from assuming an initially Gaussian density and velocity field and extrapolating the continuity equation for the matter flow into the nonlinear regime with the linear velocity fluctuations (see Coles & Jones 1991). Since the lognormal field is not able to describe caustics, we expect this distribution to fail below some threshold smoothing scale. Kayo et al. (2001) demonstrated that the lognormal distribution is a good approximation up to overdensties of about $\delta \sim$100.

Shortly after the success of Wiener filtering in large–scale structure reconstruction (see Zaroubi et al. 1995), which assumes a Gaussian prior for the matter field, a reconstruction filter based on the lognormal prior distribution was proposed (see Sheth 1995). With such a filter nonlinearities in the density field should be better recovered. In Sheth (1995) the Wiener filter was generalized to be applied to a lognormal distribution by a variable transformation. A problem with this approach is that the noise covariance has a complex form even for the simple Poisson likelihood assumption and is difficult to efficiently apply to realistic data–sets.

⋆ E-mail: kitaura@sissa.it, kitaura@mpa-garching.mpg.de

The idea of modeling the galaxies as Poisson–sampled from a lognormal underlying field was first applied to data by Saunders & Ballinger (2000). They proposed to use a filtering scheme based on an expansion of the logarithm of the matter field as a sum of harmonics. The density reconstruction using this technique as presented by Saunders et al. (2000) is fairly smooth and nonlinear structures cannot be easily recognised. This could be due to the sparse sampling of the PSCz catalogue which was used in their study or to the truncation of the harmonic series.

As demonstrated in Kitaura & Enßlin (2008), the Poissonian likelihood can be easily regularised by combining it with a prior when estimating the maximum a posteriori. They showed this calculation for Gaussian and entropic prior distribution functions.

The idea of using the full Poissonian likelihood without remaining at second order approximations using only the noise covariance matrix is based on the Richardson–Lucy deconvolution algorithm (see Richardson 1972; Lucy 1974). Shepp & Vardi (1982) showed that this filter comes from the maximum likelihood estimate of the Poissonian likelihood. Nusser & Haehnelt (1999) proposed using this method to recover the density field from the Lyman alpha forest. The problem that arose here was that the algorithm requires truncation as it assumes a flat prior for the matter field and thus the deconvolution of the response operator is not regularised. However, as Kitaura & Enßlin (2008) pointed out, this kind of problem can be solved by introducing a prior. Enßlin et al. (2008) proposed to calculate higher order corrections to obtain an estimate for the mean of the posterior distribution by employing a generating functional formalism with a Poissonian process on top of a lognormal field for the galaxy distribution.

In this work we present a general expression for the Poisson–lognormal filter given an arbitrary nonlinear galaxy bias. We derive this filter as the maximum a posteriori solution assuming a lognormal prior distribution for the matter field with a constant mean field and modeling the observed galaxy distribution by a Poissonian process. We have performed a three–dimensional implementation of this filter with a very efficient Newton–Krylov inversion scheme extending the ARGO computer code to perform nonlinear inversions (see Kitaura & Enßlin 2008). Furthermore, we have tested it for a linear galaxy bias relation and compared the results with other density field estimators commonly used in the literature (e.g. the inverse weighting scheme and the least squares (LSQ) Wiener filter). The one–dimensional lognormal probability distribution is known to fit the matter distribution well up to overdensities of about $\delta \sim$100 as found by Kayo et al. (2001). Our results show, however, good agreement for overdensities even above $\delta \sim$1000 which exceeds by one order of magnitude the expected regime in which the lognormal is expected to be valid. The reason for this apparent disagreement is that for the filter presented here the lognormal assumption enters as a prior distribution function, but the maximum a posteriori solution is also conditioned on the data. For the same reason Kitaura et al. (2009) obtained a highly non–Gaussian distributed matter field after using LSQ–Wiener filtering which according to the Bayesian formalism assumes a Gaussian distribution. We find that the Poisson–lognormal filter has a range of applicability in recovering matter density fields down to scales of about $\gtrsim 2$ Mpc/h. However, the matter statistics show that the Poisson–lognormal filter fails to recover underdense regions $\delta \lesssim -0.6$ with very few data. In addition, we test the maximum a posteriori assuming a Gaussian prior and found that it is not capable of recovering the density field when $\delta \gg 1$ and gives negative densities in low density regions which makes this filter unreliable for recovering densities of $\delta \lesssim 1$.

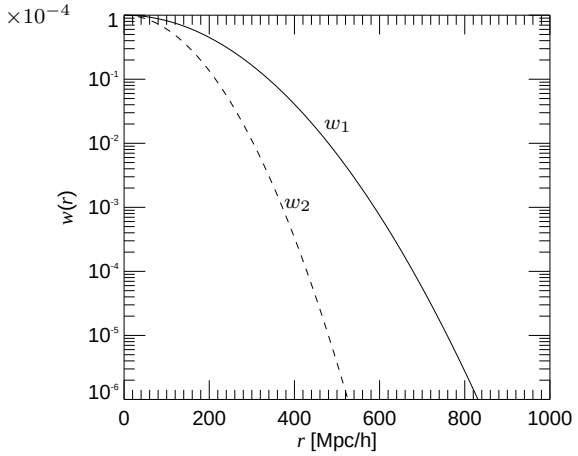

Figure 1. Two models of completeness w emulating apparent magnitude limit effects (continuous curve: w_1 and dashed curve: w_2) dependent on the distance r to the observer in Mpc/h.

Finally, we show in appendix A that the LSQ filter is the optimal linear filter under a Poisson noise assumption using up to second order statistics and does not neglect any signal to noise correlation, contrary to what has been assumed in the literature (see for example Zaroubi et al. 1995; Seljak 1998; Erdoğdu et al. 2004; Kitaura & Enßlin 2008). We also derive in appendix B a filter with a lognormal model for the underlying signal and an additive, signal–independent and Gaussian distributed noise which could be of interested in other fields of astronomy.

The paper is structured as follows. In section 2 we present the Bayesian approach used in this work. After defining the likelihood for the galaxy sample and the prior distributions for the matter field we calculate the maximum a posteriori (MAP) estimates for the underlying density field. Then the numerical scheme is presented in section 3 which permits us to solve the MAP estimates. We then present in section 4 a series of numerical experiments which show the performance of the different density estimators. Finally, we discuss our results.

2 BAYESIAN APPROACH

A Bayesian approach requires the definition of a likelihood and a prior. A full Bayesian analysis would require the complete characterization of the posterior distribution using sampling schemes (see e.g. Wandelt et al. 2004). We leave such an approach for a forthcoming publication and restrict ourselves here to calculate the extrema which leads to the maximum a posteriori expressions. This permits us to get a fast estimate of the density field. In this work, we consider a Poissonian likelihood for the observed distribution of galaxies and combine it with a Gaussian and a lognormal prior distribution for the overdensity field. In the next subsections these distribution functions are presented and the calculation of the different MAP–estimators are shown in detail.

2.1 Poissonian likelihood

The likelihood represents the observation process which leads to the data. It is the probability distribution function that describes the nature of the observable. In this case we look for a model that

accounts for the discrete nature of a galaxy distribution, the so–called shot noise. This kind of noise is traditionally modeled by a Poissonian distribution (see for example Peebles 1980a). Such a model assumes that each cell of the Universe in which we count some number of galaxies (maybe according to a certain luminosity type) is statistically independent from each other. However, the variance of counts in cells including a correlation term predicts the non–Poissonian character of the distribution of galaxies (see Peebles 1980a). Hierarchical structure formation models assume that galaxies form inside dark matter halos via the energy dissipation by baryons (see e. g. White & Rees 1978). Somerville et al. (2001) showed based on numerical N–body simulations that in regions of lower than average overdensity, the scatter in the halo biasing (the relation between the dark matter halos and the underlying dark matter distribution) is generally smaller than the mean Poisson shot noise, and in overdense regions it is larger (for a review on the halo model see Cooray & Sheth 2002). Mo & White (1996) already pointed out that halo–exclusion can cause sub–Poisson variance. Casas-Miranda et al. (2002) demonstrated with higher resolved N–body simulations that the galaxy biasing process, as well as the halo biasing process, is not only determined by the local value of the mass density field, but also by other local quantities, such as clumpiness, and by non–local properties, such as the large–scale tidal field. Accounting for all these effects is out of scope of this work, but should certainly be further investigated.

Here, we will restrict ourselves to a model in which the observed distribution of galaxies is given by an inhomogeneous Poisson realization of a continuous density field. We define the likelihood function as (see Peebles 1980b):

$$\mathcal{L}(\langle \boldsymbol{N}^{\rm o}_{\rm g}\rangle_{\rm g}|\boldsymbol{N}^{\rm o}_{\rm g}) = \Pi_{i=1}^{N_{\rm cells}} \exp\left[-\langle N^{\rm o}_{{\rm g},i}\rangle_{\rm g}\right] \frac{\langle N^{\rm o}_{{\rm g},i}\rangle_{\rm g}^{N^{\rm o}_{{\rm g},i}}}{N^{\rm o}_{{\rm g},i}!}, \quad (1)$$

with $N^{\rm o}_{{\rm g},i}$ denoting the number count of observed galaxies in cell i, and $N_{\rm cells}$ being the total number of cells. Here $\langle\{\,\}\rangle_{\rm g} \equiv \langle\{\,\}\rangle_{(N^{\rm o}_{\rm g}|\lambda^{\rm o})} \equiv \sum_{N^{\rm o}_{\rm g}=0}^{\infty} P_{\rm Pois}(N^{\rm o}_{\rm g} \mid w\lambda)\{\,\}$ denotes an ensemble average over the Poissonian distribution with the expected number of galaxy counts given by the Poissonian ensemble average: $\lambda^{\rm o} \equiv w\lambda \equiv \langle N^{\rm o}_{\rm g}\rangle_{\rm g}$. The expected number count is related to the underlying continuous galaxy overdensity field $\delta_{{\rm g},i}$ through:

$$\langle N^{\rm o}_{{\rm g},i}\rangle_{\rm g} \equiv \overline{N}_{\rm g} w_i (1+\delta_{{\rm g},i}), \quad (2)$$

where $\overline{N}_{\rm g}$ is the mean number count of galaxies and w_i the completeness at cell i. The logarithm of the likelihood can be written as:

$$\ln \mathcal{L}_i = -\overline{N}_{\rm g} w_i (1+\delta_{{\rm g},i}) + N^{\rm o}_{{\rm g},i} \ln(\overline{N}_{\rm g} w_i (1+\delta_{{\rm g},i})) - \ln(N^{\rm o}_{{\rm g},i}!). \quad (3)$$

2.2 Gaussian prior

The prior probability distribution function describes the statistical nature of the signal one wants to infer from the observed data. Here the physical model of the underlying matter field comes in. As inflationary scenarios predict a close to Gaussian distribution function for the initial density fluctuations (see Guth 1981; Guth & Pi 1982; Starobinsky 1982; Hawking 1982; Linde 1982; Albrecht & Steinhardt 1982; Bardeen et al. 1983) and linear theory preserves this property throughout cosmic evolution it is reasonable to assume a Gaussian prior to model the large–scale matter field. Note, however, that this can only be true for $|\delta| \ll 1$ since otherwise the Gaussian distribution predicts unphysical negative densities. Here we follow Bardeen et al. (1986) to describe the prior probability distribution of the density field by a multivariate Gaussian distribution function:

$$\mathcal{P}(\boldsymbol{\delta}_{\rm M}|\boldsymbol{p}) = \frac{1}{\sqrt{(2\pi)^{N_{\rm cells}}\det(\mathbf{S})}} \exp\left[-\frac{1}{2}\boldsymbol{\delta}^{\dagger}_{\rm M}\mathbf{S}^{-1}\boldsymbol{\delta}_{\rm M}\right], \quad (4)$$

with $\boldsymbol{p}$ being the set of cosmological parameters which determine the autocorrelation matrix $\mathbf{S}$ and $\delta_{\rm M}$ is the overdensity in mass. The application of the autocorrelation matrix $\boldsymbol{S}$ to a vector $\boldsymbol{x}$ is a convolution of the form: $\mathbf{S}\boldsymbol{x} \equiv S(r) \circ x(r)$ (with r being the cell coordinates in configuration space of the box $r = \{i_X, i_Y, i_Z\}$ and with "$\circ$" denoting the convolution operation).

Note that the Fourier transform of the autocorrelation matrix $\mathbf{S}$ is equal to the power spectrum: $\hat{\mathbf{S}}(\boldsymbol{k},\boldsymbol{k}') \equiv (2\pi)^3 P(\boldsymbol{k}')\delta_{\rm D}(\boldsymbol{k}-\boldsymbol{k}')$ (using the same Fourier definitions as in Kitaura & Enßlin 2008). The logarithm of the prior distribution function can be written as:

$$\ln \mathcal{P}(\boldsymbol{\delta}_{\rm M}|\boldsymbol{p}) = -\frac{1}{2}\boldsymbol{\delta}^{\dagger}_{\rm M}\mathbf{S}^{-1}\boldsymbol{\delta}_{\rm M} + c, \quad (5)$$

with c being the logarithm of the normalization.

The posterior distribution function P is proportional to the product of the prior $\mathcal{P}$ and the likelihood $\mathcal{L}$. To find the maximum a posteriori (MAP) we need to calculate the extremum. Performing the derivative of the posterior with respect to the matter overdensity field $\delta_{\rm M}$ yields:

$$\frac{\partial \ln P}{\partial \boldsymbol{\delta}_{\rm M}} \propto \frac{\partial \ln \mathcal{P}}{\partial \boldsymbol{\delta}_{\rm M}} + \frac{\partial \ln \mathcal{L}}{\partial \boldsymbol{\delta}_{\rm M}} = 0. \quad (6)$$

The derivative of the prior leads to:

$$\frac{\partial \ln \mathcal{P}}{\partial \boldsymbol{\delta}_{\rm M}} = -\mathbf{S}^{-1}\boldsymbol{\delta}_{\rm M}. \quad (7)$$

Since the likelihood is expressed as a function of the galaxy density field, we need to define the bias between the galaxy and matter fields.

2.2.1 *Linear bias*

As a particular case, let us consider a linear bias function given by:

$$\delta_{{\rm g},i} = \sum_j b_{i,j}\delta_{{\rm M},j}, \quad (8)$$

which relates the corresponding power spectra in the following way: $\hat{b}(\boldsymbol{k}) = \sqrt{P_{\rm g}(\boldsymbol{k})/P_{\rm M}(\boldsymbol{k})}$, with $P_{\rm g}(\boldsymbol{k})$ being the galaxy power–spectrum and $P_{\rm M}(\boldsymbol{k})$ being the matter power–spectrum.

The derivative of the likelihood with respect to the matter overdensity fields is then given by:

$$\sum_i \frac{\partial \ln \mathcal{L}_i}{\partial \delta_{{\rm M},k}} = \sum_i b_{i,k}\left[-\overline{N}_{\rm g} w_i + \frac{N^{\rm o}_{{\rm g},i}}{1+\sum_l b_{i,l}\delta_{{\rm M},l}}\right]. \quad (9)$$

Adding this result to the prior term Eq. (7) we obtain the MAP equation:

$$\delta^{\rm G}_{{\rm M},i} = \sum_j S_{i,j}\sum_l b_{l,j}\left(\frac{N^{\rm o}_{{\rm g},l}}{1+\sum_k b_{l,k}\delta^{\rm G}_{{\rm M},k}} - \overline{N}_{\rm g} w_l\right), \quad (10)$$

with the superscript G standing for the Gaussian prior assumption.

2.2.2 *Unity bias*

Let us consider the special case when the matter field is equal to a *continuous* galaxy field:

$$\delta_{{\rm g},i} = \delta_{{\rm M},i}, \quad (11)$$

then the MAP equation reads:

$$\delta^{\rm G}_{{\rm M},i} = \sum_j S_{i,j} \left(\frac{N^{\rm o}_{{\rm g},j}}{1+\delta^{\rm G}_{{\rm M},j}} - \overline{N}_{\rm g} w_j \right). \quad (12)$$

2.3 Lognormal prior

Here we introduce the lognormal prior distribution as proposed by Coles & Jones (1991):

$$\mathcal{P}(\boldsymbol{s}|\boldsymbol{p}) = \frac{1}{\sqrt{(2\pi)^{N_{\rm cells}}\det(\mathbf{S}_{\rm L})}} \exp\left[-\frac{1}{2}(\boldsymbol{s}-\boldsymbol{\mu})^\dagger \mathbf{S}_{\rm L}^{-1}(\boldsymbol{s}-\boldsymbol{\mu})\right], \quad (13)$$

with $\boldsymbol{s}$ being the logarithm of the weighted matter density:

$$s_i \equiv \log(1+\delta_{{\rm M}i}), \quad (14)$$

and $\mathbf{S}_{\rm L}$ the corresponding autocorrelation matrix. Note that the lognormal autocorrelation function $\mathbf{S}_{\rm L}$ applied to a vector $\boldsymbol{x}$ is again a convolution: $\mathbf{S}_{\rm L}\boldsymbol{x} \equiv S_{\rm L}(r) \circ x(r)$. The transformation of the correlation function corresponding to the overdensity field to the signal s is given by:

$$S_{\rm L}(r) \equiv \log(1+S(r)). \quad (15)$$

The mean field $\boldsymbol{\mu}$ is taken to be:

$$\mu_i \equiv -\sigma_0^2/2, \quad (16)$$

with $\sigma_0^2 \equiv S_{\rm L}(0)$ as used by Kayo et al. (2001) to ensure an overdensity field with zero mean[1] (for a formal derivation see Coles & Jones 1991). The logarithm of the prior distribution yields:

$$\ln \mathcal{P}(\boldsymbol{s}|\boldsymbol{p}) = -\frac{1}{2}(\boldsymbol{s}-\boldsymbol{\mu})^\dagger \mathbf{S}_{\rm L}^{-1}(\boldsymbol{s}-\boldsymbol{\mu}) + c, \quad (17)$$

with c being some constant term. In this case, we look for the extremum with respect to the signal $\boldsymbol{s}$:

$$\frac{\partial \ln P}{\partial \boldsymbol{s}} = \frac{\partial \ln \mathcal{P}}{\partial \boldsymbol{s}} + \frac{\partial \ln \mathcal{L}}{\partial \boldsymbol{s}} = 0. \quad (18)$$

The derivative of the prior has now an additional term due to the mean–field $\boldsymbol{\mu}$:

$$\frac{\partial \ln \mathcal{P}}{\partial \boldsymbol{s}} = -\mathbf{S}_{\rm L}^{-1}(\boldsymbol{s}-\boldsymbol{\mu}). \quad (19)$$

Now we need to relate the galaxy field to the matter field in order to express the likelihood as a function of the signal statistically defined through the prior distribution function.

2.4 General nonlinear bias

Let us consider here a general nonlinear relation between the galaxy field and the matter field.

$$\delta_{{\rm g},i} = B(\boldsymbol{\delta}_{\rm M})_i, \quad (20)$$

The derivative of the likelihood with respect to the signal $\boldsymbol{s}$ which we want to recover can be written as:

$$\sum_i \frac{\partial \ln \mathcal{L}_i}{\partial s_k} = \sum_i \sum_l \frac{\partial \ln \mathcal{L}_i}{\partial \ln(1+\delta_{{\rm g},l})} \frac{\partial \ln(1+\delta_{{\rm g},l})}{\partial \ln(1+\delta_{{\rm M},k})}. \quad (21)$$

The derivative of the likelihood with respect to the logarithm of the normalized galaxy field $\ln(1+\delta_{{\rm g},l})$ yields:

$$\frac{\partial \ln \mathcal{L}_i}{\partial \ln(1+\delta_{{\rm g},l})} = \left(N^{\rm o}_{{\rm g},i} - \overline{N}_{\rm g} w_i \left(1 + B(\boldsymbol{\delta}_{\rm M})_i\right)\right) \delta^{\rm K}_{i,l}. \quad (22)$$

[1] Here generalized to a multivariate lognormal distribution.

The factor relating the galaxy field to the matter field yields:

$$\frac{\partial \ln(1+\delta_{{\rm g},l})}{\partial \ln(1+\delta_{{\rm M},k})} = \frac{\partial \ln(1 + B\left(\exp(\boldsymbol{s}) - \vec{1}\right)_l)}{\partial s_k} = \frac{\partial B(\boldsymbol{\delta}_{\rm M})_l}{\partial \delta_{{\rm M},k}} \frac{(1+\delta_{{\rm M},k})}{1+B(\boldsymbol{\delta}_{\rm M})_l}. \quad (23)$$

The final result for the derivative of the likelihood with respect to the logarithm of the normalized matter field $\boldsymbol{s}$ assuming a general nonlinear bias is given by:

$$\sum_i \frac{\partial \ln \mathcal{L}_i}{\partial s_k} = \quad (24)$$

$$\sum_i \frac{\partial B\left(\boldsymbol{\delta}_{\rm M}\right)_i}{\partial \delta_{{\rm M},k}} \frac{(1+\delta_{{\rm M},k})}{1+B(\boldsymbol{\delta}_{\rm M})_i} \left(-\overline{N}_{\rm g} w_i \left(1 + B(\boldsymbol{\delta}_{\rm M})_i\right) + N^{\rm o}_{{\rm g},i}\right).$$

Combining this result with the prior term (Eq. 19) we obtain the MAP equation:

$$\sum_j S^{-1}_{{\rm L}i,j} \left(\ln(1+\delta^{\rm L}_{{\rm M},j}) - \mu_j\right) = \quad (25)$$

$$\sum_l \frac{\partial B\left(\boldsymbol{\delta}^{\rm L}_{\rm M}\right)_l}{\partial \delta^{\rm L}_{{\rm M},i}} \frac{(1+\delta^{\rm L}_{{\rm M},i})}{1+B(\boldsymbol{\delta}^{\rm L}_{\rm M})_l} \times \left(N^{\rm o}_{{\rm g},l} - \overline{N}_{\rm g} w_l \left(1 + B(\boldsymbol{\delta}^{\rm L}_{\rm M})_l\right)\right),$$

with the superscript L standing for the lognormal prior.

2.4.1 *Linear bias*

For the linear bias case the derivative of the likelihood reduces to the following expression:

$$\sum_i \frac{\partial \ln \mathcal{L}_i}{\partial s_k} = \sum_i \frac{b_{i,k}(1+\delta_{{\rm M},k})}{1+\sum_{l'} b_{i,l'}\delta_{{\rm M},l'}} \times \left(-\overline{N}_{\rm g} w_i \left(1 + \sum_j b_{i,j}\delta_{{\rm M},j}\right) + N^{\rm o}_{{\rm g},i}\right). \quad (26)$$

Accordingly, the MAP equation reads:

$$\sum_j S^{-1}_{{\rm L}i,j} \left(\ln\left(1+\delta^{\rm L}_{{\rm M},j}\right) - \mu_j\right) = \quad (27)$$

$$\sum_k \frac{b_{k,i}(1+\delta^{\rm L}_{{\rm M},k})}{1+\sum_{l'} b_{k,l'}\delta^{\rm L}_{{\rm M},l'}} \left(N^{\rm o}_{{\rm g},k} - \overline{N}_{\rm g} w_k \left(1 + \sum_l b_{k,l}\delta^{\rm L}_{{\rm M},l}\right)\right).$$

2.4.2 *Unity bias*

For an unity bias the derivative of the likelihood reduces to:

$$\delta_{{\rm g},i} = \delta_{{\rm M},i}. \quad (28)$$

Then, the derivative of the likelihood reduces to:

$$\sum_i \frac{\partial \ln \mathcal{L}_i}{\partial s_k} = -\overline{N}_{\rm g} w_k \exp(s_k) + N^{\rm o}_{{\rm g},k}, \quad (29)$$

and the MAP equation reads:

$$\sum_j S^{-1}_{{\rm L}i,j}(s_j - \mu_j) = N^{\rm o}_{{\rm g},i} - \overline{N}_{\rm g} w_i \exp(s_i). \quad (30)$$

Using the definitions: $\overline{N}_{\rm g} w_j \exp(s_j) = \overline{N}_{\rm g} w_j (1+\delta_{{\rm g},i}) = \langle N^{\rm o}_{\rm g}\rangle_{\rm g}$ and $\epsilon^{\rm o}_j \equiv N^{\rm o} - \langle N^{\rm o}_{\rm g}\rangle_{\rm g}$ we can rewrite the MAP equation as:

$$\sum_j S^{-1}_{{\rm L}i,j}(s_j - \mu_j) = \epsilon^{\rm o}_i, \quad (31)$$

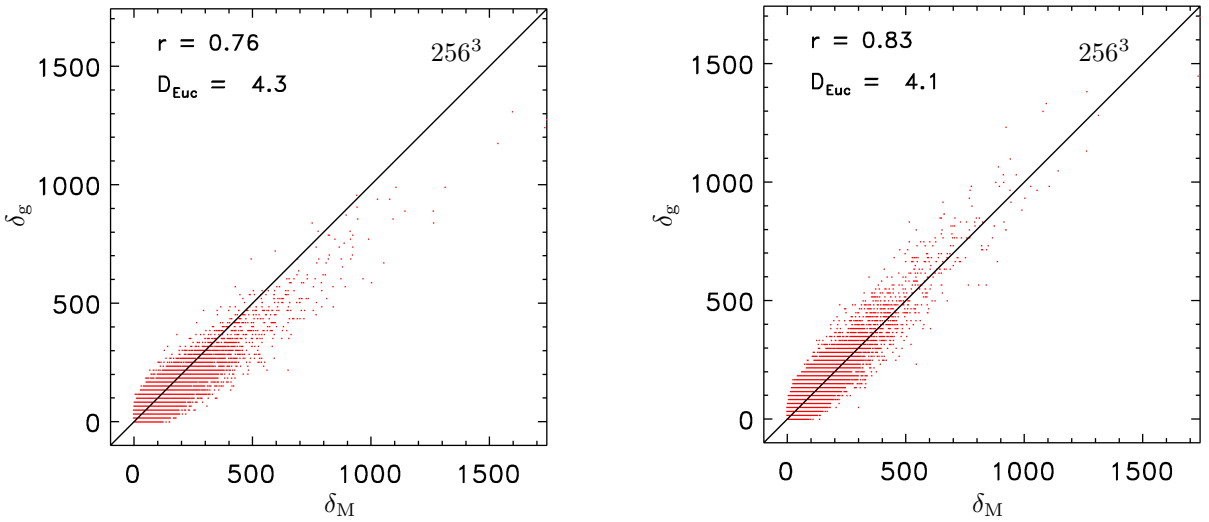

Figure 2. Cell-to-cell overdensity correlation between the mock galaxy sample and the matter field based on the Millenium run (Springel et al. 2005). On the left: mock galaxy catalogue by De Lucia & Blaizot (2007). On the right: Poisson sample over the matter field with a constant completeness of 10^{-4}. Also given are the correlation coefficient r and Euclidean distance to the underlying matter field $\mathrm{D_{Euc}}$.

The signal $\boldsymbol{s}$ is thus given by the propagation of the noise $\boldsymbol{\epsilon}^{\mathrm{o}}$ given the correlation $\mathbf{S}_{\mathrm{L}}$ up to a shift due to the mean $\boldsymbol{\mu}$. Expressing it as a function of the matter overdensity-field we obtain:

$$\sum_j S^{-1}_{\mathrm{L}i,j}\left(\ln\left(1+\delta^{\mathrm{L}}_{\mathrm{M},j}\right)-\mu_j\right)=\left(N^{\mathrm{o}}_{\mathrm{g},i}-\overline{N}_{\mathrm{g}}w_i(1+\delta^{\mathrm{L}}_{\mathrm{M},i})\right). \tag{32}$$

3 NUMERICAL APPROACH

The problem we are studying here requires the solution of a nonlinear system of 256^3 (about $17\cdot 10^6$) coupled equations. Note, that each cell introduces an equation. Thus, to find the MAP solution (Eqns. 12 and 26) we apply an operator based iterative inversion scheme as proposed in Kitaura & Enßlin (2008) which reduces the most expensive operations to FFTs. In particular, we use a nonlinear Newton–Krylov scheme which is briefly presented in the next subsections (for a reference see e.g. Kitaura & Enßlin 2008).

3.1 Method

Let us write a system of nonlinear equations as: $A(\boldsymbol{x})=\boldsymbol{f}$, with $\boldsymbol{A}$ being the nonlinear operator dependent on $\boldsymbol{x}$ and $\boldsymbol{f}$ some constant vector. We then define the gradient of the quadratic approximation as:

$$\nabla Q(\boldsymbol{x})\equiv A(\boldsymbol{x})-\boldsymbol{f}. \tag{33}$$

The corresponding Hessian matrix is then given by the second derivative of the gradient of Q:

$$\mathbf{H}\equiv\nabla\nabla Q(\boldsymbol{x}). \tag{34}$$

The basic Newton–Raphson solver scheme is given by:

$$\boldsymbol{x}^{j+1}=\boldsymbol{x}^j-\left(\mathbf{H}^j\right)^{-1}\nabla Q(\boldsymbol{x}^j). \tag{35}$$

This scheme turns out to be extremely inefficient. Therefore, we implement a Krylov step in which the solution is updated in the following way:

$$\boldsymbol{x}^{j+1}=\boldsymbol{x}^j+\tau^j\xi^j, \tag{36}$$

with the stepsize τ given by (for a derivation see Kitaura & Enßlin 2008):

$$\tau^j=-\frac{\boldsymbol{\xi}^{j\dagger}\nabla Q(\boldsymbol{x}^j)}{\boldsymbol{\xi}^{j\dagger}\mathbf{H}^j\boldsymbol{\xi}^j}. \tag{37}$$

and ξ being the searching vector (for schemes to calculate ξ see Kitaura & Enßlin 2008). In the next subsections we give the particular expressions for the quantities required to calculate the MAP given a Gaussian prior first and finally given a lognormal prior.

3.1.1 *Gaussian prior*

Rewriting Eqn. 12 as:

$$\boldsymbol{\delta}^{\mathrm{G}}_{\mathrm{M}}-\mathbf{S}\,\mathrm{diag}(\mathbf{1}+\boldsymbol{\delta}^{\mathrm{G}}_{\mathrm{M}})^{-1}\boldsymbol{N}^{\mathrm{o}}_{\mathrm{g}}=\overline{N}_{\mathrm{g}}\mathbf{S}\boldsymbol{w}_{\mathrm{g}}. \tag{38}$$

We can identify $A(\boldsymbol{s})=\boldsymbol{\delta}^{\mathrm{G}}_{\mathrm{M}}-\mathbf{S}\,\mathrm{diag}(\mathbf{1}+\boldsymbol{\delta}^{\mathrm{G}}_{\mathrm{M}})^{-1}\boldsymbol{N}^{\mathrm{o}}_{\mathrm{g}}$ and $\boldsymbol{f}=\overline{N}_{\mathrm{g}}\mathbf{S}\boldsymbol{w}_{\mathrm{g}}$. The corresponding gradient of the quadratic form is given by:

$$\nabla Q(\boldsymbol{s})=\boldsymbol{\delta}^{\mathrm{G}}_{\mathrm{M}}-\mathbf{S}\left(\mathrm{diag}(\mathbf{1}+\boldsymbol{\delta}^{\mathrm{G}}_{\mathrm{M}})^{-1}\boldsymbol{N}^{\mathrm{o}}_{\mathrm{g}}-\overline{N}_{\mathrm{g}}\boldsymbol{w}_{\mathrm{g}}\right), \tag{39}$$

and the Hessian yields:

$$\mathbf{H}=\mathbb{1}+\mathbf{S}\,\mathrm{diag}(\mathbf{1}+\boldsymbol{\delta}^{\mathrm{G}}_{\mathrm{M}})^{-2}\boldsymbol{N}^{\mathrm{o}}_{\mathrm{g}}. \tag{40}$$

3.1.2 *Lognormal prior*

We formulate Eq. 32 in an analogous way to the previous subsection as:

$$\mathbf{S}^{-1}_{\mathrm{L}}(\boldsymbol{s}-\boldsymbol{\mu})+\mathrm{diag}\left(\overline{N}_{\mathrm{g}}\boldsymbol{w}_{\mathrm{g}}\right)\exp(\boldsymbol{s})=\boldsymbol{N}^{\mathrm{o}}_{\mathrm{g}}, \tag{41}$$

with $A(\boldsymbol{s})=\mathbf{S}^{-1}_{\mathrm{L}}(\boldsymbol{s}-\boldsymbol{\mu})+\mathrm{diag}\left(\overline{N}_{\mathrm{g}}\boldsymbol{w}_{\mathrm{g}}\right)\exp(\boldsymbol{s})$ and $\boldsymbol{f}=\boldsymbol{N}^{\mathrm{o}}_{\mathrm{g}}$. Thus, the gradient of the quadratic form is given by:

$$\nabla Q(\boldsymbol{s})=\mathbf{S}^{-1}_{\mathrm{L}}(\boldsymbol{s}-\boldsymbol{\mu})+\mathrm{diag}\left(\overline{N}_{\mathrm{g}}\boldsymbol{w}_{\mathrm{g}}\right)\exp(\boldsymbol{s})-\boldsymbol{N}^{\mathrm{o}}_{\mathrm{g}}, \tag{42}$$

and the corresponding Hessian matrix reads:

$$\mathbf{H}=\mathbf{S}^{-1}_{\mathrm{L}}+\mathrm{diag}\left(\overline{N}_{\mathrm{g}}\boldsymbol{w}_{\mathrm{g}}\right)\mathrm{diag}\left(\exp(\boldsymbol{s})\right). \tag{43}$$

	LIKELIHOODS	
	Gaussian	**Poissonian**
PRIORS		
Flat	(a)	$\delta^{\rm IW}_{{\rm g},i} = \frac{N^{\rm o}_{{\rm g},i}}{w_i \overline{N}_{\rm g}} - 1$
Gaussian	$\delta^{\rm LSQ}_{{\rm M},i} = \sum_j \left(S^{-1}_{i,j} + w_j \overline{N}_{\rm g} \delta^{\rm K}_{i,j}\right)^{-1} \overline{N}_{\rm g} \delta^{\rm o}_{{\rm g},j}$	$\delta^{\rm G}_{{\rm M},i} = \sum_j S_{i,j} \left(\frac{N^{\rm o}_{{\rm g},j}}{1+\delta^{\rm G}_{{\rm M},j}} - \overline{N}_{\rm g} w_j\right)$
Lognormal	(b)	$\sum_j S^{-1}_{{\rm L}i,j} \left(\ln\left(1+\delta^{\rm L}_{{\rm M},j}\right) - \mu_j\right) = N^{\rm o}_{{\rm g},i} - \overline{N}_{\rm g} w_i (1+\delta^{\rm L}_{{\rm M},i})$

Table 1. Filters which are used in this work classified by the assumed likelihood and prior (with the exception of (a) and (b)). Note, that the bias has been set to one. (a) COBE-filter used in CMB mapping (see Janssen & Gulkis 1992). (b) for a derivation of this filter see appendix B.

4 NUMERICAL EXPERIMENTS

In this section we investigate the performance of the Poisson–lognormal filter. We construct the mock observed galaxy distributions by making a Poisson sample over the dark matter particles of the Millennium run according to different completeness models (see subsection 4.2). This permits us to avoid the galaxy biasing and redshift distortions problems in our tests.

We also test the Poisson–lognormal filter against other filters. We make a comparison to the MAP with a Gaussian prior assumption, to the inverse weighted data, and to the LSQ–Wiener filter (see section below 4.1.1). For an overview of the filters used in this work see table 1.

Finally, we test the quality of the reconstruction by making a cell–to–cell comparison to the underlying matter field which is assumed to be given by the dark matter distribution of the Millennium run at redshift zero (see Springel et al. 2005). In addition, we study the matter density statistics of the dark matter field and the reconstructions.

4.1 Quality validation of the density reconstruction

In order to show the performance of the Poisson–lognormal filter we compare the results with two other estimates of the density field. We follow Kitaura & Enßlin (2008) and Kitaura et al. (2009) in quantitatively measuring the quality of the reconstructions.

4.1.1 Alternative density field estimators: Inverse weighting and LSQ–Wiener filtering

Let us first introduce a representation of the data which tries to compensate for the selection function effect which we call inverse weighting (IW). We define the inverse weighted galaxy number count per cell i as:

$$N^{\rm IW}_{{\rm g},i} \equiv \frac{1}{w_i} N^{\rm o}_{{\rm g},i}. \tag{44}$$

The corresponding inverse weighted overdensity is calculated as follows:

$$\delta^{\rm IW}_{{\rm g},i} \equiv \frac{N^{\rm IW}_{{\rm g},i}}{\overline{N}_{\rm g}} - 1. \tag{45}$$

Note that the inverse weighting scheme can be derived as the maximum likelihood estimator assuming a Poissonian likelihood (for a derivation see Kitaura et al. 2009). As discussed in Kitaura et al. (2009) IW boosts the estimated density field at low completeness. Therefore it includes in general an additional smoothing step which lessons this effect (see e.g. Erdoğdu et al. 2004).

For an additional comparison let us introduce the least squares version of the Wiener filter (or LSQ filter for short) given by Kitaura et al. (see 2009):

$$\boldsymbol{\delta}^{\rm LSQ}_{\rm M} \equiv \mathbf{B}^{-1} \left(\mathbf{S}^{-1} + \mathbf{W}^\dagger \mathbf{N}^{-1} \mathbf{W}\right)^{-1} \mathbf{W}^\dagger \mathbf{N}^{-1} \boldsymbol{\delta}^{\rm o}_{\rm g}, \tag{46}$$

with $\mathbf{W}$ being the three dimensional mask operator defined by: $W_{i,j} \equiv w_j \delta^{\rm K}_{i,j}$ ($\delta^{\rm K}_{i,j}$ is the Kroenecker delta), and the Fourier transform of $\mathbf{B}$ given by: $\hat{\hat{B}}_{k,k'} \equiv b_{k'} \delta^{\rm K}_{k,k'}$ as introduced in subsection 2.4.1. We define the observed galaxy overdensity which we use as the input vector for the LSQ reconstruction by:

$$\delta^{\rm o}_{{\rm g},i} \equiv \frac{N^{\rm o}_{{\rm g},i}}{\overline{N}_{\rm g}} - w_i. \tag{47}$$

The noise term in Eq. 46 has the following form (see Kitaura et al. 2009):

$$N_{i,j} \equiv \frac{w_i}{\overline{N}_{\rm g}} \delta^{\rm K}_{i,j}. \tag{48}$$

Note that the LSQ–Wiener filter is the optimal linear filter using up to second order statistics even for the Poisson–noise assumption for which the noise is signal dependent. There is not an additional assumption or approximation by neglecting the signal to noise correlation. This point has been unclear in the literature (see for example Zaroubi et al. 1995; Seljak 1998; Erdoğdu et al. 2004). We show that the signal and the noise are indeed uncorrelated in the appendix. The LSQ–Wiener filter also happens to be the MAP filter for a Gaussian likelihood and a Gaussian prior as indicated in table 1.

In our numerical experiments we use a unity bias: $b_k = 1$. Thus, the Fourier transform of $\mathbf{S}$ yields: $P_{\rm g}(\boldsymbol{k}) = P_{\rm M}(\boldsymbol{k})$. The power–spectrum $P_{\rm M}(\boldsymbol{k})$ is given by a nonlinear fit that also describes the effects of virialised structures with a halo term as given by Smith et al. (2003) at redshift $z = 0$. We choose the concordance ΛCDM–cosmology with $\Omega_{\rm m} = 0.24$, $\Omega_{\rm K} = 0$ and $\Omega_\Lambda = 0.76$ (Spergel et al. 2007). In addition, we assumed a Hubble constant with $h = 73$ and a spectral index $n_s = 1$.

4.1.2 Quantitative measures

Let us define the correlation coefficient r between the reconstructed and original matter density fields by:

$$\mathrm{r}(\delta^{\rm rec}, \delta_{\rm M}) \equiv \frac{\sum_i^{N_{\rm cells}} \delta_{{\rm M},i} \delta^{\rm rec}_i}{\sqrt{\sum_i^{N_{\rm cells}} \left(\delta_{{\rm M},i}\right)^2} \sqrt{\sum_j^{N_{\rm cells}} \left(\delta^{\rm rec}_j\right)^2}}, \tag{49}$$

and the Euclidean distance

$$\mathrm{D}_{\rm Euc}(\delta^{\rm rec}, \delta_{\rm M}) \equiv \sqrt{\frac{1}{N_{\rm cells}} \sum_i^{N_{\rm cells}} \left(\delta^{\rm rec}_i - \delta_{{\rm M},i}\right)\right)^2}. \tag{50}$$

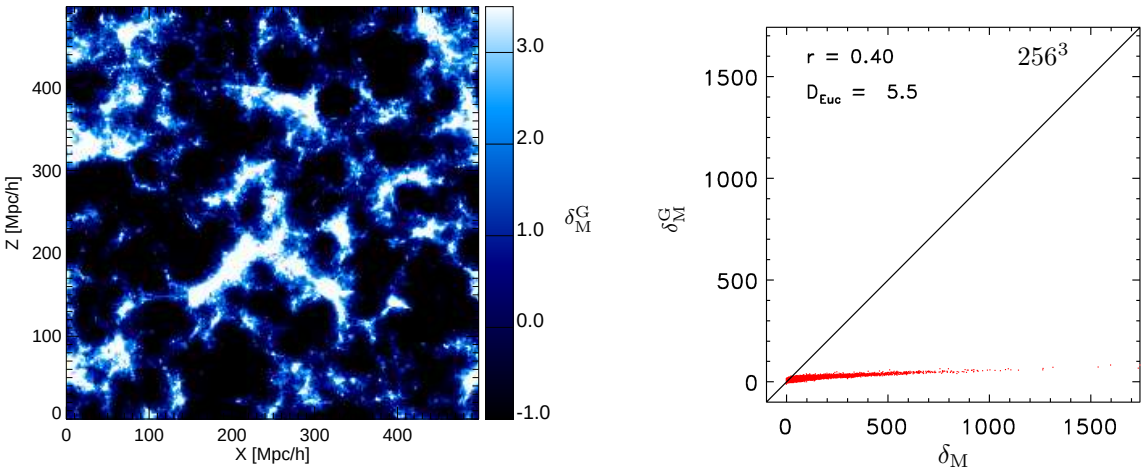

Figure 3. Reconstruction assuming a Gaussian prior. Left panel: mean over 15 neighboring slices around Y$\sim$ 176 Mpc through a 500 Mpc cube box with a 256^3 grid without smoothing. Right panel: cell–to–cell correlation between the overdensity of the full three–dimensional reconstruction and the matter field.

4.2 Input data setup

We construct the mock observed galaxy distribution taking a random subsample of the particles in the Millennium run at redshift zero (see Springel et al. 2005) which was gridded on a 256^3 mesh. Later we also investigate the resolution dependence using a 128^3, and a 64^3 mesh. As already stated above, our setup permits us to avoid the galaxy biasing problem in our tests. Note, that we also avoid the redshift distortions by considering the dark matter particles in real–space. In this way we generate three different input mock galaxy catalogues. One has about 1 Million particles and has been produced as a Poisson sampling with a homogeneous completeness of $w = 10^{-4}$. The other two mocks were generated with a radial selection function using two exponential decaying models of completeness w (see Fig. 1) emulating apparent magnitude limit effects (see radial selection function in Kitaura et al. 2009). The final mock galaxy samples have 350961 and 123679 particles using the softer and steeper decaying selection functions respectively. The observer was set at the center of the box, i.e. at coordinates: X=250 Mpc/h, Y=250 Mpc/h, and Z=250 Mpc/h.

The left panel on Fig. 2 shows a cell–to–cell comparison between the dark matter distribution of the Millennium run ($\sim 10^{10}$ particles) gridded on a 256^3 mesh and a subsample of 1 Million homogeneously selected mock galaxies of the De Lucia & Blaizot (2007) catalogue. One can clearly see a deviation of the pixels with respect to the perfect slope of 45°. This effect is due to galaxy biasing. The right panel shows the analougous comparison with a Poisson sampling using a homogeneous completeness of $w = 10^{-4}$ which leaves about 1 Million particles. Here, we see a nearly perfect scatter around the 45° slope demonstrating that our mocks do not include biasing.

4.3 MAP results

Here we calculate the maximum a posteriori solutions which we derived in the previous theoretical sections. There we assumed two different prior distributions for the matter field: a Gaussian and a lognormal prior.

4.3.1 Gaussian prior and Poissonian likelihood

In this subsection we present the results given by the maximum a posteriori solution assuming a Gaussian prior and a Poissonian likelihood. The solution of Eqn. 12 leads to a matter field which dramatically underestimates large overdensities (see Fig. 3). This shows that the Gaussian prior cannot fit the underlying matter field which has a clearly non–Gaussian distribution with a minimum overdensity of ~ -1 up to maximal overdensities of about 1500 at the resolution we are looking at ($\sim$ 2 Mpc/h cell side length). The density peaks are highly suppressed with a Gaussian prior. This effect is known from the Wiener filter as traditionally applied where the noise covariance is dependent on the signal (see discussion in Kitaura et al. 2009). Note, that the filter we are using here is a more accurate being based on the full Poissonian distribution and not only on the second order term as in the Wiener filter.

4.3.2 Lognormal prior and Poissonian likelihood

Here we present the results of the maximum a posteriori solution assuming a lognormal prior and a Poissonian likelihood. For that we solve the MAP Eqn. 32.

We show in Fig. 4 the performance of the Poisson–lognormal filter with a homogeneous completeness. Panel a in Fig. 4 shows a slice through the matter distribution from the Millennium run. Panel b shows the mock galaxy sample. Panels c and d show the Poisson–lognormal filter and the LSQ filter reconstruction respectively. The performance depicted in cell–to–cell correlation plots shown in panels e and f demonstrate the superior behaviour of the Poisson–lognormal filter reconstruction in terms of higher correlation, smaller Euclidean distances and better alignment along the perfect correlation slope. The Poisson–lognormal filter recovers the density field up to overdensities above 1500 whereas the LSQ filter tends to underestimate the density field.

We study the inhomogeneous completeness effects by selecting dark matter particle subsamples with two different radial selection functions depicted in Fig. 1. In the upper panel of Fig. 5, the inverse weighting scheme is shown to overestimate the density at low completeness (at the borders and corners of the cube). This is in agreement with tests performed by Kitaura et al. (2009). The

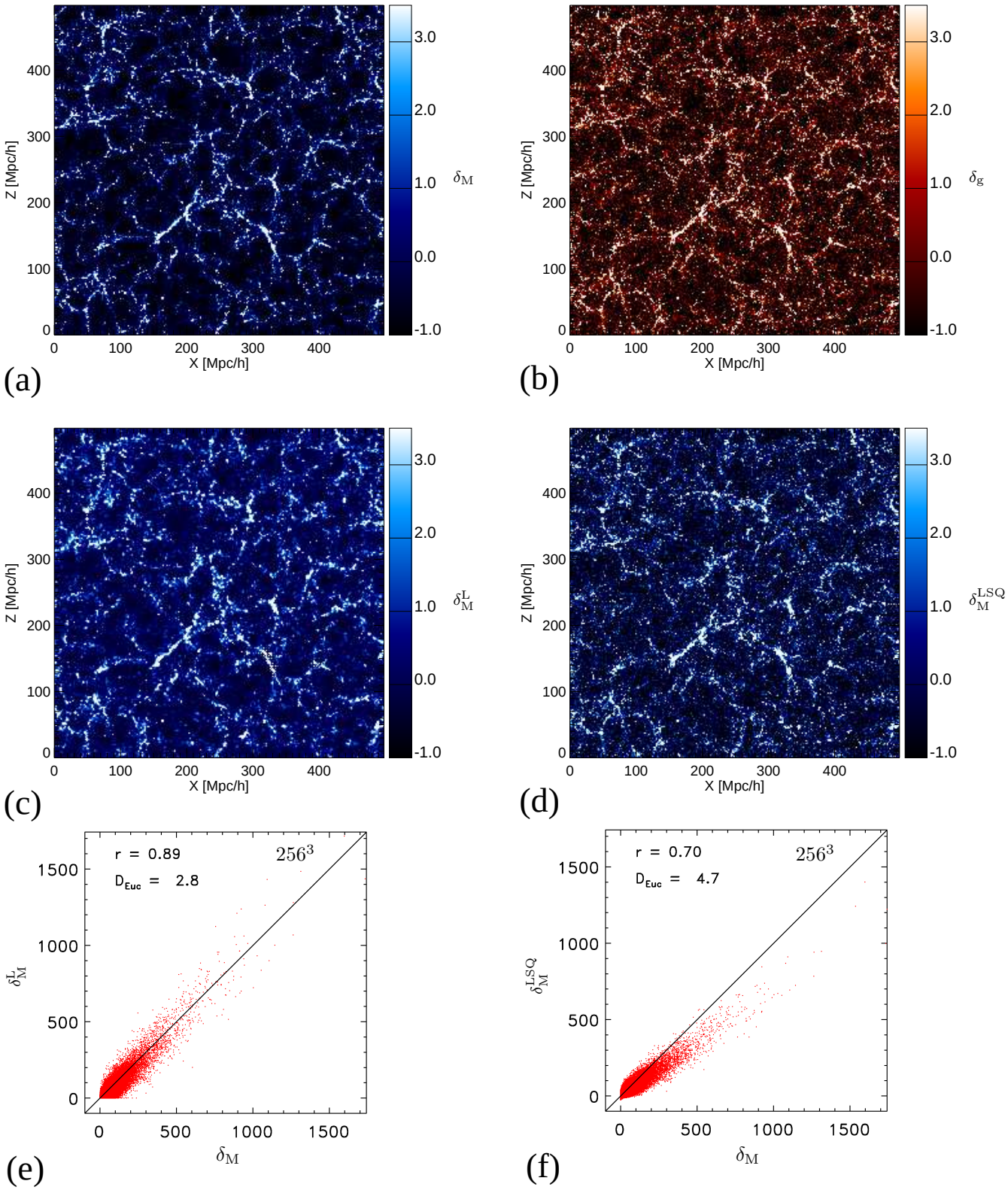

Figure 4. Panel a: slice through the Millenium run dark matter particle simulation. Panel b: mock galaxy distribution with 10^6 particles. Panel c: reconstruction with the lognormal filter. Panel d: reconstruction with the LSQ-Wiener filter. Panel e: cell–to–cell correlation between the overdensity of the full three–dimensional reconstruction with the Lognormal filter (panel c) and the matter field (panel a). Panel f: cell–to–cell correlation between the full three–dimensional overdensity of the reconstruction with the LSQ-Wiener filter (panel d) and the matter field (panel a). The cell–to–cell correlation between the mock galaxy distribution (panel b) and the dark matter distribution can be seen on the right panel of Fig. 2. The plots were produced by calculating the mean over 15 neighboring slices around slice 218 (Y$\sim$ 176 Mpc/h) through a 500 Mpc/h cube box with a 256^3 grid.

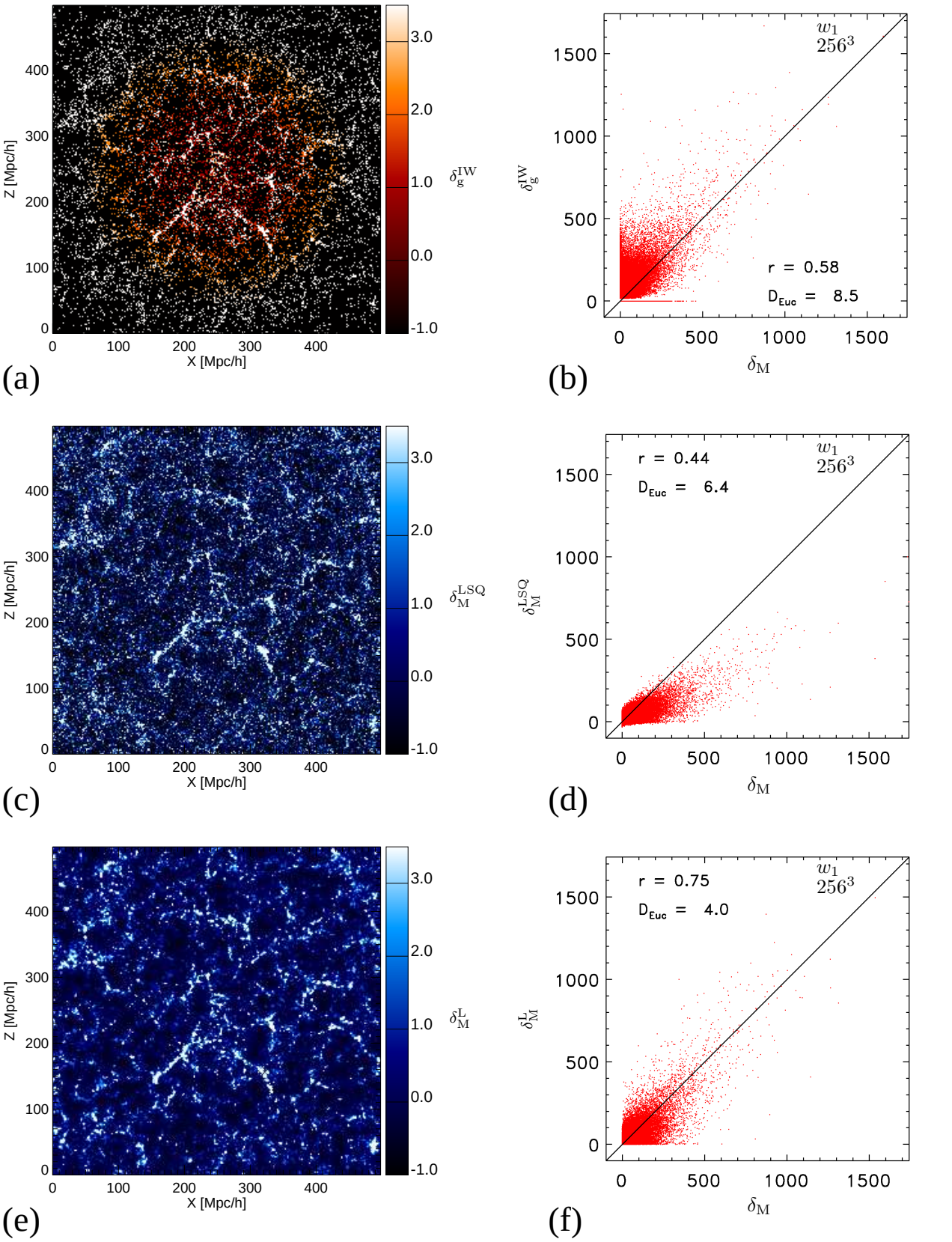

Figure 5. Panel a: inverse weigted mock galaxy distribution after applying a radial selection function (w_1) leaving 350961 galaxies. Panels c: LSQ-Wiener filter reconstruction. Panels e: Lognormal filter reconstruction. The plots were produced by calculating the mean over 15 neighboring slices around slice 218 (Y$\sim$ 176 Mpc/h) through a 500 Mpc/h cube box with a 256^3 grid. The performance depicted in cell–to–cell correlation plots are shown in the right hand side panels b, e and f.

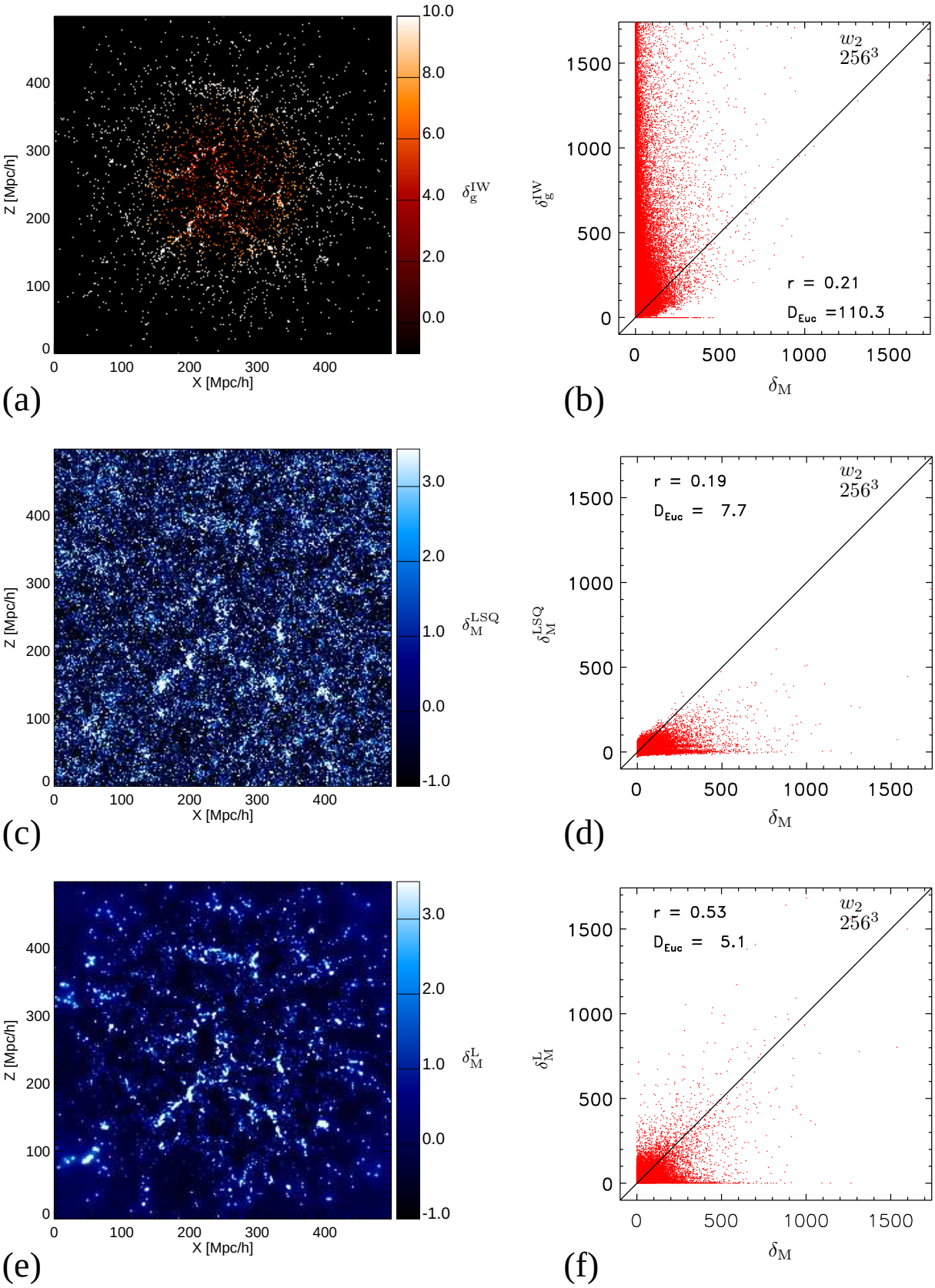

Figure 6. The same as Fig. 5 corresponding to the radial selection function w_2 with a mock galaxy distribution of 123679 particles.

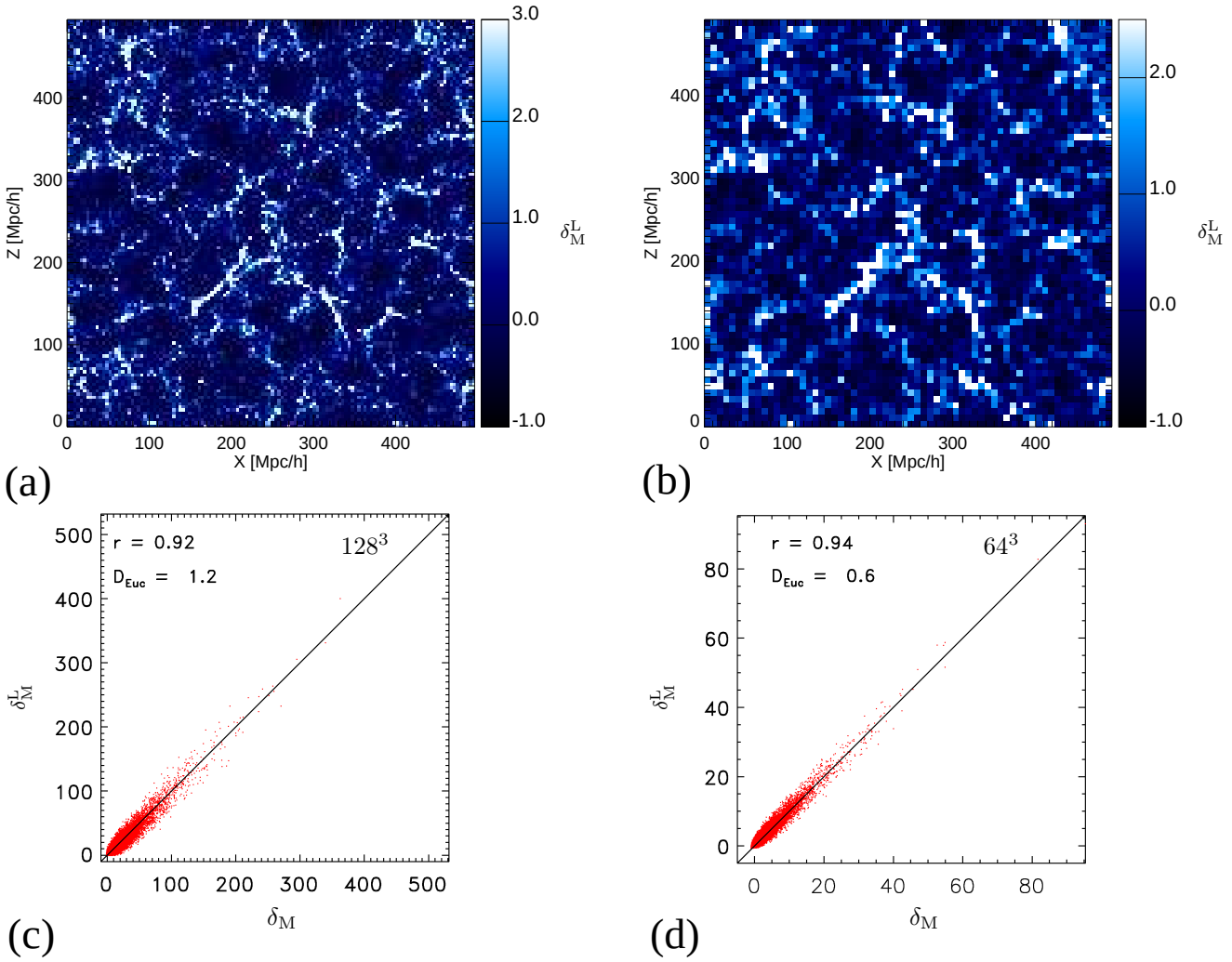

Figure 7. Matter field reconstructions with the lognormal filter on a grid mesh with 128^3 and 64^3 cells for a uniform selection using about 10^6 mock galaxies. Panel a: mean over 9 slices through the reconstruction on a mesh with 128^3 cells around slice 109 ($Y \sim 179$ Mpc/h). Panel b: mean over 5 slices through the reconstruction on a mesh with 64^3 cells around slice 55 ($Y \sim 179$ Mpc/h). Panels c and d show the cell–to–cell statistics corresponding to the full reconstructions shown in panels a and b, respectively.

LSQ filter, on the other hand, smooths the density more strongly in at low completeness regions and leads to a significantly lower Euclidean distance. The correlation coefficient is lower since the LSQ filtering suppresses the signal and gives a smooth version of the density field which is valid on larger scales (see panels c and d of Fig. 5), but does not reproduce small-scale features. The lower panels show the results coming from the Poisson–lognormal filter reconstruction. The density at low completeness is suppressed to zero due to the mean field used for this calculation (see Eq. 16). In regions of very low completeness the filter tends to favor the mean density. The statistical correlation shows to be clearly superior to the previous cases and the Euclidean distance with respect to the underlying matter field is far smaller (see panel f). The cell–to–cell correlation plot shows a scatter around the 45 slope and reproduces even the highest overdensities like the one at $\sim$1600 which can also be seen in panel b. We perform the analogous study with the steeper radial selection function (see Fig. 1). The results are shown in Fig. 6 and are consistent with the previously discussed ones.

We perform the same study for two more resolutions: a mesh with 128^3 cells and a mesh with 64^3 cells for the same comoving box. First we grid the dark matter field coming from the Millenium run on the lower resolution mesh and then we apply the radial selection $w(r)$ using $w = 10^{-4}$, $w(r) = w_1(r)$ and $w(r) = w_2(r)$. The results of the Poisson–lognormal filter reconstructions are shown in Figs. 7 and 8. We see a clear tendency to better recover the underlying matter field when using a lower resolution (compare Fig. 4 with Fig. 7 and Figs. 5, 6 with Fig. 8).

4.3.3 Matter statistics

Finally, we calculate the matter statistics and the corresponding skewness and kurtosis for the dark matter field and the lognormal reconstructions corresponding to the incomplete mocks with selection functions w_1 and w_2 (for particular expressions to calculate the matter statistics, the skewness and the kurosis see Kitaura et al. 2009). The matter statistics represented in Fig. 9 shows consistent results for different grid resolutions (compare left, middle and right panels). After convolving the matter fields with a Gaussian kernel using a smoothing radius of 10 Mpc/h the matter distribution appears to be closely lognormal distributed for all resolutions (see

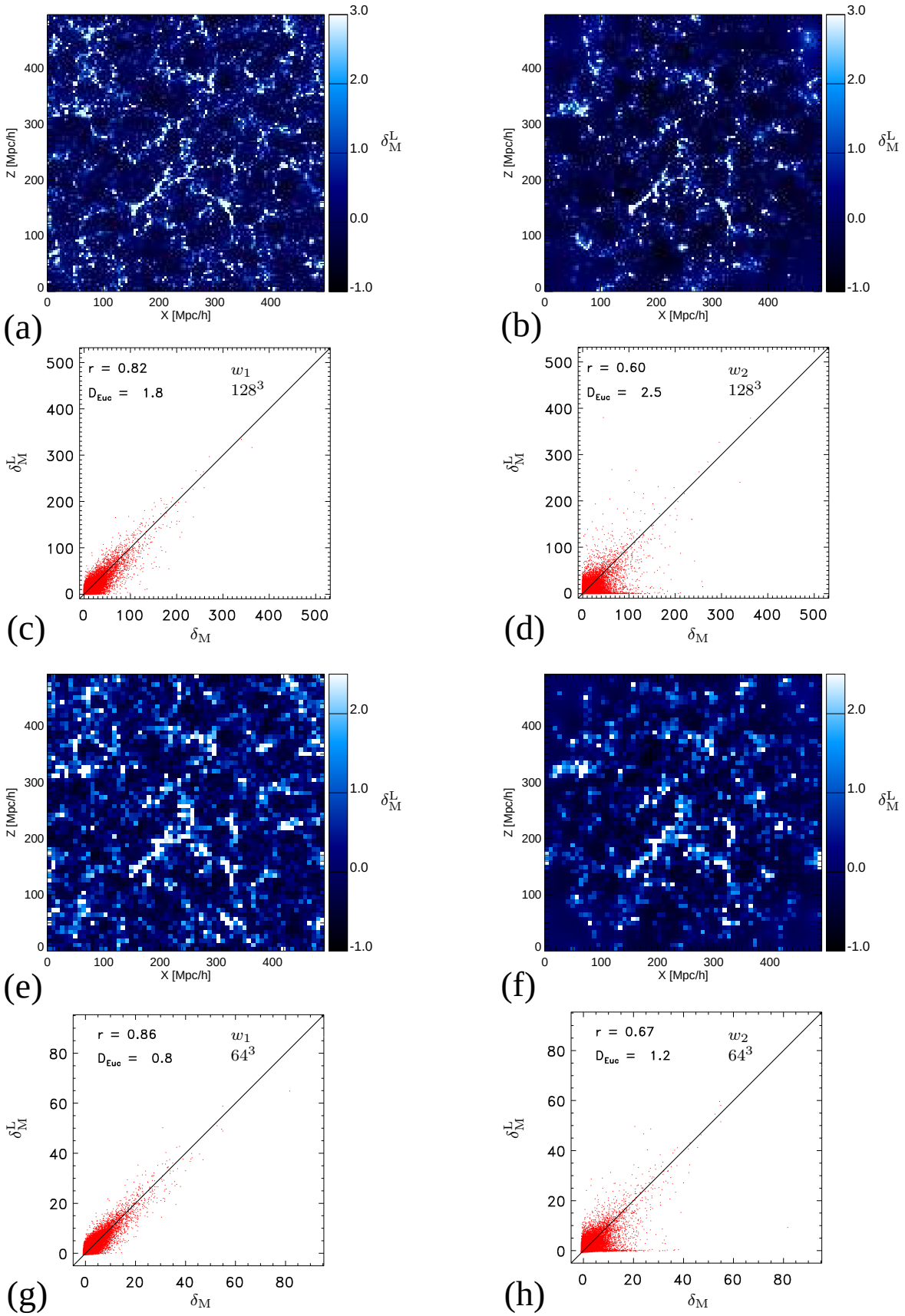

Figure 8. Matter field reconstructions with the lognormal filter on a grid mesh with 128^3 and 64^3 cells for both w_1 and w_2 selection criteria. Panel a: same as panel a in previous figure for the case w_1. Panel b: same as panel a for the case w_2. Panels c and d show the cell–to–cell statistics corresponding to the full reconstructions shown in panels a and b, respectively. Panel e: same as panel a on a mesh with 64^3. Panel f: same as panel e for the case of w_2. Panels g and h show the cell–to–cell statistics corresponding to the full reconstructions shown in panels e and f, respectively.

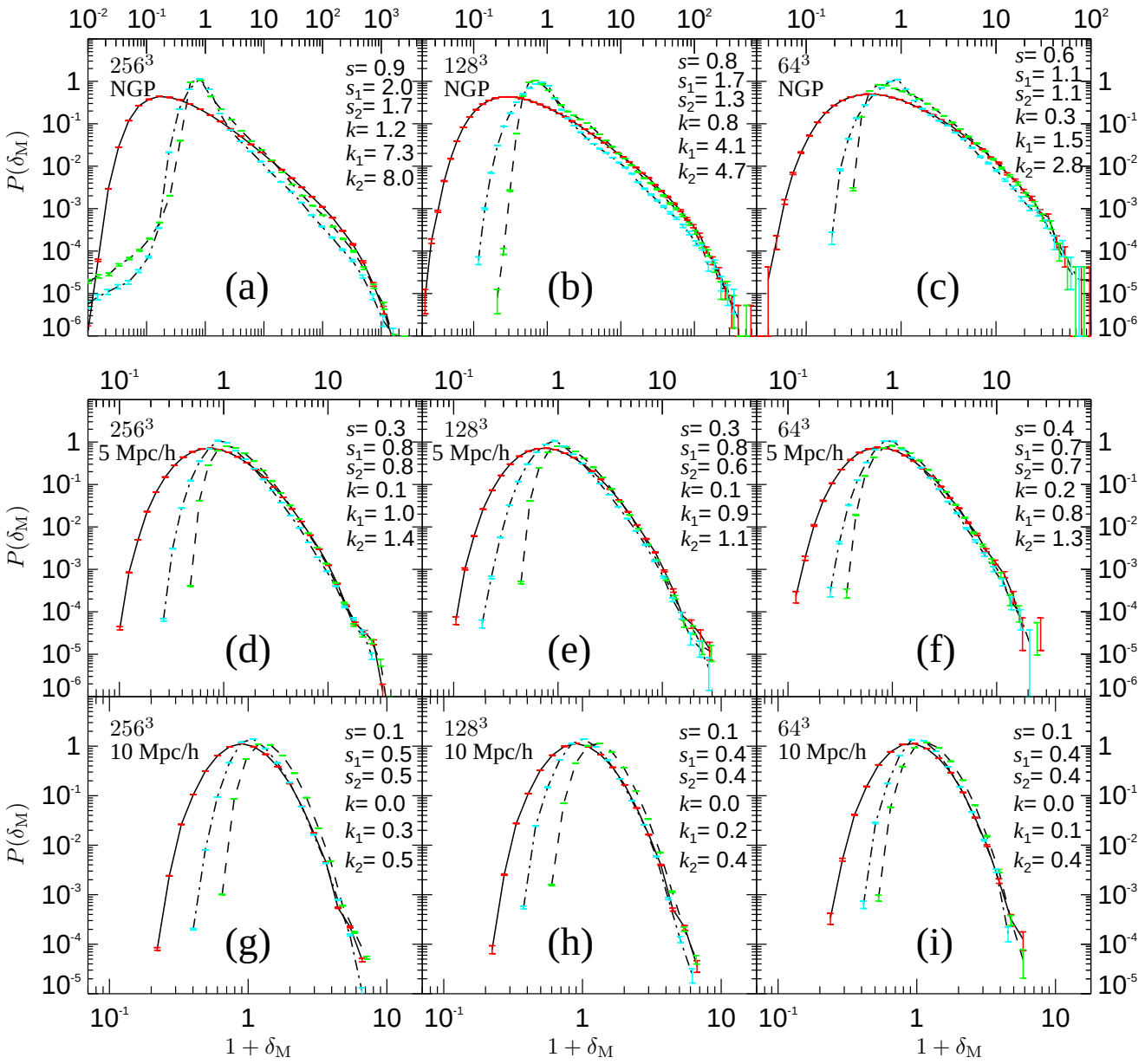

Figure 9. Matter statistics for the dark matter field from the Millenium run (black curve, red error bars) using about $\sim 10^{10}$ particles and the corresponding reconstructions using the selected mocks with radial completeness w_1 (dashed curve, green error bars) and w_2 (dashed–dotted curve, cyan error bars) having about $\sim 10^5$ particles for different resolutions (256^3: left panels, 128^3: middle panels and 64^3: right panels). Upper panels (a, b and c): without smoothing. Lower panels: after convolution with a Gaussian kernel with smoothing radii of 5 Mpc/h (panels d, e and f) and 10 Mpc/h (panels g, h and i). The number of cells was counted for a logarithmic density binning of 0.2 in $\ln(1+\delta_{\rm M})$ for all cases except for panel a for which a binning of 0.4 was used. Also skewness (s, s_1 and s_2) and kurtosis (k, k_1 and k_2) in $\ln(1+\delta_{\rm M})$ are shown corresponding to the matter field, the reconstructions for the case w_1 and the case w_2, respectively. The error bars are given by the shot noise caused by the number counts of cells in each density bin without taking into account the uncertainties introduced by the completeness or the reconstruction method itself.

panels at the bottom). The skewness (s, s_1 and s_2) and the kurtosis (k, k_1 and k_2) show some deviation from zero particularly in the reconstructed fields (skewness and kurtosis without subindex correspond to the dark matter field, with the subindex "1" to the selected sample with w_1 and with the subindex "2" to the selected sample with w_2). However, their values are small which means that the distributions are not especially peaked or have significantly longer tails with respect to the lognormal distribution. This result is consistent with observations (Kitaura et al. 2009) where for a similar smoothing radius the matter field obtained from the Sloan Digital Sky survey (data release 6) was found to be close to lognormal distributed. When convolving with a Gaussian kernel with a 5 Mpc/h smoothing radius (panels d, e and f) the distribution shows a tail towards larger densities with higher skewness and kurtosis than for the panels at the bottom which cannot be attributed to the uncertainty at the high densities shown by the large error bars caused by the low number counts in that regime. This deviation from the lognormal distribution is even better demonstrated in the upper panels which show the matter statistics without any additional smoothing. The results show that the multivariate lognormal prior distribution

does not impose a lognormal matter field statistics to the recovered density field. Here the prior is subdominant with respect to the data similarly to the case of the LSQ–Wiener reconstruction which can lead to non–Gaussian statistics even though it is based on a Gaussian prior (see Kitaura et al. 2009).

However, we also observe several effects causing a deviation in the reconstructed fields from the *true* matter field statistics. The reconstructions tend to overestimate the number of cells around the mean density (see the peaks of the distribution in the upper panels of Fig. 9). This trend is more acute for the stronger sampled mock for which w_2 was used. This can be seen by comparing the dashed–dotted curves with the dashed curves and the kurtosis k_1 with k_2 (k_2 is always larger than k_1). We also observe that the reconstructions underestimate the number of cells in the extremely underdense regions ($\delta \lesssim -0.6$). In addition, we investigated the statistics of the reconstructed fields based on the homogeneously sampled mock data using $w = 10^{-4}$ and found that the densities are distributed in a very similar way to the reconstructed matter fields with w_1. These effects are caused by the conservative character of the reconstruction method. The maximum a posteriori solution leads to a stronger smoothing in the undersampled low–density regions and produces a larger number of cells with densities closer to the mean.

5 CONCLUSIONS

In this work we have presented a general expression for the Poisson–lognormal filter given an arbitrary nonlinear galaxy bias. We derived this filter as the maximum a posteriori solution assuming a lognormal prior distribution for the matter field with a constant mean field and modeling the observed galaxy distribution by a Poissonian process (see Eq. 26).

We have performed a three–dimensional implementation of this filter with a very efficient Newton–Krylov inversion scheme (see section 3). Furthermore, we have tested it for a linear galaxy bias relation and compared the results with other density field estimators commonly used in the literature (e.g. the inverse weighting scheme and the least squares (LSQ) Wiener filter (see section 4)).

We also found that the solution of Eqn. 12, assuming a Gaussian prior distribution for the matter field, leads to a reconstruction which clearly underestimates large overdensities (see Fig. 3). This shows that the Gaussian prior cannot fit the underlying matter field which has a clearly non–Gaussian distribution with a minimum overdensity of $\delta \sim -1$ up to maximal overdensities of about $\delta \sim 1700$ for a resolution of ~ 2 Mpc/h. The density peaks are highly suppressed with the Gaussian prior. This effect is known from the Wiener filter as traditionally applied in which the noise covariance is dependent on the signal (see discussion in Kitaura et al. 2009).

However, we have seen that even the LSQ–Wiener filter fails for high overdensities ($\delta_{\rm M} \gtrsim 100$). We showed in appendix A that the LSQ–filter is the optimal linear filter under a Poisson noise assumption and does not neglect any signal to noise correlation, contrary to what has been assumed in literature (see for example Zaroubi et al. 1995; Seljak 1998; Erdoğdu et al. 2004; Kitaura & Enßlin 2008). The LSQ filter is the optimal linear filter only up to second order statistics and thus is less well suited to distributions with high skewness or long tails such as the lognormal distribution. Another reason for the inferior performance of the LSQ–filter with respect to the Poisson–lognormal filter is its linearity. Note, that the relation at overdensities $\delta \gg 1$ is highly nonlinear.

The one–dimensional lognormal probability distribution is known to fit well the matter distribution up to overdensities of about $\delta \sim$100 as found by Kayo et al. (2001). Our results show, however, good agreement for overdensities even above $\delta \sim$1000 which exceeds by one order of magnitude the regime in which the lognormal is expected to be valid. This is because in our filter the lognormal assumption enters as a prior distribution function, but the maximum a posteriori solution is also conditioned on the data. In a similar way Kitaura et al. (2009) was able to recover a highly non–Gaussian distributed matter field from the SDSS dr6 after using LSQ–Wiener filtering which according to the Bayesian formalism assumes a Gaussian distribution. Assuming that the galaxy bias is known we find that the Poisson–lognormal filter is able to recover the matter density fields down to scales of about $\gtrsim 2$ Mpc/h. However, our study of the matter statistics comparing the dark matter with the reconstructed fields shows that the Poisson–lognormal filter fails to recover underdense regions for $\delta \lesssim -0.6$. At lower densities the recovered field is smoothed out due to the conservative maximum a posteriori solution.

Our work shows a great improvement with respect to previous filters in recovering the matter density field from a point source distribution. Still much work has to be done to further analyse the statistical properties of the cosmological structure. Nevertheless, the nonlinear reconstruction method derived in this work could be of great interest for large scale structure density field reconstructions taking a galaxy distribution or even some other observables like the Lyman alpha forest.

ACKNOWLEDGEMENTS

We thank Ofer Lahav and Benjamin D. Wandelt for suggesting as several years ago to study the lognormal filter. Special thanks to Rien van de Weijgaert and Bernard J. T. Jones for discussions about the lognormal prior at the conference in Santander 2007. We also thank Simon D. M. White, Carlos Hernández Monteagudo, Torsten Enßlin and Gerard Lemson for encouraging conversations.

The authors thank the Intra-European Marie Curie fellowship and the Transregio TR33 Dark Universe, as well as the Munich cluster *Universe* for supporting this project and both the Max Planck Institute for Astrophysics in Munich and the Scuola Internazionale Superiore di Studi Avanzati in Trieste for generously providing the authors with all the necessary facilities.

We finally thank the German Astrophysical Virtual Observatory (GAVO), which is supported by a grant from the German Federal Ministry of Education and Research (BMBF) under contract 05 AC6VHA, for providing us with mock data.

REFERENCES

Albrecht A., Steinhardt P. J., 1982, Physical Review Letters, 48, 1220
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Bardeen J. M., Steinhardt P. J., Turner M. S., 1983, Phys. Rev. D, 28, 679
Casas-Miranda R., Mo H. J., Sheth R. K., Boerner G., 2002, MNRAS, 333, 730
Coles P., Jones B., 1991, MNRAS, 248, 1
Cooray A., Sheth R., 2002, Physics Reports, 372, 1
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
Enßlin T. A., Frommert M., Kitaura F. S., 2008, ArXiv e-prints

Erdoğdu P., Lahav O., Zaroubi S., Efstathiou G., Moody S., Peacock J. A., Colless M., Baldry I. K., et al. 2004, MNRAS, 352, 939

Guth A. H., 1981, Phys. Rev. D, 23, 347

Guth A. H., Pi S.-Y., 1982, Physical Review Letters, 49, 1110

Hawking S. W., 1982, Communications in Mathematical Physics, 87, 395

Hubble E., 1934, ApJ, 79, 8

Janssen M. A., Gulkis S., 1992, in Signore M., Dupraz C., eds, NATO ASIC Proc. 359: The Infrared and Submillimetre Sky after COBE Mapping the sky with the COBE differential microwave radiometers. pp 391–408

Kayo I., Taruya A., Suto Y., 2001, ApJ, 561, 22

Kitaura F. S., Enßlin T. A., 2008, MNRAS, 389, 497

Kitaura F. S., Jasche J., Li C., Ensslin T. A., Metcalf R. B., Wandelt B. D., Lemson G., White S. D. M., 2009, ArXiv e-prints

Linde A. D., 1982, Physics Letters B, 108, 389

Lucy L. B., 1974, AJ, 79, 745

Mo H. J., White S. D. M., 1996, MNRAS, 282, 347

Nusser A., Haehnelt M., 1999, MNRAS, 303, 179

Peebles P. J. E., 1980a, The large-scale structure of the universe. Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p.

Peebles P. J. E., 1980b, The large-scale structure of the universe. Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p.

Richardson W. H., 1972, Journal of the Optical Society of America (1917-1983), 62, 55

Rybicki G. B., Press W. H., 1992, ApJ, 398, 169

Saunders W., Ballinger W. E., 2000, in Kraan-Korteweg R. C., Henning P. A., Andernach H., eds, Mapping the Hidden Universe: The Universe behind the Mily Way - The Universe in HI Vol. 218 of Astronomical Society of the Pacific Conference Series, Interpolation of Discretely-Sampled Density Fields. pp 181–+

Saunders W., D'Mellow K. J., Valentine H., Tully R. B., Carrasco B. E., Mobasher B., Maddox S. J., Hau G. K. T., Sutherland W. J., Clements D. L., Staveley-Smith L., 2000, in Kraan-Korteweg R. C., Henning P. A., Andernach H., eds, Mapping the Hidden Universe: The Universe behind the Mily Way - The Universe in HI Vol. 218 of Astronomical Society of the Pacific Conference Series, The IRAS View of the Local Universe. pp 141–+

Seljak U., 1998, ApJ, 503, 492

Shepp L. A., Vardi Y., 1982, IEEE Trans. Med. Imaging, 1, 113

Sheth R. K., 1995, MNRAS, 277, 933

Smith R. E., Peacock J. A., Jenkins A., White S. D. M., Frenk C. S., Pearce F. R., Thomas P. A., Efstathiou G., Couchman H. M. P., 2003, MNRAS, 341, 1311

Somerville R. S., Lemson G., Sigad Y., Dekel A., Kauffmann G., White S. D. M., 2001, MNRAS, 320, 289

Spergel D. N., Bean R., Doré O., Nolta M. R., Bennett C. L., Dunkley J., Hinshaw G., Jarosik N., et al. 2007, .Rev.Astrn.Astrophys., 170, 377

Springel V., White S. D. M., Jenkins A., Frenk C. S., Yoshida N., Gao L., Navarro J., Thacker R., Croton D., Helly J., Peacock J. A., Cole S., Thomas P., Couchman H., Evrard A., Colberg J., Pearce F., 2005, Nature, 435, 629

Starobinsky A. A., 1982, Physics Letters B, 117, 175

Tegmark M., 1997, ApJ, 480, L87+

Wandelt B. D., Larson D. L., Lakshminarayanan A., 2004, Phys. Rev. D, 70, 083511

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wiener N., 1949, Extrapolation, Interpolation, and Smoothing of Stationary Time Series. New York: Wiley

Wild V., Peacock J. A., Lahav O., Conway E., Maddox S., Baldry I. K., Baugh C. M., et al 2005, MNRAS, 356, 247

Zaroubi S., Hoffman Y., Fisher K. B., Lahav O., 1995, ApJ, 449, 446

APPENDIX A: THE LSQ FILTER

Here we show a derivation of the LSQ filter which does not require a data degradation model with an additive noise term. Let us adopt here the usual notation for the data: $\boldsymbol{d} \equiv \boldsymbol{\delta}_{\rm g}^{\rm o}$. The data vector is accordingly defined by:

$$\boldsymbol{d} \equiv \frac{\boldsymbol{N}_{\rm g}^{\rm o}}{\overline{N}_{\rm g}} - \mathbf{W}\vec{1}, \tag{A1}$$

for a definition of the mask operator $\mathbf{W}$ see section 4.1.1. We define here the signal vector $\boldsymbol{s}$ as the matter overdensity field: $\boldsymbol{s} \equiv \boldsymbol{\delta}_{\rm M}$. In the linear approximation we try to find a filter $\mathbf{F}$ which applied to the data $\boldsymbol{d}$ gives an estimate of the signal $\boldsymbol{s}$ of the form:

$$\langle \boldsymbol{s} \rangle_{\rm LSQ} \equiv \mathbf{F}\boldsymbol{d}. \tag{A2}$$

This filter should minimize the following quantity in the least squares approach (see Wiener 1949; Rybicki & Press 1992; Zaroubi et al. 1995):

$$\begin{aligned} \mathcal{A} &\equiv \langle (\mathbf{F}\boldsymbol{d} - \boldsymbol{s})^2 \rangle \\ &= \mathbf{F}\langle \boldsymbol{d}\boldsymbol{d}^\dagger \rangle \mathbf{F}^\dagger - \mathbf{F}\langle \boldsymbol{d}\boldsymbol{s}^\dagger \rangle - \langle \boldsymbol{s}\boldsymbol{d}^\dagger \rangle \mathbf{F}^\dagger + \langle \boldsymbol{s}\boldsymbol{s}^\dagger \rangle. \end{aligned} \tag{A3}$$

As Kitaura & Enßlin (2008) pointed out it is important to note that the ensemble average $\langle \{\,\} \rangle$ goes over the galaxy and matter field realizations and the filter is thus different from the Wiener filter as derived in a Bayesian framework. We define here the global ensemble average by: $\langle \{\,\} \rangle = \langle \langle \{\,\} \rangle_{\rm g} \rangle_{\rm M}$. Here $\langle \{\,\} \rangle_{\rm g} \equiv \langle \{\,\} \rangle_{(N_{\rm g}^{\rm o}|\lambda^{\rm o})} \equiv \sum_{N_{\rm g}^{\rm o}=0}^{\infty} P_{\rm Pois}(N_{\rm g}^{\rm o} \mid w\lambda)\{\,\}$ denotes an ensemble average over the Poissonian distribution with the expected number of galaxy counts given by the Poissonian ensemble average: $\lambda^{\rm o} \equiv w\lambda \equiv \langle N_{\rm g}^{\rm o} \rangle_{\rm g}$, and $\langle \{\,\} \rangle_{\rm M} \equiv \langle \{\,\} \rangle_{(\delta_{\rm M}|\boldsymbol{p}_{\rm M})} \equiv \int {\rm d}\delta_{\rm M} P(\delta_{\rm M} \mid \boldsymbol{p}_{\rm M})$ being the ensemble average over all possible matter density realizations with some prior distribution $P(\delta_{\rm M} \mid \boldsymbol{p}_{\rm M})$ with $\boldsymbol{p}_{\rm M}$ being a set of parameters which determine the matter field, say the cosmological parameters. We impose $\langle \delta_{\rm M} \rangle_{\rm M} = 0$.

Recalling the derivations done by Wiener (1949); Rybicki & Press (1992); Zaroubi et al. (1995) we find minimizing the action with respect to the filter:

$$\frac{\partial \mathcal{A}}{\partial \mathbf{F}} = 0, \tag{A4}$$

the following LSQ filter expression:

$$\mathbf{F} = \langle \boldsymbol{s}\boldsymbol{d}^\dagger \rangle \langle \boldsymbol{d}\boldsymbol{d}^\dagger \rangle^{-1}. \tag{A5}$$

Traditionally one would then define a data degradation model with an additive noise term of the form: $\boldsymbol{d} = \mathbf{R}\boldsymbol{s} + \boldsymbol{\epsilon}$, with $\mathbf{R}$ being some response operator. Then substituting this data model in Eq. A5 and neglecting noise to signal correlation terms one would obtain a final expression for the LSQ filter (see Zaroubi et al. 1995).

A1 Signal to noise correlation

One can show that the noise is actually uncorrelated with the signal by making the following definition:

$$\epsilon_i^{\rm o} \equiv N_{{\rm g},i}^{\rm o} - \langle N_{{\rm g},i}^{\rm o} \rangle_{\rm g}, \tag{A6}$$

and then calculating the correlation:

$$\langle \epsilon_i^{\rm o} \langle N_{{\rm g},j}^{\rm o} \rangle_{\rm g} \rangle_{\rm g} = \langle N_{{\rm g},i}^{\rm o} \langle N_{{\rm g},j}^{\rm o} \rangle_{\rm g} - \langle N_{{\rm g},i}^{\rm o} \rangle_{\rm g} \langle N_{{\rm g},j}^{\rm o} \rangle_{\rm g} \rangle_{\rm g} = 0. \tag{A7}$$

Note, that this implies: $\langle \epsilon_i^{\rm o} \langle \delta_{{\rm g},j}^{\rm o} \rangle_{\rm g} \rangle_{\rm g} = 0$ and thus also $\langle \boldsymbol{\epsilon}\boldsymbol{s}^\dagger \rangle_{\rm g} = 0$.

A2 LSQ filter derivation without the additive noise assumption

However, one does not even need to use the additive noise assumption to derive the LSQ filter. Let us show here how to make such a derivation. We define the observed galaxy number counts per cell i as:

$$N^{\rm o}_{{\rm g},i} \equiv \overline{N}_{\rm g}(w_i + \delta^{\rm o}_{{\rm g},i}). \tag{A8}$$

The corresponding ensemble average over all possible galaxy realizations is:

$$\langle N^{\rm o}_{{\rm g},i}\rangle_{\rm g} \equiv \overline{N}_{\rm g} w_i (1+\delta_{{\rm g},i}). \tag{A9}$$

Recalling the linear bias relation:

$$\delta_{{\rm g},i} = \sum_j b_{i,j}\delta_j, \tag{A10}$$

we can then calculate with the above definitions the signal to data correlation matrix:

$$\begin{aligned}\langle s_i d_j\rangle &\equiv \langle \delta_{{\rm M},i}\delta^{\rm o}_{{\rm g},j}\rangle = \langle\langle \delta_{{\rm M},i}\delta^{\rm o}_{{\rm g},j}\rangle_{\rm g}\rangle_{\rm M} \\ &= \langle \delta_{{\rm M},i}\langle\delta^{\rm o}_{{\rm g},j}\rangle_{\rm g}\rangle_{\rm M} = w_j\sum_{j'} b_{j,j'}\langle\delta_{{\rm M},i}\delta_{{\rm M},j'}\rangle_{\rm M}. \end{aligned}\tag{A11}$$

We also have to calculate the data autocorrelation matrix:

$$\begin{aligned}\langle d_i d_j\rangle \equiv \langle \delta^{\rm o}_{{\rm g},i}\delta^{\rm o}_{{\rm g},j}\rangle &= \langle\langle\left(\frac{N^{\rm o}_{{\rm g},i}}{\overline{N}_{\rm g}} - w_i\right)\left(\frac{N^{\rm o}_{{\rm g},j}}{\overline{N}_{\rm g}} - w_j\right)\rangle_{\rm g}\rangle_{\rm M} \\ &= \frac{\langle\langle N^{\rm o}_{{\rm g},i}N^{\rm o}_{{\rm g},j}\rangle_{\rm g}\rangle_{\rm M}}{\overline{N}^2_{\rm g}} - w_i w_j. \end{aligned}\tag{A12}$$

Here we need a model for the two-point number count statistics. Note, that we can introduce here Poissonity:

$$\langle N^{\rm o}_{{\rm g},i}N^{\rm o}_{{\rm g},j}\rangle_{\rm g} \equiv \langle N^{\rm o}_{{\rm g},i}\rangle_{\rm g}\langle N^{\rm o}_{{\rm g},j}\rangle_{\rm g} + \langle N^{\rm o}_{{\rm g},i}\rangle_{\rm g}\delta^{\rm K}_{i,j}. \tag{A13}$$

With the additional matter field ensemble average we get:

$$\langle\langle N^{\rm o}_{{\rm g},i}\rangle_{\rm g}\langle N^{\rm o}_{{\rm g},j}\rangle_{\rm g}\rangle_{\rm M} = \overline{N}^2_{\rm g} w_i w_j\left(1 + \sum_k b_{i,k}\sum_l b_{j,l}\langle\delta_{{\rm M},k}\delta_{{\rm M},l}\rangle_\delta\right). \tag{A14}$$

We can define the noise covariance matrix as:

$$\begin{aligned}N_{i,j} &\equiv \frac{1}{\overline{N}^2_{\rm g}}\langle\langle N^{\rm o}_{{\rm g},i}N^{\rm o}_{{\rm g},j}\rangle_{\rm g} - \langle N^{\rm o}_{{\rm g},i}\rangle_{\rm g}\langle N^{\rm o}_{{\rm g},j}\rangle_{\rm g}\rangle_{\rm M} \\ &= \frac{1}{\overline{N}^2_{\rm g}}\langle\langle N^{\rm o}_{{\rm g},i}\rangle_{\rm g}\rangle_{\rm M}\delta^{\rm K}_{i,j} = \frac{w_i}{\overline{N}_{\rm g}}\delta^{\rm K}_{i,j}. \end{aligned}\tag{A15}$$

The LSQ filter can be then written as:

$$\begin{aligned}F_{i,j} = &\sum_{j'} w_{j'}\sum_l b_{j',l}\langle\delta_{{\rm M},i}\delta_{{\rm M},l}\rangle_{\rm M} \\ &\times\left(w_{j'}\sum_k b_{j',k}\sum_{k'} b_{j,k'}\langle\delta_{{\rm M},k}\delta_{{\rm M},k'}\rangle_{\rm M} w_j + \frac{w_j}{\overline{N}_{\rm g}}\delta^{\rm K}_{j',j}\right)^{-1}.\end{aligned}\tag{A16}$$

The corresponding matrix notation of the LSQ filter yields:

$$\mathbf{F} = \mathbf{S}\mathbf{R}^\dagger\left(\mathbf{R}\mathbf{S}\mathbf{R}^\dagger + \mathbf{N}\right)^{-1}, \tag{A17}$$

with $\mathbf{S} \equiv \langle\boldsymbol{\delta}_{\rm M}\boldsymbol{\delta}^\dagger_{\rm M}\rangle_{\rm M}$ and $\mathbf{R} \equiv \mathbf{W}\mathbf{B}$ (see section 4.1.1 for a definition of the bias operator $\mathbf{B}$). This data-space expression is equivalent to the signal-space representation (for a demonstration see appendix C in Kitaura & Enßlin 2008):

$$\mathbf{F} = \left(\mathbf{S}^{-1} + \mathbf{R}^\dagger\mathbf{N}^{-1}\mathbf{R}\right)^{-1}\mathbf{R}^\dagger\mathbf{N}^{-1}. \tag{A18}$$

We conclude that the LSQ filter is the optimal linear filter under a Poisson noise assumption. We have shown that this filter does not neglect any signal to noise correlation.

APPENDIX B: LOGNORMAL PRIOR AND GAUSSIAN LIKELIHOOD

For completeness we derive a nonlinear filter which assumes a lognormal prior and a Gaussian likelihood. Following Sheth (1995) one could use the Wiener filter with a data transformation and apply it to recover non-Gaussian distributed fields. The problem in such a model is that one requires a multiplicative noise assumption of the form:

$$\begin{aligned} d'_i &\equiv \delta_i\epsilon'_i + \epsilon'_i \\ \ln(d'_i) &= \ln(1+\delta_i) + \ln(\epsilon'_i) \\ d_i &\equiv s_i + \epsilon_i, \end{aligned}\tag{B1}$$

for each cell i, with $d_i \equiv \ln(d'_i)$, $s_i \equiv \ln(1+\delta_i)$ and $\epsilon_i \equiv \ln(\epsilon'_i)$. Note, that with such a data model one could easily apply the Wiener filter assuming that the signal s and the noise ϵ are Gaussian distributed.

B1 Additive noise model

However, one may rather prefer a data model with an additive noise term as commonly used in the literature (see e.g. Zaroubi et al. 1995; Tegmark 1997). We define therefore a data model of the form:

$$\boldsymbol{d} \equiv \mathbf{R}\boldsymbol{\delta} + \boldsymbol{\epsilon}, \tag{B2}$$

including in the signal higher order terms:

$$\boldsymbol{\delta} = \exp(\boldsymbol{s}) - \vec{1}. \tag{B3}$$

One can then assume the signal to be lognormal distributed and the noise to be Gaussian distributed and signal-independent.

B2 Gaussian likelihood

Let us write the log-likelihood as:

$$\ln\mathcal{L} \propto -\frac{1}{2}\left(\boldsymbol{\epsilon}^\dagger\mathbf{N}^{-1}\boldsymbol{\epsilon} - \ln\left(\det\left(\mathbf{N}\right)\right)\right) + c. \tag{B4}$$

Making the substitution $\boldsymbol{\epsilon} = \boldsymbol{d} - \mathbf{R}\boldsymbol{\delta}$ we get:

$$\boldsymbol{\epsilon}^\dagger\mathbf{N}^{-1}\boldsymbol{\epsilon} = \boldsymbol{d}^\dagger\mathbf{N}^{-1}\boldsymbol{d} + \boldsymbol{\delta}^\dagger\mathbf{R}^\dagger\mathbf{N}^{-1}\mathbf{R}\boldsymbol{\delta} - \boldsymbol{\delta}^\dagger\mathbf{R}^\dagger\mathbf{N}^{-1} - \boldsymbol{d}^\dagger\mathbf{N}^{-1}\mathbf{R}\boldsymbol{\delta}. \tag{B5}$$

To find the maximum a posteriori solution we have to calculate the derivative of the likelihood with respect to the signal $\boldsymbol{s}$:

$$\sum_i\frac{\partial\ln\mathcal{L}_i}{\partial s_k} = \sum_i\sum_l\frac{\partial\ln\mathcal{L}_i}{\partial\delta_l}\frac{\partial\delta_l}{\partial s_k}. \tag{B6}$$

From Eq. B3 we get:

$$\frac{\partial\delta_l}{\partial s_k} = \exp(\boldsymbol{s}_i)\delta^{\rm K}_{i,k}. \tag{B7}$$

Assuming a signal independent noise yields:

$$\frac{\partial\ln\mathcal{L}_i}{\partial\delta_k} = -\sum_{jlm}\delta_j R_{j,l}N^{-1}_{l,m}R_{m,i} + \sum_{jl} d_j N^{-1}_{j,l}R_{l,i}. \tag{B8}$$

Combining these results with the derivative of the lognormal prior (Eq. 19) leads to:

$$\sum_j S^{-1}_{\mathrm{L}j,k}\,(s_j - \mu_j) = \tag{B9}$$
$$\left(\sum_{jlm}\left(\exp(s_j) - \vec{1}\right) R_{j,l} N^{-1}_{l,m} R_{m,k} + \sum_{jl} d_j N^{-1}_{j,l} R_{l,k}\right) \exp(s_k).$$