Bayesian weak lensing tomography: Reconstructing the 3D large-scale distribution of matter with a lognormal prior

Vanessa Böhm,[1] Stefan Hilbert,[2,3] Maksim Greiner,[1,2] and Torsten A. Enßlin[1,3,2]

[1]*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Strasse 1, 85748 Garching, Germany*
[2]*Exzellenzcluster Universe, Boltzmannstrasse 2, 85748 Garching, Germany*
[3]*Ludwig-Maximilians-Universität, Universitäts-Sternwarte, Scheinerstrasse 1, 81679 München, Germany*

We present a Bayesian reconstruction algorithm that infers the three-dimensional large-scale matter distribution from the weak gravitational lensing effects measured in the image shapes of galaxies. The algorithm is designed to also work with non-Gaussian posterior distributions which arise, for example, from a non-Gaussian prior distribution. In this work, we use a lognormal prior and compare the reconstruction results to a Gaussian prior in a suite of increasingly realistic tests on mock data. We find that in cases of high noise levels (i.e. for low source galaxy densities and/or high shape measurement uncertainties), both normal and lognormal priors lead to reconstructions of comparable quality, but with the lognormal reconstruction being prone to mass-sheet degeneracy. In the low-noise regime and on small scales, the lognormal model produces better reconstructions than the normal model: The lognormal model (1) enforces non-negative densities, while negative densities are present when a normal prior is employed, (2) better traces the extremal values and the skewness of the true underlying distribution, and (3) yields a higher pixel-wise correlation between the reconstruction and the true density.

I. INTRODUCTION

Weak gravitational lensing of galaxies offers a unique way to study the distribution of matter in the Universe (see [1] for a review on weak gravitational lensing and [2–4] for recent reviews on weak galaxy lensing). Lensing by structures along the line of sight causes distortions in the images of distant galaxies (which in this context are often referred to as sources), which leads to correlations between the apparent shapes of these galaxies. The dominant and most easily detectable image distortion that lensing induces is a shearing of the galaxy images. Because of this, the effect is often referred to as cosmic shear.

Galaxy shape measurements allow to constrain the clustering of matter at different scales and redshifts, which can then be translated into constraints on cosmological models and their parameters. The integrated lensing signal is mostly sensitive to a combination of the cosmic mean matter density Ω_m and the matter power spectrum amplitude σ_8 [5]. Tomographic methods can yield additional constraints on all parameters guiding the growth rate of structure [6,7], notably also on the properties of dark energy [8–10], but rely on low photometric redshift uncertainties [11]. Since lensing is a direct probe of the total matter, luminous and dark, it can be combined with measurements of the luminous matter distribution in order to learn about the relationship between baryons and dark matter. Another important feature of weak galaxy lensing lies in its ability to probe the matter distribution over a wide range of scales, from many tens of Mega parsec, where structure formation is still linear today and comparably easy to model, down to nonlinear sub-Mega parsec scales. Due to its sensitivity to such a wide range of scales, lensing can provide a large amount of information to constrain models of nonlinear structure formation and cosmology, in particular if also higher-order statistics are considered.

First firm statistical detections of cosmic shear were reported in 2000 by four different groups [12–15]. Since then, the field has seen a tremendous increase in the amount and quality of lensing data [e.g., [16–20]] as well as a notable improvement in analysis techniques [e.g. [21–30]]. Estimates of cosmological parameters have been inferred from these data by comparing the power spectrum of the fully projected 2D shear field (or related quantities) as well as auto- and cross-spectra in a number of redshift bins to theoretical predictions [27,31–39]. Several authors have also investigated the additional constraining power that can be achieved by incorporating third-order statistics and/or shear peak counts and correlations [5,40–48], which helps to break parameter degeneracies.

The full information content, however, lies in the three-dimensional nonlinear shear field. 3D weak shear analysis methods have been proposed by a number of authors [49–51] and were recently applied to data from the Canada France Hawaii Telescope Lensing Survey (CFHTLenS) [52]. Furthermore, the measured shear field can be used for 3D reconstructions of the underlying density field. This can then be directly compared to models of structure formation and simplifies the cross correlation with other tracers of matter. Algorithms that invert the lens equation to obtain the density have been worked out by a number of authors [53–55]. Since this inversion is underconstrained,

123510-1

it requires some regularization method or choice of prior on the density field. Most of the algorithms employ a Wiener filter, which corresponds to a normal (Gaussian) prior, that can be complemented with information about galaxy clustering [56,57].

We aim to extend the work on tomographic reconstruction of the 3D matter distribution by designing a fully Bayesian reconstruction algorithm that uses a lognormal prior on the density field. The algorithm is designed to reconstruct the 3D cosmic density fluctuation field $\delta(\mathbf{x}, \tau)$ from weak galaxy lensing data, i.e. a measurement of galaxy ellipticities at different (photometrically measured) redshifts. Its derivation is based on the language of information field theory [58], which has already been used to address similar tomography problems [59]. We do not make use of the flat-sky approximation i.e. lines of sight are allowed to be nonperpendicular to a fixed 3D coordinate grid. Further, we do not bin the data into pixels but take each galaxy into account as an individual contribution to the likelihood. This allows us to incorporate distance uncertainties of individual galaxies instead of sample redshift distributions.

In contrast to a normal prior for the density field, as it has often been assumed before, a lognormal prior automatically enforces the strict positivity of the field and allows to capture some of the non-Gaussian features that are imprinted on the density distribution by nonlinear structure formation. Hubble was the first to notice that galaxy number counts could be well approximated by a lognormal distribution [60]. Characterizing the matter overdensities in the Universe as a lognormal field was first assessed by Coles and Jones in 1991 [61]. Subsequent studies showed that a logarithmic mapping of the nonlinear matter distribution can partly re-Gaussianize the field, and that nonlinear features in the matter power spectrum can be reproduced by a lognormal transformation of the linear matter power spectrum [62,63]. A lognormal prior has already been used and shown to be superior to a Gaussian one in Bayesian algorithms that reconstruct the large-scale matter distribution from the observed galaxy distribution [58,64,65]. Lognormal distributions have also already been considered in the context of weak lensing: Analyses of ray-tracing simulations and the Dark Energy Survey (DES) Science Verification data showed that the 1-point distribution function of the lensing convergence is better described by a lognormal than a Gaussian model [66,67]. Also the cosmic shear covariance can be modeled to better accuracy under the assumption that the underlying convergence field follows a lognormal distribution instead of a Gaussian one [68].

Bayesian inference methods are widely used in weak shear analyses, most prominently in the context of shear measurements from galaxy images [21,26,69]. Recently, notable effort has been put into developing a fully Bayesian analysis pipeline that propagates all uncertainties

consistently from the raw image to the inferred cosmological parameters [70,71].

This paper is organized as follows: This introduction is followed by a short section, Sec. II, in which we briefly introduce the notations and coordinate systems that will be used in the derivation of the formalism. Our lognormal prior model for the density is described in detail in Sec. III. In Sec. IV, we present the data model, i.e. the lensing formalism that connects the data from a cosmic shear measurement to the underlying density field and give a brief overview over its implementation in Sec. V. The maximum *a posteriori* estimator that is used to infer the matter distribution is introduced in Sec. VI and extended to include redshift uncertainties of individual sources in Sec. VII. In Sec. VIII, we show results of the density reconstruction on increasingly realistic mock data. We conclude this work with a summary and discussion in Sec. IX.

II. COORDINATE SYSTEMS AND NOTATIONAL CONVENTIONS

In the derivation of the formalism we work with three different types of coordinate systems. First, we use three-dimensional purely spatial comoving coordinates $\mathbf{x} = (x_0, x_1, x_2)$ at fixed comoving lookback time τ, that, combined with the time coordinate, form the four-dimensional coordinate system $(\mathbf{x}, \tau)$.

Second, we use a 3D comoving coordinate system on the light cone of an observer at the origin. Vectors on the light cone are marked by a prime, e.g. $\mathbf{x}'$. Since $\mathbf{x}'$-coordinates implicitly define a comoving lookback time $\tau(\mathbf{x}') = |\mathbf{x}'|/c$ (where c denotes the speed of light), we omit spelling out the time explicitly and write $A(\mathbf{x}') = A(\mathbf{x}', \tau)$ for any quantity A that is defined on the light cone. The operation that links quantities on the light cone to their corresponding quantities in 4D spacetime can be encoded in a projection operator with kernel

$$\mathcal{C}_{\mathbf{x}'(\mathbf{x}, \tau)} = \delta_{\mathrm{D}}(\mathbf{x}' - \mathbf{x})\delta_{\mathrm{D}}(\tau - |\mathbf{x}|/c). \qquad (1)$$

Third, we employ a set of coordinate systems on the light cone, in which each system is orientated such that one axis points into the direction of a source galaxy. These line of sight coordinate systems are centered on the observer and spanned by the vectors $(\hat{\mathbf{r}}_0^i, \hat{\mathbf{r}}_1^i, \hat{\mathbf{r}}_2^i)$, where $\hat{\mathbf{r}}_0^i(\mathbf{x}')$ points into the direction of the ith source galaxy and the normal vectors $(\hat{\mathbf{r}}_1^i, \hat{\mathbf{r}}_2^i)$ span the two-dimensional plane perpendicular to $\hat{\mathbf{r}}_0^i$. The radial comoving distance of each galaxy from the observer is denoted $r^i = |\mathbf{r}^i| \equiv |\mathbf{r}_0^i|$. The transformation from the light cone system into the line of sight (LOS) system of a source galaxy i is achieved by a rotation of the x_0'-axis into the ith line of sight, while x_1' and x_2' get aligned with the image coordinates of the galaxy observation. We label the corresponding transformation operators $\mathcal{R}^i_{\mathbf{r}\mathbf{x}'}$.

III. PRIOR MODEL

The comoving density $\rho(\mathbf{x}, \tau)$ can be split into its time-independent spatial mean $\bar{\rho}$, and a perturbation $\delta\rho(\mathbf{x}, \tau) = \bar{\rho}\delta(\mathbf{x}, \tau)$. The fractional overdensity $\delta(\mathbf{x}, \tau)$ is commonly modeled as a homogeneous, isotropic Gaussian field with zero mean and power spectrum $P_\delta(k, \tau)$. This is an excellent description at early times where fluctuations are very small, as e.g., shown by observations of the cosmic microwave background radiation (CMB). At linear level, valid for $\delta \ll 1$, the time evolution of the density field can be described by

$$\rho(\mathbf{x}, \tau) = \bar{\rho}[1 + \delta(\mathbf{x}, \tau)]$$

$$= \bar{\rho}\left[1 + \int_{\mathbf{y}} D(|\mathbf{x} - \mathbf{y}|, \tau, \tau_0)\varphi(\mathbf{y}, \tau_0)\right]$$

$$= \bar{\rho}[1 + D_{\mathbf{xy}}(\tau, \tau_0)\varphi_{\mathbf{y}}(\tau_0)]. \qquad (2)$$

In the last expression we have introduced a shorthand notation that will be used in the rest of the paper: repeated indices are integrated over if they do not appear on both sides of the equation. In Eq. (2), $D(|\mathbf{x} - \mathbf{y}|, \tau, \tau_0)$ is the integration kernel of a linear homogeneous and isotropic, but possibly scale-dependent, growth operator. The field φ is an isotropic and homogeneous random field whose values are drawn from a multivariate normal distribution with mean $\bar{\varphi}$ and covariance Φ,

$$\varphi \hookleftarrow \mathcal{N}(\varphi|\bar{\varphi}, \Phi). \qquad (3)$$

Here, it describes the three dimensional density fluctuations at time τ_0 and is translated to other times τ by the growth operation $D_{\mathbf{xy}}(\tau, \tau_0)$. This implies $\bar{\varphi} = 0$ for this linear Gaussian model.

The description Eq. (2) breaks down when $\delta \ll 1$ such that nonlinearities become important. A possible way to account for nonlinearities is to include higher-order terms from a perturbation series expansion of the full nonlinear evolution equations. However, a further shortcoming of the model (2) is that it allows arbitrarily negative density contrasts, which physically can not be smaller than -1. To obtain a strictly positive density field, we instead modify Eq. (2) by an exponential:

$$\rho(\mathbf{x}, \tau) = \bar{\rho}[1 + \delta(\mathbf{x}, \tau)] = \bar{\rho}\exp[D_{\mathbf{xy}}(\tau, \tau_0)\varphi_{\mathbf{y}}(\tau_0)]. \qquad (4)$$

Since the expectation value of the density $\rho(\mathbf{x}, \tau)$ must equal $\bar{\rho}$, the mean of φ must be set to

$$\bar{\varphi}_{\mathbf{x}}(\tau) = -\frac{1}{2}D_{\mathbf{xy}}^{-1}(\tau, \tau_0)[D(\tau, \tau_0)\Phi(\tau_0)D(\tau, \tau_0)]_{\mathbf{yy}}, \quad (5)$$

for every time τ. Note that we integrate over the index $\mathbf{y}$ in this expression, i.e. the diagonal of the composite operator in square brackets is treated as a field.

For a local growth operator, $D_{\mathbf{xy}}(\tau, \tau_0) = D(\tau, \tau_0)\delta_{\mathrm{D}}(\mathbf{x} - \mathbf{y})$, this mean correction simplifies to

$$\bar{\varphi}_{\mathbf{x}}(\tau) = -\frac{1}{2}D(\tau, \tau_0)\hat{\Phi}_{\mathbf{x}}(\tau_0), \qquad (6)$$

where we defined $\hat{\Phi}_{\mathbf{x}}(\tau_0) \equiv \Phi_{\mathbf{xx}}(\tau_0)$.

The Gaussian field φ and the growth operator D can be related to known quantities. To see this, consider the expansion of the Fourier modes of δ in Eulerian perturbation theory (see e.g. [72]),

$$\delta(\mathbf{k}, \tau) = \sum_{n=1} D^{(n)}(\tau)\delta^{(n)}(\mathbf{k}), \qquad (7)$$

where $\delta^{(n)}$ are convolutions of n initial fields $\delta(\mathbf{k}, \tau_0)$ with an integration kernel that changes from order to order. The first term in this series is $D^{(1)}(\tau)/D^{(1)}(\tau_0)\delta_0(\mathbf{k}, \tau_0)$, where $D^{(1)}(\tau)$ is the growing solution to the linearized growth equation [73].

We use this analogy to motivate the simplest possible form of the growth operation in the lognormal model and write

$$D_{\mathbf{xy}}(\tau) = \delta_{\mathrm{D}}(\mathbf{x} - \mathbf{y})D^{(1)}(\tau) \qquad (8)$$

where we have set $D^{(1)}(\tau_0) = 1$. This approximation erases any *a priori* assumption of scale-dependent growth and mode-coupling of the log field φ. Such a simplification is viable since the model in Eq. (8) describes only our prior assumptions about the density field ρ. The algorithm will find the most probable realization of φ for a fixed growth operator D given the data. If a scale-dependence is favored by the data, it will be recovered, at least partially, in the estimate of φ.

Our algorithm also allows to incorporate a more general growth operation at the expense of computation time and memory usage. The application of the most general $D_{\mathbf{xy}}(\tau)$ generates a four-dimensional field: three-dimensional spatial comoving volumes for every time-slice τ. This very large volume is then restricted to a three-dimensional cut by application of the light cone operator [Eq. (1)]. The prior model for the matter overdensity on the light cone then becomes

$$\delta_{\mathbf{x}'} = \mathcal{C}_{\mathbf{x}'(\mathbf{x}, \tau)}\exp[D_{\mathbf{xy}}(\tau)\varphi_{\mathbf{y}}(\tau)] - 1. \qquad (9)$$

For the purpose of comparing the lognormal to the Gaussian density model we choose the simplest form of D, Eq. (8), since it can be applied to the three-dimensional field on the light cone

$$\delta_{\mathbf{x}'} = \exp[D_{\mathbf{x}'\mathbf{y}'}\varphi_{\mathbf{y}'}] - 1. \qquad (10)$$

More complicated models, e.g. motivated by perturbation theories for large-scale structure, could be envisioned but

lie beyond the scope of this work. A growth model that is better informed about the spatial correlations in the density field could potentially improve the reconstruction in regions with low signal to noise (see also the discussion in Sec. IX.

IV. DATA MODEL

In this section we establish the analytic relation between the signal field φ, the field of overdensities δ that we aim to reconstruct, and the data that is obtained from a weak lensing measurement.

Weak galaxy lensing surveys produce galaxy image ellipticities that can be quantified, e.g., by a complex number $\epsilon = \epsilon_1 + i\epsilon_2$ such that $\epsilon = (a - b)/(a + b) \exp(2i\eta)$ for an ellipse with major axis a, minor axis b, and position angle η [e.g. [74]]. We use the common approximation that the components of the intrinsic source galaxy ellipticity $\boldsymbol{\epsilon}^{s} = (\epsilon_1^{s}, \epsilon_2^{s})$, which define the shape that would be observed in the absence of lensing,[1] follow a global bivariate Gaussian distribution with zero mean and variance σ_ϵ^2 per component:

$$\boldsymbol{\epsilon}^{s} \hookleftarrow \mathcal{N}(\boldsymbol{\epsilon}^{s}|0, N^{s}), \qquad N_{ij}^{s} = \delta_{ij}\sigma_\epsilon^2. \qquad (11)$$

This approximation has shortcomings (see e.g. [23]), but serves for the proof of our concept, since we create the mock data on which we test the algorithm with exactly this shape noise model. In the future, more elaborated (Bayesian hierarchical) shear estimators, that e.g. take into account galaxy properties, can be incorporated into the algorithm [21,26,69].

Lensing distorts the galaxy images in shape and size [74–78]. If the distortion is small, i.e. in the limit of weak lensing, the relation between intrinsic source ellipticity and observed ellipticity can be linearized and simplifies to [79,80]

$$\epsilon = g + \epsilon^{s}, \qquad (12)$$

where g is the reduced shear. The reduced shear combines the effect of anisotropic lensing distortions, encoded in the shear $\gamma = \gamma_1 + i\gamma_2$, and the isotropic distortion, encoded in the convergence κ

$$g = \frac{\gamma}{1 - \kappa} \approx \gamma. \qquad (13)$$

If $\kappa \ll 1$, which is often the case for galaxy lensing, the reduced shear can be approximated by the shear itself $g \approx \gamma$.

The shear and convergence at angular position $\boldsymbol{\theta}$ are related to the lensing potential ψ by

---

[1]We use vector notation, e.g. $\boldsymbol{\epsilon}$, to denote the tuple of real and imaginary part (ϵ_1, ϵ_2) of a complex number $\epsilon = \epsilon_1 + i\epsilon_2$.

$$\gamma_1(\boldsymbol{\theta}) = \frac{1}{2}(\partial_1^2 - \partial_2^2)\psi(\boldsymbol{\theta}); \qquad \gamma_2(\boldsymbol{\theta}) = \partial_1\partial_2\psi(\boldsymbol{\theta});$$

$$\kappa(\boldsymbol{\theta}) = \frac{1}{2}(\partial_1^2 + \partial_2^2)\psi(\boldsymbol{\theta}). \qquad (14)$$

The lensing potential ψ is an integrated measure of scalar perturbations to the background metric along the perturbed photon geodesic. Integrating the perturbations along the unperturbed geodesic turns out to be an excellent approximation [80,81]. Working in this so-called Born approximation, choosing the Newtonian gauge [82] and assuming no anisotropic stress, the lensing potential can be written as a weighted projection of the peculiar Newtonian gravitational potential ϕ along the line of sight. For a source at LOS distance r^i, this integration reads

$$\psi(\boldsymbol{\theta}) = \frac{2}{c^2} \int_0^{r^i} dr \frac{r^i - r}{rr^i} \phi(r, r\theta_1, r\theta_2), \qquad (15)$$

where we have assumed a spatially flat Universe. Applying the angular derivatives in Eq. (14) to the expression for the lensing potential in Eq. (15), we get

$$\partial_k\partial_l\psi(\boldsymbol{\theta}) = \frac{2}{c^2} \int_0^{r^i} dr W(r; r^i)\partial_{r_k}\partial_{r_l}\phi(r, r\theta_1, r\theta_2), \qquad (16)$$

where $k, l \in (1, 2)$, and the lensing efficiency

$$W(r; r^i) = \frac{r(r^i - r)}{r^i}. \qquad (17)$$

In practice the distance to the source r^i cannot be determined directly but follows from the photometrically measured redshift z^i. Photometrically measured redshifts are associated with a relatively high error, $\sigma_z/(1 + z) \approx 0.04$–$0.06$ [30]. In its most simple form the algorithm ignores this uncertainty. We will use this simplified model to validate the functionality of the algorithm in terms of reconstructing non-linear structures in the lognormal approximation. Redshift uncertainties will be included later in Sec. VII.

The lensing shear is completely determined by the second derivatives of the lensing potential perpendicular to the LOS. The tidal tensor $\partial_{r_k}\partial_{r_l}\phi(\mathbf{r})$ along the LOS of the ith source galaxy is obtained by rotating the tidal tensor on the global coordinate grid $\mathbf{x}'$

$$T_{ij}(\mathbf{x}') = \partial_{x_i'}\partial_{x_j'}\phi(\mathbf{x}'), \qquad i, j \in (0, 1, 2), \qquad (18)$$

into the specific coordinate system (with coordinates $\mathbf{r}^i$) that points into the direction of this ith galaxy and projecting it onto the $(r_1^i - r_2^i)$-plane perpendicular to the LOS.

The last relation required to connect the data to the density fluctuations is Poisson's equation. It relates the potential $\phi(\mathbf{x}')$ to the density fluctuations $\delta(\mathbf{x}')$,

$$\nabla^2 \phi(\mathbf{x}') = \frac{3}{2}\Omega_{\mathrm{m}}H_0^2 \frac{\delta(\mathbf{x}')}{a(|\mathbf{x}'|/\mathrm{c})}, \qquad (19)$$

where H_0 denotes the Hubble constant (which will be parametrized by $h = H_0/(100\ \mathrm{km\,s^{-1}\,Mpc^{-1}})$ in our test simulations), and $a(|\mathbf{x}'|/\mathrm{c}) = a(\tau)$ denotes the scale factor at the time τ corresponding to LOS distance $r^i(x')$.

V. IMPLEMENTATION

The implementation, not only of the data model, but of the entire algorithm, is based on NIFTy [83], a versatile software package for the development of inference algorithms. We further compute cosmology-dependent quantities, like power spectra and distance-redshift relations, with the publicly available CLASS code [84]. To summarize the data model we introduce shorthand notations for each operation in terms of operators.

In its most general form, the prior and data model, that connect the Gaussian field φ with a data vector of N_s measured source ellipticities, are as follows: The growth operator $D_{\mathbf{xy}}(\tau)$ imprints a growth structure on the Gaussian field. The resulting four-dimensional field is plugged into the exponential of $E(\cdot) = \exp(\cdot) - 1$ [see Eq. (4)] to obtain the fractional overdensity $\delta_{\mathbf{x}}(\tau)$. The overdensity induces the potential $\phi_{\mathbf{y}}(\tau)$ by the Poisson equation, encoded in the operator $\mathcal{P}_{\mathbf{yx}}(\tau) = \Delta^{-1}\frac{3}{2}\Omega_{\mathrm{m}}H_0^2 \delta(\mathbf{x}')/a(|\mathbf{x}'|/\mathrm{c})$ [Eq. (19)]. The potential can be computed efficiently in Fourier space. The gravitational potential is then restricted to the light cone of the observer by the light cone operator $\mathcal{C}_{\mathbf{y}'(\mathbf{y},\tau)}$ [Eq. (1)]. We compute the tidal tensor of the resulting 3D field [(Eq. (18)] by application of a global differential operator, which we denote $\mathcal{T}_{\mathbf{z}'\mathbf{y}'}$. The resulting tidal tensor is then rotated into each galaxy's LOS coordinate system by a rotation operator, $\mathcal{R}^i_{\mathbf{z}'}$. An integration operator $\mathcal{I}_{j\mathbf{r}}$, which applies the integration in Eq. (16), integrates the components of each of the resulting N_s tidal tensors along the unperturbed photon geodesic. For this operator, we use an adapted version of the implementation that was already successfully used in a similar reconstruction method [59]. The application of $\mathcal{I}_{j\mathbf{r}}$ yields derivatives of the lensing potential for each galaxy location. From this we can compute the shear components by a linear operator $\mathcal{L}_{ij}$ that comprises the equations in Eq. (14). Rotation and integration map the three dimensional continuous signal space into the discrete space (one point for every galaxy) of the data. The shear components are thus automatically computed at the locations of the galaxies.

In the simplified implementation that we use for this work, we avoid the 4D coordinate grid $(\mathbf{x}, \tau)$ and work on the three-dimensional light cone from the beginning. In the prior, we model the Gaussian log-density φ with the power spectrum of matter fluctuations today $P_{\mathrm{lin}}(k, a = 1)$, where a denotes the scale factor. The growth operator is diagonal

in configuration space and only a function of comoving distance to the observer $D_{\mathbf{x}'\mathbf{y}'} = D^{(1)}(\tau)\delta_{\mathrm{D}}(\mathbf{x}' - \mathbf{y}')$, where $D^{(1)}(\tau)$ is the growing solution of the linearized growth equation. The Poisson operator is split into two parts. First, a multiplication with $3/2\Omega_{\mathrm{m}}H_0^2/a(|\mathbf{x}'|/\mathrm{c})$, i.e. an operation that is diagonal in configuration space. Second, the inverse Laplace operation Δ^{-1}, which is diagonal in Fourier space. The inverse Laplacian in the Poisson equation is a nonlocal operation that should strictly be applied to 3D spatial volumes at fixed time. Here, we apply it on the light cone noting that the induced error in radial direction should be small (roughly of order a^2) if we apply the first part of the Poisson operator first and if $D^{(1)}(\tau)$ is approximately proportional to the scale factor. In this case, the first order term of the exponential expression

$$\exp(D_{\mathbf{x}'\mathbf{y}'}\varphi_{\mathbf{y}'}) - 1 = D_{\mathbf{x}'\mathbf{y}'}\varphi_{\mathbf{y}'} + \mathcal{O}(\varphi^2)$$
$$\approx a(|\mathbf{x}'|)\delta_{\mathrm{D}}(\mathbf{x}' - \mathbf{y}')\varphi_{\mathbf{y}'} + \mathcal{O}(\varphi^2), \quad (20)$$

and therefore the first order time dependence of the overdensity field will partly be canceled by the $1/a$ in the Poisson equation before the inverse Laplacian rescales the field. This cancellation corresponds to the commonly known fact that the comoving gravitational potential is constant in a matter dominated Universe.

After computing the six independent entries of the tidal shear tensor, we integrate each of its components along the LOS. Only after the integration we rotate the resulting tensors into their LOS coordinate systems and project out all entries which are not in the plane perpendicular to $\mathbf{r}$. This change of order allows to efficiently combine three operations, that is the rotation into the LOS, projecting out nonperpendicular components of the tidal tensor, and the computation of shear components and convergence, in a single linear operation. We denote the corresponding operator $\mathcal{G}$ (gamma-projection-rotation).

For one data point the implemented data model in operator notation reads

$$d_i = \boldsymbol{\epsilon}_i = (R(\varphi))_i + \boldsymbol{\epsilon}^s_i + \mathbf{n}_i$$
$$= \mathcal{G}_{ij}\mathcal{I}_{j\mathbf{r}'}\mathcal{T}_{\mathbf{r}'\mathbf{z}'}\mathcal{P}_{\mathbf{z}'\mathbf{y}'}[\exp(D_{\mathbf{y}'\mathbf{x}'}\varphi_{\mathbf{x}'}) - 1_{\mathbf{y}'}] + \boldsymbol{\epsilon}^s_i + \mathbf{n}_i.$$
$$(21)$$

The total data vector d has dimensions $2 \times N_s$, i.e. two ellipticity components for each of the N_s source galaxies. We use the letter R to encode the total response of the data to a signal φ, i.e. the composite action of all operators. We have also added an experimental noise $\mathbf{n}$ here for completeness. In general this will be subdominant to the shape noise $\boldsymbol{\epsilon}^s$ and we ignore it in the following. Note, however, that the formalism allows for an incorporation of several independent noise sources.

VI. MAP ESTIMATOR

Our aim is to obtain a maximum *a posteriori* (MAP) estimate of the signal field φ. The posterior distribution is related to the likelihood and the prior by Bayes' theorem

$$P(\varphi|d) = \frac{P(d|\varphi)P(\varphi)}{P(d)}. \quad (22)$$

The prior probability $P(\varphi)$ is modeled as a Gaussian distribution with covariance Φ. In a first simple approximation Φ can be taken to be diagonal in Fourier space with the usual power spectrum of matter overdensities $P_{\mathrm{lin}}(k, a)$.

To obtain the likelihood, we marginalize over all realizations of the shape noise

$$P(d|\varphi) = \int \mathcal{D}\boldsymbol{\epsilon}^s P(d|\varphi, \boldsymbol{\epsilon}^s)P(\boldsymbol{\epsilon}^s) = \mathcal{N}[d|R(\varphi), N], \quad (23)$$

where the covariance N of the shape noise was defined in Eq. (11) and we neglect any other sources of measurement noise. With this, the negative log-posterior becomes

$$-\ln P(\varphi|d) \hat{=} \frac{1}{2}[d - R(\varphi)]^{\dagger} N^{-1}[d - R(\varphi)]$$
$$+ \frac{1}{2}(\varphi - \bar{\varphi})^{\dagger}\Phi^{-1}(\varphi - \bar{\varphi}), \quad (24)$$

where we have dropped most terms that are independent of the field of interest, φ.

The maximum of the posterior distribution is found by minimizing the expression in Eq. (24). Note that the posterior distribution is not Gaussian. Due to the exponential in the response R it is not quadratic in φ. To find the minimum of the negative log-posterior we apply a Newton-like minimization scheme [85]. This requires the derivative of the negative log-posterior with respect to φ

$$-\frac{\delta \ln P(\varphi|d)}{\delta \varphi_{\mathbf{u}}} = \Phi_{\mathbf{uq}}^{-1}(\varphi - \bar{\varphi})_{\mathbf{q}}$$
$$+ [d - R(\varphi)]_i N_{ij}^{-1}\mathcal{G}_{jk}\mathcal{I}_{k\mathbf{r}'}\mathcal{T}_{\mathbf{r}'\mathbf{z}'}\mathcal{P}_{\mathbf{z}'\mathbf{y}'}$$
$$\times [\exp(D_{\mathbf{y}'\mathbf{x}'}\varphi_{\mathbf{x}'}) * D_{\mathbf{y}'\mathbf{u}}], \quad (25)$$

where the star denotes a pointwise product in position space, i.e. $(\alpha * \beta)_{\mathbf{x}} = \alpha_{\mathbf{x}}\beta_{\mathbf{x}}$.

VII. REDSHIFT-MARGINALIZED LIKELIHOOD AND POSTERIOR

We can take into account the source redshift uncertainty by generalizing the marginalized likelihood in Eq. (23) to include the probability of the source redshifts z_s given their measured photometric redshifts z_{photo}. This probability is given by the posterior redshift distribution function $P(z^s|z_{\mathrm{photo}})$, where z^s and z_{photo} are vectors of the redshifts

of all sources. Including this redshift posterior in the likelihood yields,

$$P(d|\varphi) = \int \mathcal{D}\boldsymbol{\epsilon}^s \int \mathcal{D}z^s P(d|\varphi, \boldsymbol{\epsilon}^s, z^s)P(\boldsymbol{\epsilon}^s)P(z^s|z_{\mathrm{photo}})$$
$$= \int \mathcal{D}z^s \mathcal{N}[d|R^{z^s}(\varphi), N]P(z^s|z_{\mathrm{photo}}). \quad (26)$$

In most cases, this integration cannot be done analytically and the resulting distribution will in general not be Gaussian. In the absence of further information about the general shape of this distribution, a Gaussian characterized by the first and second central moment of the full distribution is the most conservative approximation one can use (it is the maximum entropy approximation and therefore adds the least additional information). This motivates the use of a Gaussian to model the redshift-marginalized likelihood.

To obtain the first and second moment, we need to calculate $\langle d \rangle_{P(d|\varphi)}$ and $\langle dd^{\dagger} \rangle_{P(d|\varphi)}$. For the expectation value of the data, we obtain

$$\langle d \rangle_{P(d|\varphi)} = \int \mathcal{D}z^s \int \mathcal{D}dd\mathcal{N}[d|R^{z^s}(\varphi), N]P(z^s|z_{\mathrm{photo}})$$
$$= \int \mathcal{D}z^s R^{z^s}(\varphi)P(z^s|z_{\mathrm{photo}}) \quad (27)$$

$$\langle d_i \rangle_{P(d|\varphi)}$$
$$= \mathcal{G}_{ij}\left[\int dz^s P(z^s|z_{\mathrm{photo}})\mathcal{I}_{j\mathbf{r}}^{z^s}\right]\bar{R}(\varphi)_{\mathbf{r}}$$
$$= \mathcal{G}_{ij}\left[\int_0^{\infty} dr \int_r^{\infty} dr^j W(r, r^j)P(z_j^s|z_{\mathrm{photo}_j})\frac{dz_j^s}{dr^j}\right]\bar{R}(\varphi)_{\mathbf{r}}$$
$$\equiv [\mathcal{G}\bar{\mathcal{I}}\bar{R}](\varphi)_i \equiv \tilde{R}(\varphi)_i, \quad (28)$$

where $\bar{I}$ denotes the redshift averaged integration operator defined in the square brackets in the first line of Eq. (28) and we have introduced $\bar{R}(\cdot) \equiv \mathcal{TP}(\exp[D(\cdot)] - 1)$ to summarize the action of all source-redshift independent operators.

The second moment is

$$\langle dd^{\dagger} \rangle_{P(d|\varphi)} = N + \langle (R(\varphi))(R(\varphi))^{\dagger} \rangle_{P(z|z_{\mathrm{ph}})}. \quad (29)$$

Nondiagonal elements of the second term in Eq. (29) read

$$\langle ((R(\varphi))_i(R(\varphi))_j \rangle_{P(z|z_{\mathrm{ph}})}$$
$$= \langle (R(\varphi))_i \rangle_{P(z|z_{\mathrm{ph}})}\langle (R(\varphi))_j \rangle_{P(z|z_{\mathrm{ph}})}$$
$$= (\tilde{R}\varphi)_i(\tilde{R}\varphi)_j \quad (30)$$
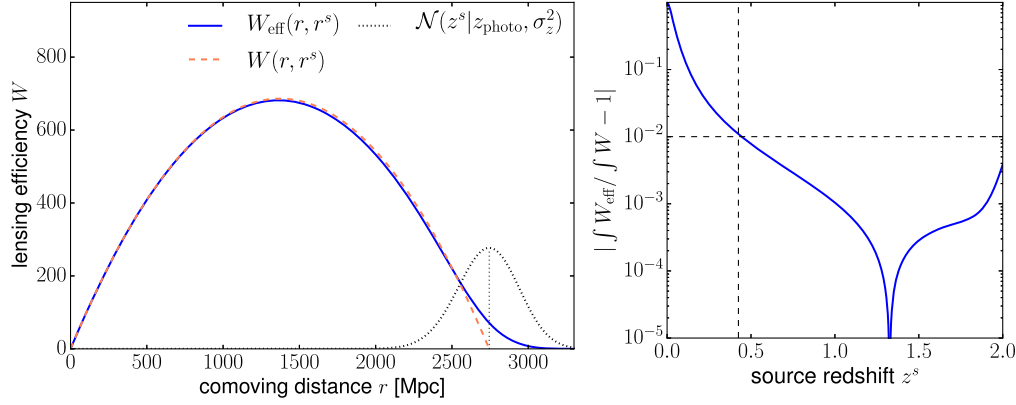
and diagonal elements are

FIG. 1.   Left panel: In the presence of redshift uncertainties the expected measured galaxy shape (i.e. the mean of its likelihood) is obtained by replacing the lensing efficiency $W[r, r^s(z^s)]$ with known source redshift z^s (blue) by the redshift-marginalized effective lensing efficiency W_{eff} (orange-dashed) in the data model (Eq. (28). In the example shown, we have placed a source at redshift $z^s = 0.8$ and assume a Gaussian photometric redshift uncertainty characterized by $\sigma_z/(1+z) \approx 0.04$. The redshift distribution function $P(z^s|z_{\text{photo}})$ is shown for comparison (not normalized). In this example, the pointwise difference between the kernels is below 4% for all $r < 0.9r^s$. Right panel: Relative ratio of the integration volumes for different source redshifts assuming the same redshift error. In this idealized case (we assume a homogeneous density contrast along the line of sight), the effect of replacing the lensing kernel is sub-percent for source redshifts $z^s > 0.43$ as indicated by the dashed lines.

$$\langle (R(\varphi))_i (R(\varphi))_i \rangle_{P(z|z_{\text{ph}})}$$

$$= \mathcal{G}_{ij}\mathcal{G}_{ij} \left[\int_0^\infty dr \int_0^\infty dr' \int_{\max(r,r')}^\infty dr^j W(r, r^j)W(r', r^j) \right.$$

$$\left. \times P(z_j^s|z_{\text{photo}_j}) \frac{dz^j}{dr^j} \right] \bar{R}(\varphi)_{\mathbf{r}} \bar{R}(\varphi)_{\mathbf{r}'}$$

$$\equiv \mathcal{G}_{ij}\mathcal{G}_{ij}\tilde{\mathcal{I}}_{j\mathbf{r},\mathbf{r}'}\bar{R}(\varphi)_{\mathbf{r}}\bar{R}(\varphi)_{\mathbf{r}'}, \qquad (31)$$

where the new operator $\tilde{\mathcal{I}}$ denotes the squared average of the integration operator, i.e. the square brackets in the first line of Eq. (31).

The Gaussian approximation to the likelihood is then $\mathcal{N}[d|\tilde{R}(\varphi), \tilde{N}]$, where $\tilde{N} = N + Q$ and Q is

$$Q = \langle (R(\varphi))_i (R(\varphi))_j \rangle_{P(z|z_{\text{ph}})} - \langle (\tilde{R}\varphi)_i \rangle_{P(z|z_{\text{ph}})} \langle (\tilde{R}\varphi)_j \rangle_{P(z|z_{\text{ph}})}$$

$$= \langle (R(\varphi))_i (R(\varphi))_i \rangle_{P(z|z_{\text{ph}})} - (\tilde{R}\varphi)_i (\tilde{R}\varphi)_i. \qquad (32)$$

This expression is still signal-dependent and we approximate it further by replacing φ by its posterior mean $\langle \varphi \rangle_{P(\varphi|d)}$. Since this mean depends on $\tilde{N}$, the resulting set of equations must be solved iteratively.

To estimate the effect of the additional redshift-marginalization, we compare the effective lensing efficiency that appears in the redshift-marginalized data model [Eq. (28)],

$$W_{\text{eff}}(r, r^s) = \int_r^\infty dr^s W(r, r^s)P(z^s|z_{\text{photo}})\frac{dz^s}{dr^s}, \qquad (33)$$

to the exact lensing efficiency $W(r, r^s)$.

In the left panel Fig. 1 we plot both kernels for a source at redshift $z^s = 0.8$ and a Gaussian redshift posterior with

variance $\sigma_z = 0.04(1 + z^s)$. For this particular configuration we find that both integration kernels are very similar over most of the distance to the source, i.e. their relative difference never exceeds 4% for $r < 0.9r^s$. Only very close to the galaxy the kernels start to differ significantly. At this point the amplitude of the kernels has already dropped by more than 80%. The contribution to the lensing signal from structures with $r > 0.9r^s$ is therefore strongly suppressed.

In the right panel of Fig. 1 we plot the relative difference of the integration volumes of the kernels assuming a uniform density contrast along the line of sight and the same Gaussian redshift posterior with $\sigma_z = 0.04(1 + z^s)$. We find that the difference between the kernels is sub-percent for sources with $z^s > 0.43$.

These tests suggest that the redshift uncertainty should change the lensing signal of individual galaxies on average at the percent level.

We caution, however, that these tests are very simplified and only serve to get a rough estimate of the effect of redshift uncertainties, especially since we have ignored catastrophic outliers or the fact that redshift posteriors can be bimodal. For the proof of the concept of our reconstruction algorithm, we ignore redshift uncertainties. Following a test of their exact impact (e.g. by comparing a reconstruction with exact spectroscopic redshifts to one with redshifts obtained from the corresponding photometric redshift posteriors), they could be included in the way that has been outlined in this section.

VIII. VALIDATION AND TESTS

To validate the implementation and assess the goodness of the tomographic reconstruction, we perform a number of increasingly realistic tests, which we denote tests A, B and

C. In tests of type A we employ an even source distribution over the entire box. These tests are the least realistic ones and serve to validate the correctness of the implementation. Test B uses a surveylike source-redshift distribution and test C adds realistic shape noise.

In all of these tests, we place the observer in the center of the bottom of the computational box[2] and resolve the underlying and reconstructed overdensity fields with 128^3 pixels. Depending on the test, we allow the physical sizes of the box to differ. The current pixel resolution is limited due to computation time and memory usage. A higher resolution will be accessible after parallelization and adaption of the code for the usage on a high-performance cluster.

The minimization is achieved through 300 steps of an LBFGS algorithm [85] followed by a steepest descent algorithm. The latter is set to have reached convergence if the maximal pixel-wise relative difference between two subsequent field estimates is smaller than 10^{-4} for three iteration in a row. If the steepest descent does not converge within 200 iterations, the same combination of LBFGS and steepest descent minimizers is repeated. We find that, at latest in this second run, all of our example reconstructions meet the convergence criterium. In the current serial implementation, the entire minimization takes between 1 day (for idealized tests at fixed redshifts) and 6 days (for a realistic source distribution and shape noise) on a 2.3 Ghz core of an Intel Xeon E5 processor.

Most of the tests are based on mock data that we create by applying the data model described in Sec. IV to nonlinear density fields obtained from N-body simulations. For tests at fixed redshifts we use snapshots of the Millennium-XXL simulation [86] from which we take the entire volume of size $[3h^{-1} \text{ Gpc}]^3$. For more realistic test cases we construct a light cone of size $[500h^{-1} \text{ Mpc}, 500h^{-1} \text{ Mpc}, 4000h^{-1} \text{ Mpc}]$ by joining snapshots of the Millennium Run [87]. The Millennium box measures $500h^{-1}$ Mpc along each side and we shift it every time $500h^{-1}$ Mpc in the z-direction have been constructed. This procedure ensures that a LOS is unlikely to hit the same structure repeatedly. In the resulting light cone we achieve a resolution of $3.9h^{-1}$ Mpc in the x- and y-directions corresponding to mildly nonlinear scales and reach a redshift of $z = 2.2$ in the z-direction. Note that the physical size of the pixels is eight times longer along the z-axis, meaning that we obtain a poorer resolution in radial direction.

We map the simulations onto our pixel grid by a nearest grid point interpolation. This procedure suppresses power on small scales and modifies the power spectrum by a kernel consisting of cardinal sine functions (see e.g. [88]). To be consistent with this modification, we apply the same kernel to the theory power spectra which we use in the prior distributions.

---

[2]This position is not fixed by the algorithm. For a reconstruction from a survey that covers a significant fraction of the sky, the observer can be placed in the center of the box, for example.

Both the Millennium-XXL and the Millennium simulation use a flat ΛCDM cosmology with $\Omega_\text{m} = 0.25$ and $h = 0.73$.

A. Fidelity of the reconstruction

The BFGS algorithm is a quasi-Newton method. While Newton methods rely on an exact evaluation of the inverse Hessian at each step, quasi-Newton methods estimate the curvature from previous evaluations of the gradient. The "L" in L-BFGS refers to the limited memory version of this algorithm in which the inverse Hessian is approximated by a sparse matrix instead of a dense one. This additional approximation becomes necessary for problems with many degrees of freedom such as this. In principle the final estimate of the inverse Hessian could be used as an approximation for the pixel-pixel covariance, but it depends on the correlated previous steps of the algorithm and only probes the curvature along the search directions. The covariance, C, of a Gaussian posterior distribution, $\mathcal{G}(\varphi|d)$, with maximum m can be estimated by sampling. A straightforward algorithm to obtain such an estimate is outlined in the following.

(1) Draw random realizations of signal and noise from the Gaussian prior and noise distributions.

$$\varphi' \hookleftarrow \mathcal{G}(\varphi, \Phi) \quad \text{and} \quad n' \hookleftarrow \mathcal{G}(n, N). \quad (34)$$

(2) Use the resulting realizations to model a data vector d'

$$d' = R\varphi' + n'. \quad (35)$$

(3) From this random data vector reconstruct the Wiener filter estimate m', i.e. the maximum of the Gaussian posterior distribution,

$$m' = CR^\dagger N^{-1} d', \quad \text{with} \quad C = [R^\dagger N^{-1} R + \Phi^{-1}]^{-1}. \quad (36)$$

(4) Repeat steps 1–3 N times and estimate C from

$$\hat{C} = \frac{1}{N} \sum_i^N (\varphi_i' - m_i')(\varphi_i' - m_i')^\dagger. \quad (37)$$

Computing the full covariance is computationally infeasible since it would require ∼9 Tb to store a full covariance matrix of size $1/2 n(n + 1)$ for a map with $n = 128^3$ pixels. Numerically feasible are, e.g., estimates of its diagonal (the pixelwise covariance), or of the full covariance of 1D slices of the reconstruction box.

The above algorithm yields an unbiased estimator of the posterior covariance only if the data model is linear in the density field. The lognormal model is inherently nonlinear and therefore the posterior non-Gaussian. However, for sufficiently small perturbation Δm around the MAP estimate m the Laplace approximation should be applicable, meaning that we can approximate the negative log posterior as a Gaussian,
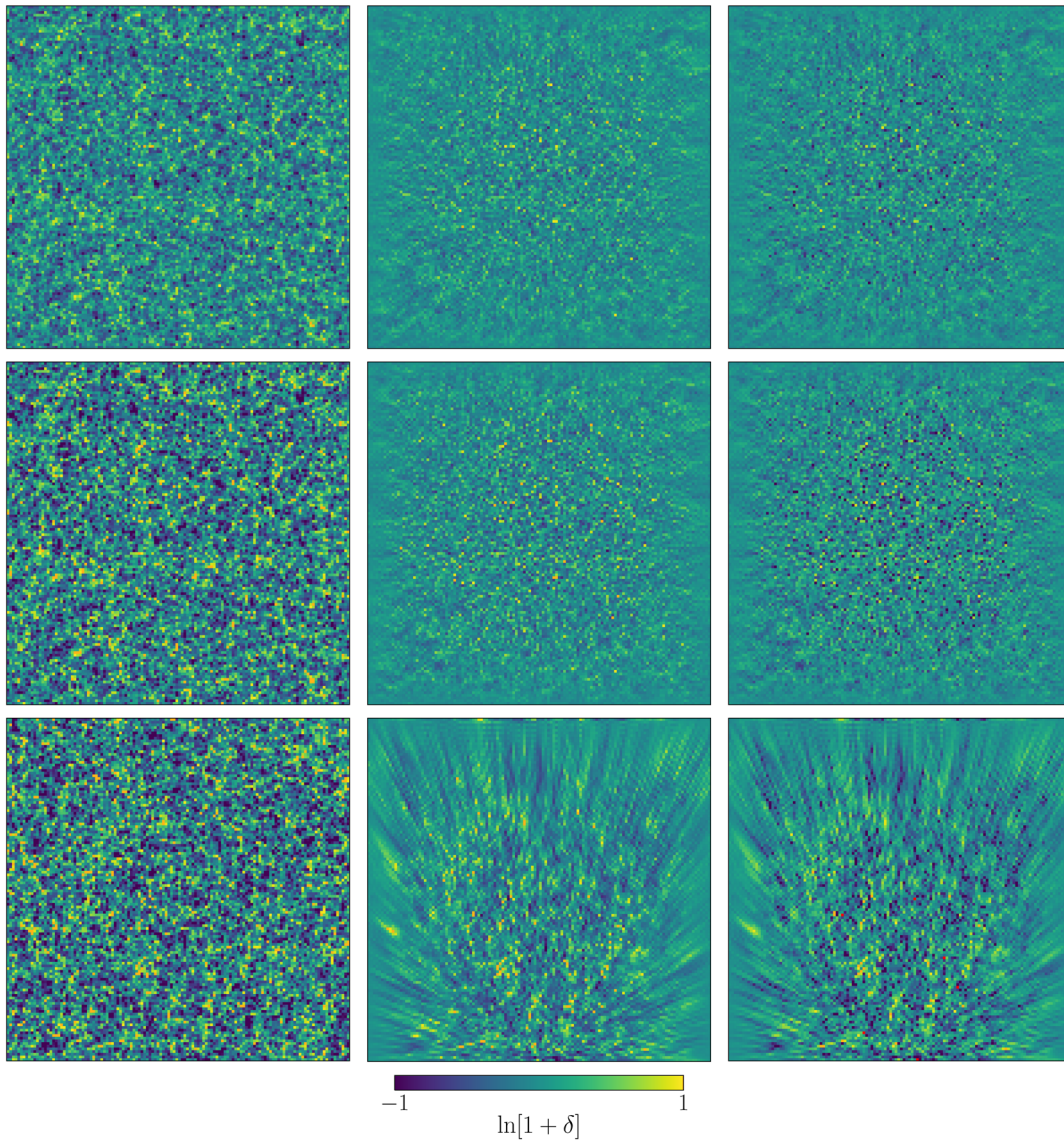
$$-1 \qquad\qquad 1$$
$$\ln[1 + \delta]$$

FIG. 2.   Qualitative comparison of reconstruction methods in the test series A2, for which we create mock data from individual simulation snapshots, apply an even source distribution and add negligible shape noise. We show central slices through the 3-dimensional fields. The observer is located in the center of the $(x - y)$-plane at the origin of the z-axis. The redshift of the snapshot decreases from top to bottom, $z = [1, 0.25, 0.]$. In the two upper rows we show the $(x - y)$ plane at $z = \max(z)/2$, in the last row we look at the $(x - z)$-plane at $y = \max(y)/2$. The fields are from left to right: the underlying density field from the simulation, its reconstruction with a lognormal prior and its reconstruction with a Gaussian prior (Wiener filter). We plot $\ln[1 + \delta]$ and mark unphysical negative densities in the Wiener Filter reconstruction in red. The resolution of the reconstruction decreases with distance to the observer. This is because (1) the density of lines of sight decreases (2) there are less sources behind the point we want to reconstruct (3) the information from these sources is suppressed by the shape of the integration kernel. In all cases, the lognormal reconstruction is superior in capturing the highest values of the density field and avoids unphysically low density contrasts below -1. The Gaussian prior seems better in identifying low density regions.

$$-\ln P(\varphi = m + \Delta m | d) \approx \frac{1}{2}(\tilde{d} - R'\tilde{\varphi})^{\dagger} N^{-1}(\tilde{d} - R'\tilde{\varphi})$$

$$+\frac{1}{2}\tilde{\varphi}^{\dagger}\Phi^{-1}\tilde{\varphi} \tag{38}$$

$$\hat{=} -\ln \mathcal{G}(\tilde{\varphi}|d). \tag{39}$$

To obtain Eq. (38), we have linearized the response operator

$$R' = \frac{\delta R}{\delta \varphi}\bigg|_{\varphi = m}, \tag{40}$$

redefined the data vector $\tilde{d} = d - R(m) + R'(m - \bar{\varphi})$ and shifted the signal $\tilde{\varphi} = m + \Delta m - \bar{\varphi}$. The covariance C' of this Gaussian approximation to the posterior is

$$C' = [R'^{\dagger} N^{-1} R' + S^{-1}]^{-1}. \tag{41}$$

This approximate covariance can be evaluated as outlined above. As an example, we estimate $\mathrm{diag}(C')$, which can be interpreted as an uncertainty map for the reconstruction, for one of our test cases in the next section.

B. Tests A: Simple geometry

In a first series of tests we distribute 500 000 sources evenly in the box and slightly beyond ($\pm 100 h^{-1}$ Mpc). We place sources outside of the reconstructed volume in order to increase the area in which we can recover the underlying density field with high resolution. The quality of the reconstruction decreases with increasing distance to the observer because of increasing distances between lines of sight and a decreasing number of background sources that help to break the LOS degeneracy of the lensing kernel. We further only add unphysically low, negligible shape noise in this first test set.

We perform three different tests of type A for which we create data from increasingly realistic input fields:

(1) A self-consistency test, where we create an over-density field from a random realization of the Gaussian field φ [Eq. (3)] in a cubic box of side length $1000 h^{-1}$ Mpc. We then apply the implemented data model and use the algorithm to recover this input field. We do not show any results from this test, since they would not provide any additional insights compared to tests A2 and A3, but simply state that we can recover the input field with high fidelity which means that the implementation is in itself consistent.

(2) A test in which we create shears from overdensity fields taken from an N-body simulation at fixed redshifts. We use three different snapshots at individual redshifts $z = 1$, $z = 0.25$ and $z = 0$. This allows us to assess the algorithm's ability to recover increasingly nonlinear fields. In each case, we compare the lognormal reconstruction to a Wiener filter (WF) reconstruction, which uses the same data model and source distribution but a Gaussian prior on the overdensity field δ. Comparisons between input and reconstructed fields in each case are shown in Fig. 2. The resolution of the reconstruction decreases with distance to the observer. This is because (1) the density of lines of sight decreases (2) there are less sources behind the point we want to reconstruct (3) the information from these sources is suppressed by the shape of the integration kernel. In Table I we compare the minimal and maximal values in the reconstructions to the extremal values in the true density field. The lognormal reconstruction is superior in capturing the highest values of the density field and avoids unphysically low density contrasts below -1. Table II summarizes pixel-wise quadratic differences and the pixel-wise Pearson correlation between underlying and reconstructed fields. For all redshifts tested the lognormal prior yields higher correlations with the

TABLE I. Quantitative comparison of reconstruction methods for tests A2 and A3. We compare the minimal and maximal values of the underlying and reconstructed fields (δ and $\hat{\delta}$, respectively).

Redshifts	max($\hat{\delta}$), max(δ)			min($\hat{\delta}$), min(δ)		
	Lognormal	WF	N-body	Lognormal	WF	N-body
1.	4.22	2.95	4.92	−0.63	−0.91	−0.79
0.25	7.38	5.00	8.49	−0.77	−1.28	−0.87
0.0	8.32	5.55	10.26	−0.99	−1.57	−0.89
light cone	32.90	20.81	70.20	−0.98	−4.28	−0.92

TABLE II. Quantitative comparison of reconstruction methods for tests A2 and A3. We compare the lognormal to the Gaussian prior in terms of the mean square pixel-wise difference (first column) and the Pearson-correlation coefficient (second column) between the reconstruction and the underlying field (which we denote $\hat{\delta}$ and δ, respectively).

Redshifts	$\langle(\hat{\delta}-\delta)^2\rangle/\sigma_\delta^2$		$\sqrt{\langle\hat{\delta}\delta\rangle/(\sigma_\delta\sigma_{\hat{\delta}})}$	
	Lognormal	WF	Lognormal	WF
1.	0.66	0.69	0.58	0.56
0.25	0.73	0.77	0.52	0.49
0.0	0.73	0.77	0.53	0.48
light cone	0.65	0.73	0.61	0.51

original field than the Gaussian prior. The difference between lognormal and WF reconstruction increases as the input field becomes more nonlinear. This is also reflected in the 1-point probability distribution functions (PDFs), which we show in Fig. 3. While the Wiener filter PDF is closer to symmetric in all cases, we can capture part of the skewness of the input field by applying the lognormal prior. A notable feature of the 1-point PDFs is that the maximum value is slightly biased in both reconstructions. The distributions of the reconstructions peak at zero, i.e. at their mean, while the underlying field peaks below. This is a feature of both priors since they prefer the mean density (or a zero density contrast) if the data does not contain enough information on the density.[3]

(3) zA test similar to A2 but with a redshift dependent density field constructed from different snapshots. The physical size of the box spans $[500 h^{-1}$ Mpc$]^2$

---

[3]To be precise one should note that the maximum of the lognormal posterior lies slightly below the mean density. However, the difference to the mean turns out to be small in all test cases. For the light cone box we have calculated that the prior distribution peaks at $\delta = -0.1$. This value is consistent with the lognormal reconstructions that we obtain in our test cases. Closely inspecting the right panels of Figs. 5 and 10, one can see that the bulge of the reconstructed densities lies slightly below 1, at $\delta + 1 \approx 0.9$).
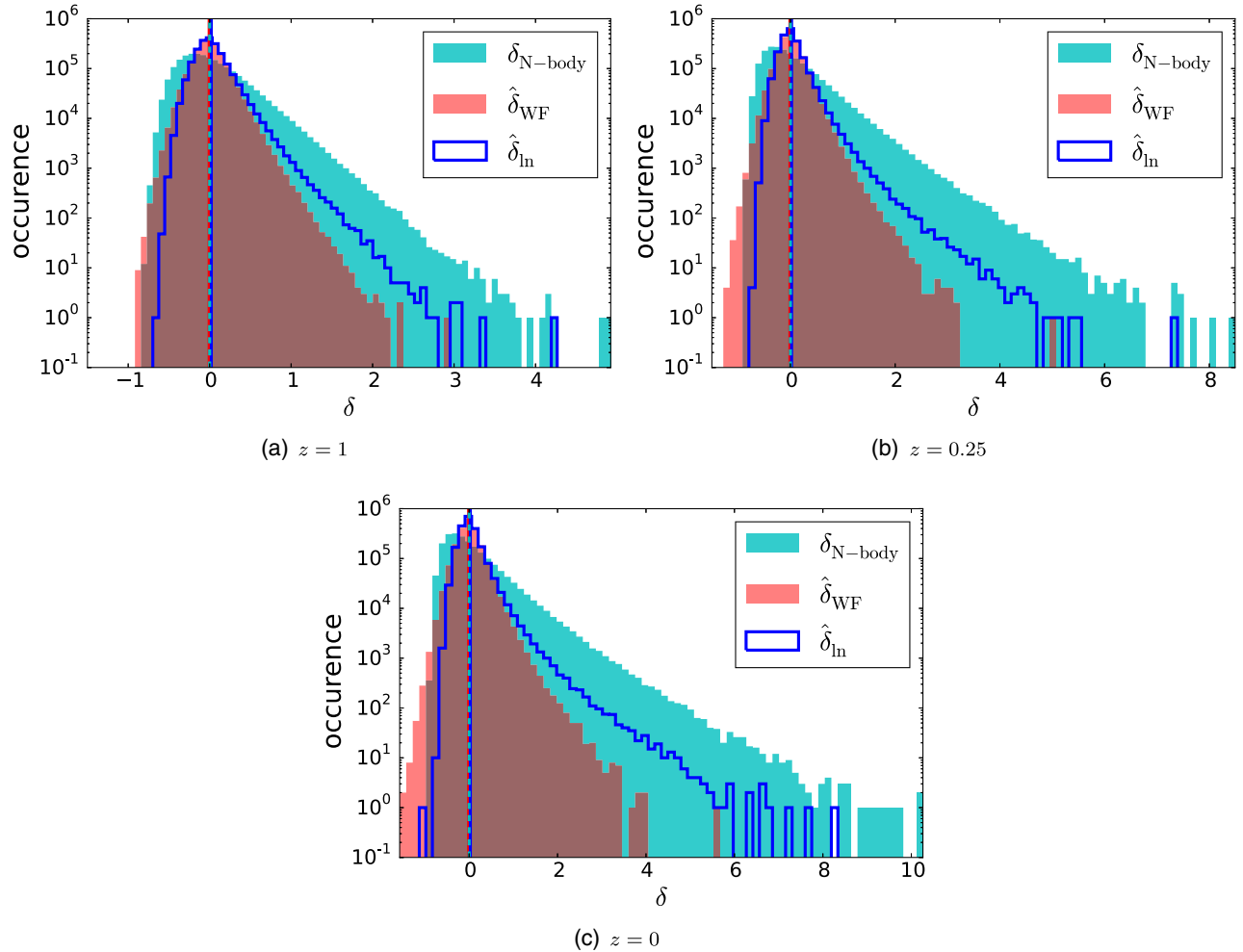
FIG. 3.   1-point PDFs of the simulated density contrasts (in cyan) and their Wiener filter and lognormal reconstructions (red and dark blue) in tests on mock data created from different snapshots of an N-body simulation with an even source distribution and negligible shape noise (A2). The PDF of the lognormal reconstruction is slightly more skewed than the Wiener filter PDF, which is closer to symmetric in all cases. Both the Wiener filter and the lognormal reconstruction are equally biased in the position of the peak, however, the mean values of both reconstructions agree well with the mean value of the underlying density. The mean values of the density contrast are indicated by vertical lines in the same color as their corresponding distribution. We also find that the Wiener filter produces unphysical density contrasts below -1 in snapshots at $z = 0.25$ and $z = 0$.

in the $(x - y)$-plane and $4000h^{-1}$ Mpc into the z-direction corresponding to a maximum redshift of $z = 2.2.$[4] Results of test A3 are shown in Fig. 4 and Fig. 5. The quantitative comparison between

---

[4]The light cone is constructed by merging planes from different snapshots along the z-direction of the box, but the algorithm assumes that the distance and redshift increase in radial direction from the observer. This leads to a mismatch between the distance to a (source) position assumed by the algorithm and the distance or rather redshift that this position corresponds to in the simulation. The distance in the simulation is $\cos\theta$ times the distance assumed by the algorithm, where θ is the angle between the LOS and the z-axis. This mismatch should only become relevant for relatively large angles $\theta > \pi/2$. Lines of sight with such angles only cover a minor fraction of the box and we therefore expect the test to not be severely affected by this approximation.

the lognormal and the corresponding WF run is listed in the last rows of Tables I and II, labeled "light cone." We find a similar match between the underlying and reconstructed fields as in tests A2 and again a superiority of the lognormal prior over the Wiener filter reconstruction in terms of correlation coefficients, pixelwise differences and extremal values. The distributions of the pixelwise matching between input and reconstructions in the left panel of Fig. 5 imply that the lognormal prior is more likely to reconstruct overdensities correctly, while the Wiener filter is better in tracing the underdensities. The full distributions in the right panel of the same figure suggest that the better reconstruction of low densities with the Wiener filter is a consequence of the symmetry of the
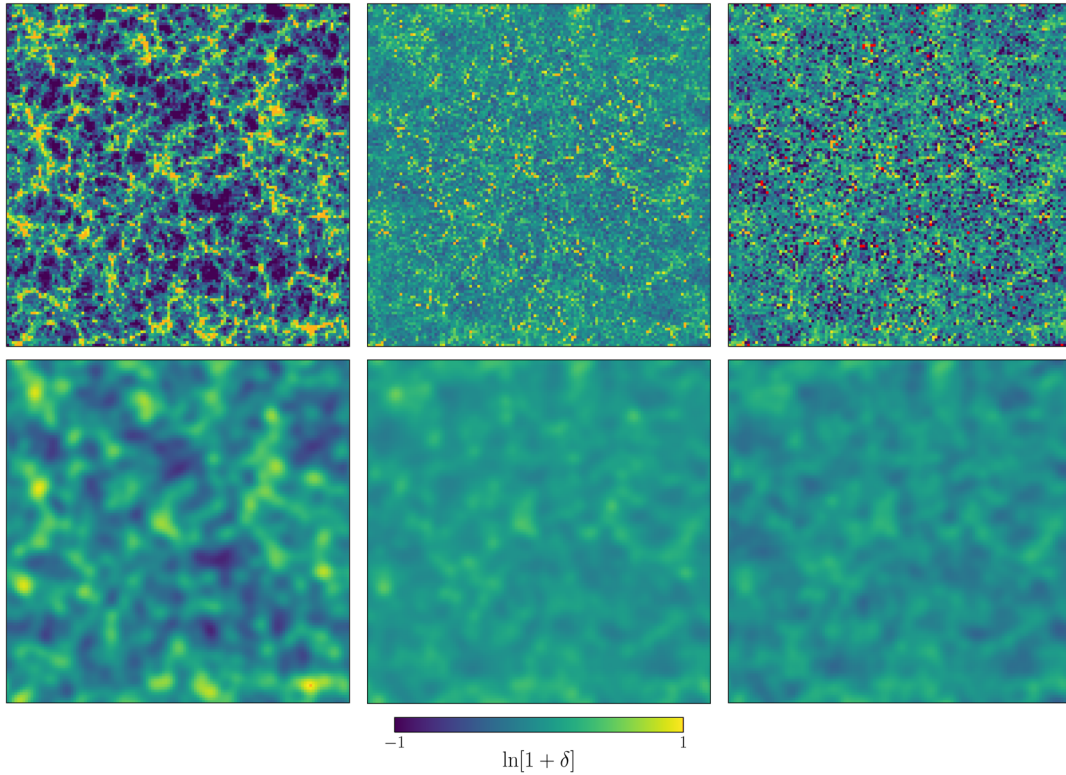
FIG. 4. Qualitative comparison of reconstruction methods in test A3. For this test, we create mock data from a simulated density field on a light cone, use an even source distribution and add negligible shape noise. The fields are from left to right: the underlying overdensity field from the simulation, the reconstructed overdensity field using a lognormal prior and the reconstructed overdensity using a Gaussian prior (Wiener filter). We show the central $(x - y)$-plane of the reconstruction box in each panel. The observer is located in the center of the $(x - y)$-plane at the origin of the z-axis. Note that we plot $\ln[1 + \delta]$ and mark negative densities in red. In the second row, we show a smoothed version of the fields. The smoothing is performed with a Gaussian kernel with $\sigma = 8h^{-1}$ Mpc.
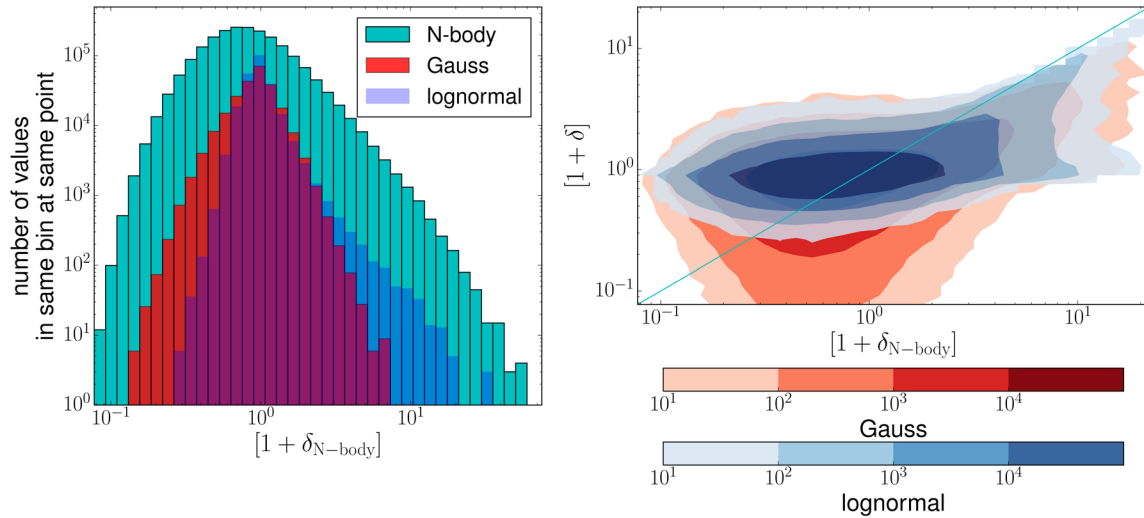
FIG. 5. Comparison of reconstruction methods in test A3: In the left panel, we have binned the underlying density into 40 bins of equal width (0.17) in ln-space and count the points in the reconstructions whose density lies in the same bin at the same location. The total number of points in a density bin in the original density field is shown for comparison in the background (cyan). This plot shows that the lognormal model is better in capturing overdensities, while it is worse than the WF reconstruction in capturing underdensities (note the log scaling). In the right panel, we plot the full distributions of the reconstructions against the underlying density. A perfect reconstruction would follow the diagonal. The lognormal reconstruction achieves this better in the high density tail and prefers values around the mean in poorly constrained regions. The Gaussian prior tends to identify more low density regions.
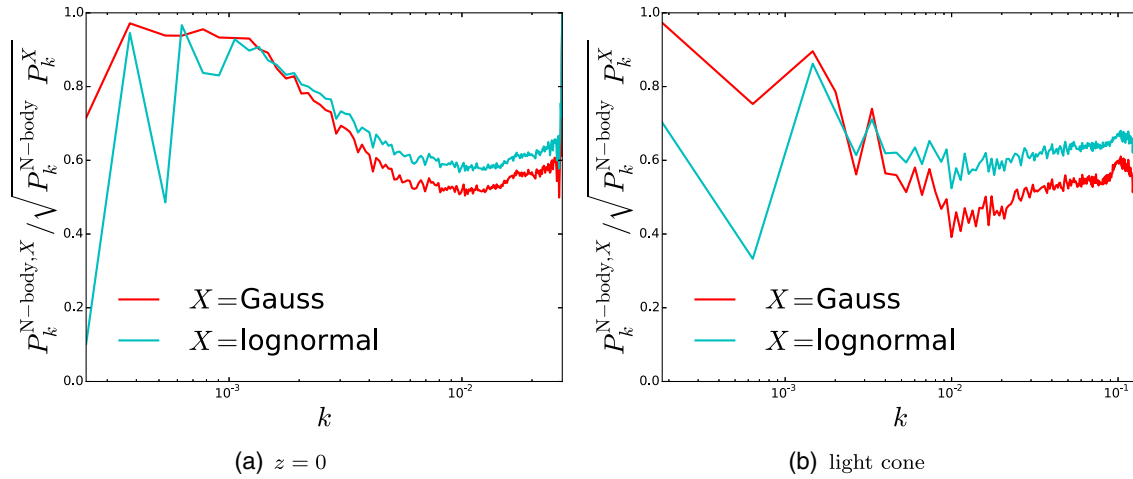
(a) $z = 0$　　　　　　　　　　　　　　　　(b) light cone

FIG. 6.　Cross spectra $(2\pi\delta_D(\mathbf{k} - \mathbf{k}')P_k^{Y,X} = \Re\langle\delta_Y(\mathbf{k})\delta_X(\mathbf{k}')\rangle)$ between the underlying density field and the reconstructions divided by the square root of the auto spectra. We show results for tests of type A at $z = 0$ (left panel) and on the light cone (right panel). Albeit adding only negligible noise in this test, we do not expect a perfect correlation due to the information loss caused by the sparse sampling of the underlying field by the data. We find that the Gaussian model is better in reconstructing the largest-scales, while the lognormal model shows higher correlations at low and intermediate scales.

Gaussian prior: The Gaussian prior suppresses both positive and negative offsets from the mean density by an equal amount. Because of this property it also allows negative densities. The lognormal prior strongly suppresses deviations towards low densities, while it is less sensitive to deviations towards higher densities (since it suppresses deviations equally on a logarithmic scale). As a result, the lognormal model refrains from assigning very low densities to regions with small signal-to-noise.

In Fig. 7 we show estimates of the reconstruction uncertainties in this test which were obtained by following the prescription in Sec. VIII A. We find that the estimator in Eq. (37) converges very quickly and use 20 samples (from the 15th sample on the average change per pixel from adding another sample is less than 0.05, adding the 20th sample changes the uncertainty estimate on average only by 0.03 per pixel). For both priors, but especially for the lognormal prior, a comparison to the input field shows that overdensities are reconstructed with higher fidelity than low density regions. The Wiener filter reconstruction is less sensitive to overdensities but slightly more sensitive to underdensities than the lognormal reconstruction.

In Fig. 6 we compare the correlation coefficient between input field and reconstructions on different scales. The Gaussian model provides better fits on the largest scales. This could be due to the aforementioned missing power in low density regions in the lognormal reconstruction, which can be best seen by eye in the lower row of Fig. 4. On intermediate and small scales the power spectrum of the lognormal reconstruction shows higher correlations with the original field. We note that the estimator in this work

was designed for the map level and is not expected to be an optimal estimator for the power spectrum.

C. Test B: Realistic survey geometry

In the test of type B, we employ a realistic survey geometry and source distribution. We use input over-densities from the same light cone as in test A3 to generate mock data and place sources in a cone that spans 7.15°. The sources are distributed according to a distribution function of form

$$n(z) = z^\alpha \exp[-(z/z_0)^\beta], \qquad (42)$$

where we fit α, β, and z_0 to the publicly available source distribution of the CFHTLenS survey.[5] The resulting source distribution function is shown in the left panel of Fig. 8. The fit does not exactly match the survey distribution, but the similarity is sufficient for this test. We then draw $\Omega[\text{arcmin}^2]\rho_s[\text{gal/arcmin}^2]$ source positions from this distribution, where Ω is the angular opening area of the cone in arcmin^2 and ρ_s the source density of the survey, which we choose to be $\rho_s = 11 \text{ gal/arcmin}^2$ corresponding to the source density in CFHTLenS. A scatter plot of the resulting spatial source distribution is shown in the right panel of Fig. 8. The reconstructions are depicted in Fig. 9.

For a realistic source distribution the lognormal model seems to better reconstruct the underlying field at lower redshifts, while the Gaussian prior better traces distant low
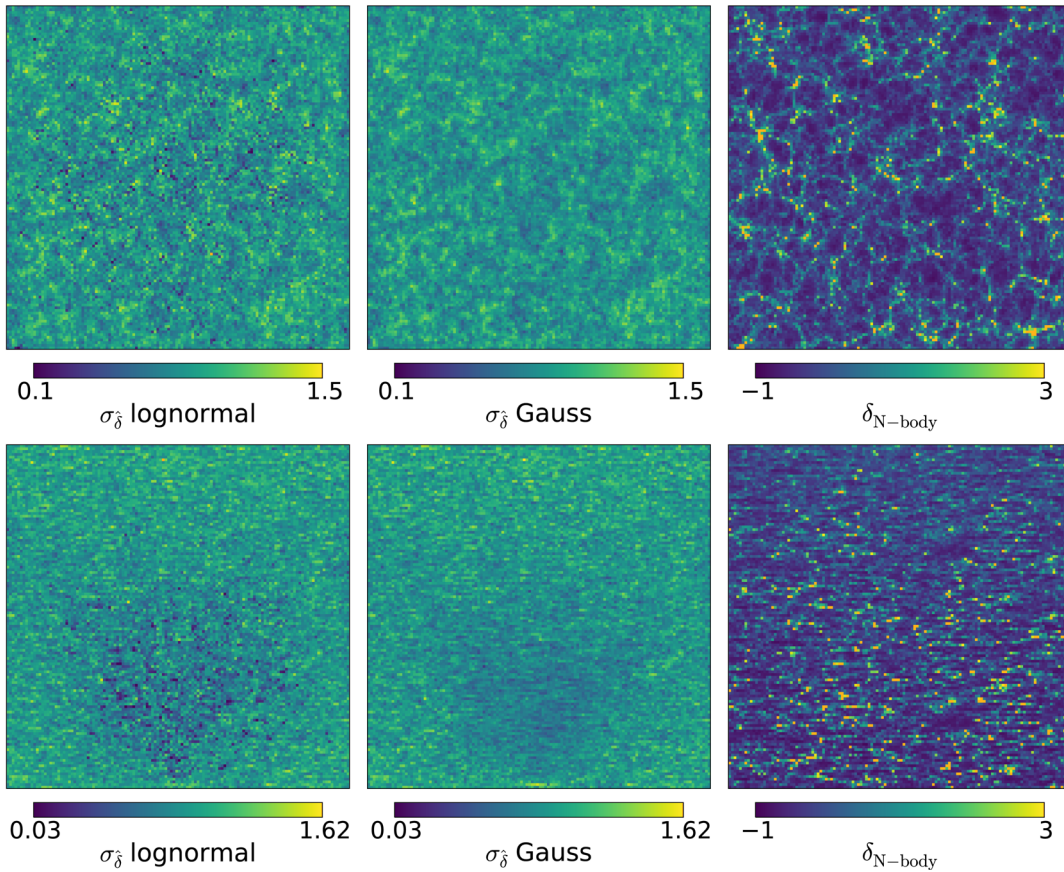
---

[5]http://www.cfhtlens.org/astronomers/cosmological-data-products.

FIG. 7. Estimated uncertainty maps, $\sigma_{\hat{\delta}} = \sqrt{\mathrm{diag}(\hat{C})}$ (c.p. Sec. VIII A), of the reconstructions in Test A(3) obtained with 20 samples. In the upper row we show uncertainty maps of the central slice through the $(x-y)$-plane, in the lower row through the $(x-z)$-plane. The columns are from left to right: estimate of $\sigma_{\hat{\delta}}$ of the lognormal reconstruction, the Wiener filter reconstruction, and the input field in linear scaling for comparison.

density regions. We interpret this is again a consequence of the symmetry of the Gaussian distribution, the same property which leads to the abundant reconstruction of negative densities, nicely visible in the 1-point PDF (Fig. 10).

D. Test C: Realistic shape noise

In the last test of our test series, we add realistic shape noise with $\sigma_\epsilon = 0.25$ to the mock data on the light cone. This value is close to the value in CFHTLenS, $\sigma_\epsilon = 0.279$ [38]. Compared to test B, we use the same opening
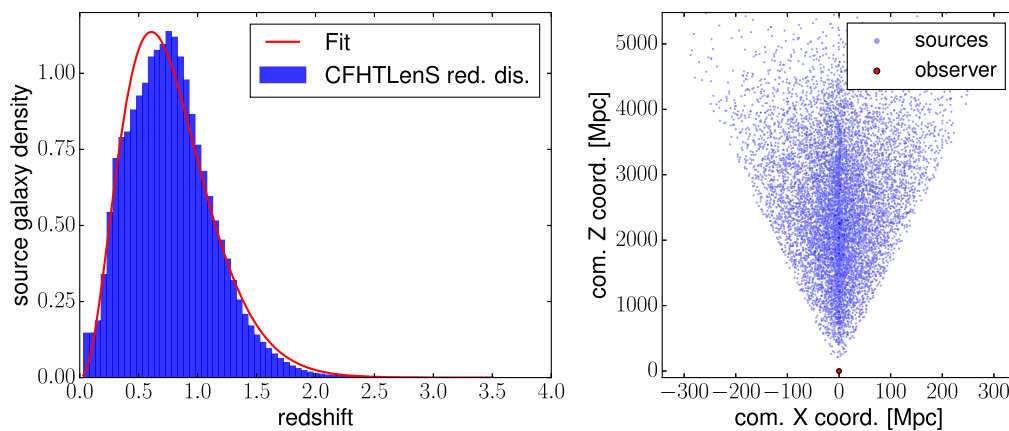

FIG. 8. The source redshift distribution used for test B: In the left panel, we show the histogram of source redshifts from CFHTLenS and our fit to this distribution of the form given in Eq. (42). The right panel depicts the projected 3D scatter plot of the galaxy sample that is drawn from the fit and used in test B (we only show every 200th source).
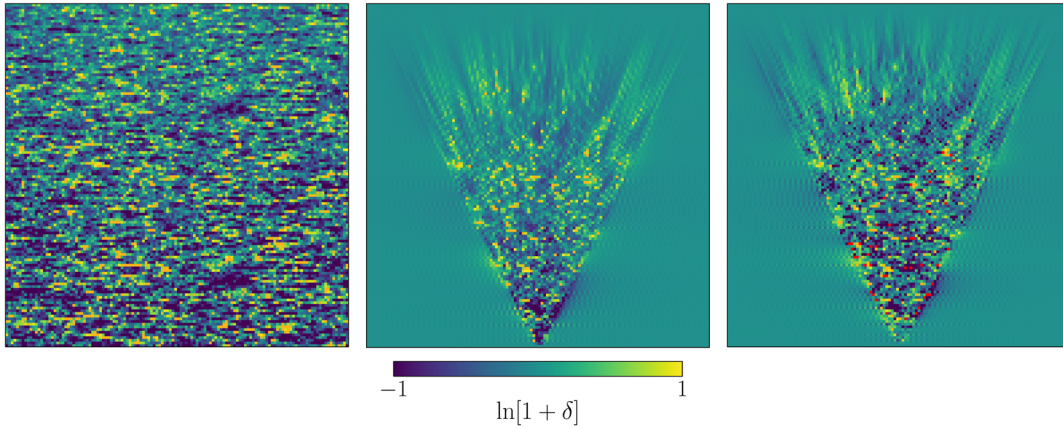
FIG. 9.   Results of test B, reconstructions of a redshift-dependent density field on the observer's light cone from mock data with a realistic source distribution but negligible shape noise. We show central slices through the y-axis of the underlying field, $\ln[\delta_{\text{N-Body}} + 1]$, in the left, the lognormal reconstruction, $\ln[\hat{\delta}_{\text{ln}} + 1]$, in the middle and the Wiener filter reconstruction, $\ln[\hat{\delta}_{\text{WF}} + 1]$ in the right panel. Negative densities are marked in red. The apparent anisotropy of the fields stems from the fact that we have squeezed them in the z-direction. We use square pixels in the plot, while their physical size is eight times larger in the z-direction than in the x-direction.
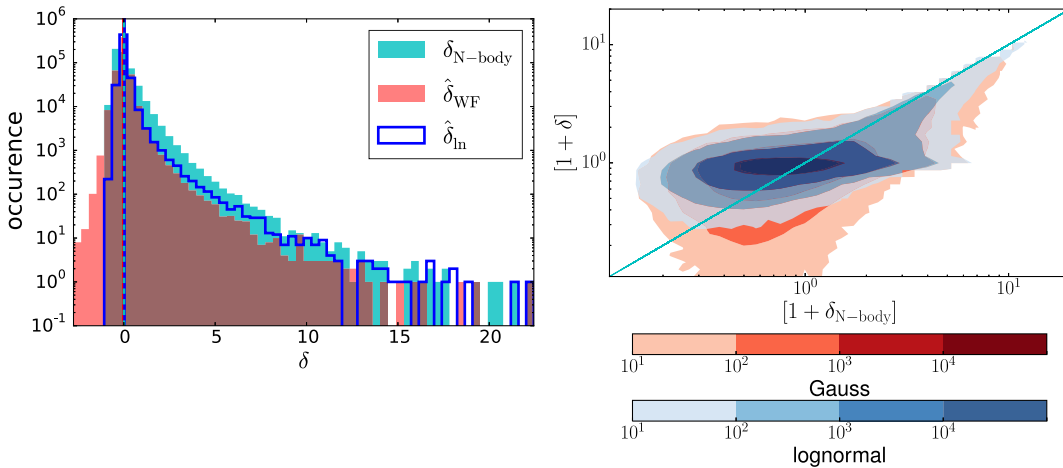

FIG. 10.   Results of test B, reconstructions from mock data with a CFHTLenS-like source distribution. We have restricted the analysis to the field values that lie within the observed cone. In the left panel we show the 1-point PDFs of the fields with mean values indicated by vertical lines in the same color as their corresponding distribution. The right plot shows the density distributions of the reconstructions against the underlying density. The lognormal model is clearly better in tracing the highest values of the true density distribution.

angle of 7.15° but a slightly higher galaxy density of 15 gal/arcmin². We show results of this final test in Figs. 11 and 12. The left panel of Fig. 11 shows that the lognormal model is more likely to reconstruct the correct density in a pixel. However, the PDF in the right panel seems to have a spurious tail of high density values. This tail does not significantly deviate from the high density tail of the underlying reconstruction and is therefore hardly suppressed by the prior. Figure 11 reveals that these high density values are due to an overestimation of densities in slices normal to the direction of observation. We have verified that these overdense sheets are not due to a wrong minimum or saddle point by starting the

algorithm at either the WF solution or the true input field and changing convergence criteria. We interpret these features as follows: Adding a constant density offset to a slice normal to the observational direction does not significantly change the likelihood, because of the well known mass-sheet degeneracy—the fact that adding a homogeneous mass sheet does not change the lensing signal. The lognormal prior does not punish or even encourages these overdensities since it expects a skewed density distribution. The Gaussian prior suppresses strong deviations from a symmetric density distribution, which is why these features are not present in the Gaussian case.
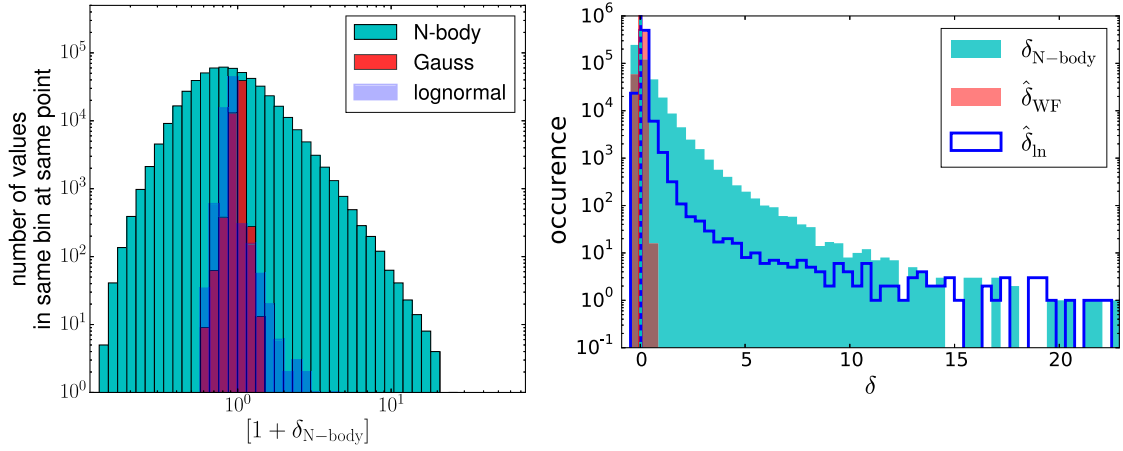
FIG. 11.   Results of test C, reconstructions from mock data with a CFHTLenS-like source distribution and realistic shape noise. We have restricted the analysis to the field values that lie within the observed cone. In the left panel, we count the pixels in which the reconstructed densities coincide with the original density field (i.e. lie within the same density bin). Similarly to test with low noise, we find that the lognormal prior is more likely to reconstruct overdensities correctly. In the right panel we show the 1-point PDFs of the fields with mean values indicated by vertical lines in the same color as their corresponding distribution. The lognormal reconstruction seems to trace the high density tail of the underlying distribution, while the Gaussian density estimates remain very conservative.
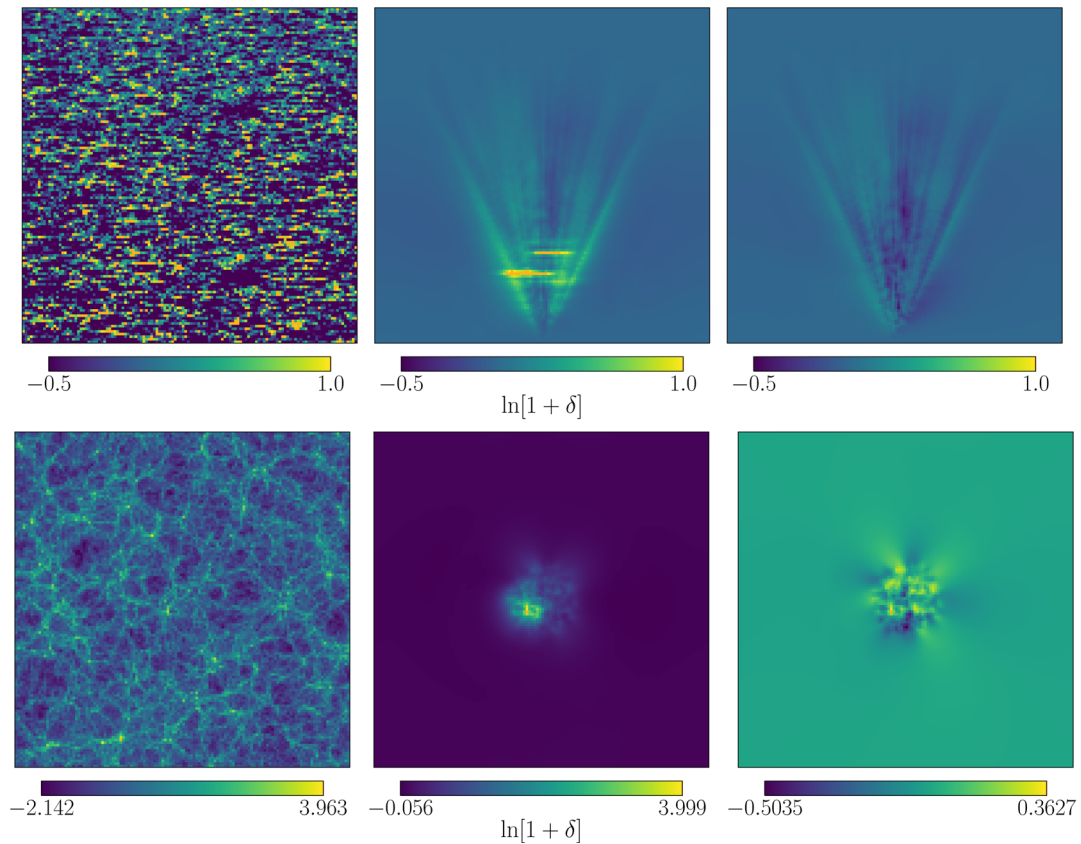

FIG. 12.   Results of test C: reconstructions from CFHTLenS-like mock data with realistic shape noise. In the left column we show slices through the underlying density field, in the central and middle columns the corresponding slices through the lognormal and WF reconstructions, respectively. In the upper row, we show a central slice through the y-axis. Both reconstructions suffer from a strong line of sight degeneracy. The Wiener filter seems to slightly better reconstruct underdensities, while the lognormal model seems more sensitive to overdensities. The latter even overestimates the densities in certain slices normal to the observational direction. We interpret this as a consequence of mass sheet degeneracy. To verify that the algorithm identifies real structures in these overdense slices, we show one of them (the lower one) in the lower row.

IX. DISCUSSION AND FUTURE IMPROVEMENTS

We have presented a novel, fully Bayesian, lensing tomography algorithm that reconstructs the three-dimensional matter distribution from measurements of individual galaxy shapes. Given a data vector of measured galaxy shears, the galaxies' angular positions and their redshifts, this algorithm finds the maximum *a posteriori* estimate of the density contrast between the lensed galaxies and the observer. Since galaxy shapes provide only limited information about the intervening structure, this reconstruction problem is in general ill-constrained and requires regularization. In case of a maximum *a posteriori* estimate this regularization is provided by the prior probability distribution—the distribution which ideally contains all available information about the underlying field before the lensing measurement.

This ideal distribution, however, is in reality unaccessible. Similar previous works have used a Gaussian prior in this context. A Gaussian prior (together with a Gaussian noise model) results in a Gaussian posterior distribution which has a distinct minimum and no saddle points. Maximizing this posterior is comparably straightforward. The algorithm presented in this work was especially designed to also handle non-Gaussian posteriors. We show this ability by using a lognormal prior which should be an improvement over the Gaussian approximation in two ways: First, it enforces the strict positivity of the density while a Gaussian prior allows unphysical, negative densities. Second, it incorporates an *a priori* knowledge of the presence of odd moments in the matter distribution, which arise as a consequence of nonlinear structure formation. These corrections are relevant since the cosmic shear signal probes structures down to scales that lie well within the nonlinear regime.

We note that neither of these priors describe the exact distribution of the nonlinearly evolved density contrast. In regions with high signal-to-noise this is not problematic since the prior is updated by the information contained in the data. In less well constrained regions the prior will paint in missing information. This can lead to systematic deviations from the true distribution. However, nonoptimal priors can be sufficient depending on the quantity and scales one is interested in reconstructing. For example a Gaussian prior will be a good choice for large scales, where the field is still Gaussian. On small scales the lognormal model will yield better reconstructions since it is informed about the skewness of the distribution and avoids negative densities.

To test the algorithm and assess the differences between a reconstruction with a lognormal and a Gaussian prior (in the latter case the reconstruction simplifies to a Wiener filter), we have applied it in both configurations to increasingly realistic mock data sets. Those data sets were produced by applying the algorithm's inherent data model to nonlinearly evolved densities from N-body simulations.

In tests with negligible shape noise, but a realistic CFHTLenS-like source distribution, we find that the reconstruction with a lognormal prior avoids negative densities that are abundantly present in the Wiener filter reconstruction. It is also better in reconstructing the highest local peaks in the density field, and it leads to a higher pointwise correlation between the true underlying density and its reconstruction. The Gaussian model seems better in tracing low density regions and capturing the largest scales. In regions with low signal-to-noise both priors tend towards the mean density.

For realistic shape noise, the data contains too little information to break the line-of-sight degeneracy of the tomographic problem with either prior. This is expected for current data sets, which are commonly analyzed in a limited number of broad redshift bins in order to yield sufficient signal to noise per bin. In addition, the lognormal model is prone to mass-sheet degeneracy: since the prior encourages a skewed distribution with more high density pixels, the reconstruction tends towards inserting sheets of higher density normal to the observational direction. The situation might slightly improve for future surveys with a higher source density. But since the shape noise scales only with the square root of the source density, even future data is unlikely to contain sufficient information for a 3D reconstruction without this degeneracy.

The fidelity of the reconstructions in this work might also be overestimated since we assumed precisely known source redshifts and point spread function, and ignored intrinsic alignments. In this work we show how photometric redshift uncertainties can be included, but leave the implementation for future work. Other sources of uncertainty could be included in a similar fashion. Another approximation we made concerns the source distribution. In reality the source distribution is correlated with the underlying density field, which we ignore in our test cases. A more clustered distribution of sources would render the fidelity of the reconstruction more inhomogeneous since structures in front of high density regions would be detected with higher signal to noise than structures in front of low density regions. Since the volume fraction of high density regions is smaller than the volume fraction of low density regions, this could decrease the overall fidelity of the reconstruction. To get an estimate of the effect of the clustering, future tests of the algorithm should account for the correlation between galaxies and density field by using the halo positions identified in the simulation by a halo finder.

Possible improvements of the algorithm could go into two directions, (1) improving the prior or data model, (2) including additional sources of information. An improved data model could be obtained by forward modeling structure formation numerically for an initially Gaussian distributed field. While this is conceptually straightforward to integrate in the algorithm, it is numerically very expensive. In general, we expect any prior model

that is informed about the relative orientation of structures (as e.g. the Zeldovich approximation) to partly break the line-of-sight degeneracy or inhibit spurious mass sheets. Additional information could come from the clustering of galaxies. This would require the inference of at least one additional bias parameter (depending on the bias model), but would add information about the relative orientation of structure, also possibly breaking the line-of-sight and mass-sheet degeneracy.

In the current implementation the algorithm assumes a fixed set of cosmological parameters for both models. For future applications the parameter estimation could be made part of the algorithm by analyzing the joint posterior of the density distribution and cosmological parameters. The cosmology enters the density estimate through the theoretical matter power spectrum, the growth factor and the background evolution (that determines e.g. the redshift-distance relation). Lensing measurements are mostly sensitive to a combination of σ_8 and Ω_m. For two cosmological parameters only, sampling of the joint posterior would be feasible already with the current serialized implementation, for one could sample the $\sigma_8 - \Omega_m$ plane (e.g. on a grid) and evaluate the minimum of $P(\delta|\sigma_8, \Omega_m, d)$ for each parameter combination in parallel. Since the total information content of the data does not depend on the analysis method, a joint reconstruction of the 3D density field and cosmological parameters should yield competitive constraints to traditional analyses that make use of redshift binning.

ACKNOWLEDGMENTS

[1] M. Bartelmann and P. Schneider, Phys. Rep. **340**, 291 (2001).

[2] M. Kilbinger, Rep. Prog. Phys. **78**, 086901 (2015).

[3] H. Hoekstra and B. Jain, Annu. Rev. Nucl. Part. Sci. **58**, 99 (2008).

[4] A. Refregier, Annu. Rev. Astron. Astrophys. **41**, 645 (2003).

[5] F. Bernardeau, L. van Waerbeke, and Y. Mellier, Astron. Astrophys. **322**, 1 (1997).

[6] W. Hu, Astrophys. J. **522**, L21 (1999).

[7] P. Simon, L. J. King, and P. Schneider, Astron. Astrophys. **417**, 873 (2004).

[8] W. Hu, Phys. Rev. D **66**, 083515 (2002).

[9] D. Huterer, Phys. Rev. D **65**, 063001 (2002).

[10] Y.-S. Song and L. Knox, Phys. Rev. D **70**, 063510 (2004).

[11] Z. Ma, W. Hu, and D. Huterer, Astrophys. J. **636**, 21 (2006).

[12] D. J. Bacon, A. R. Refregier, and R. S. Ellis, Mon. Not. R. Astron. Soc. **318**, 625 (2000).

[13] L. Van Waerbeke, Y. Mellier, T. Erben, J. C. Cuillandre, F. Bernardeau, R. Maoli, E. Bertin, H. J. McCracken, O. Le Fèvre, B. Fort, M. Dantel-Fort, B. Jain, and P. Schneider, Astron. Astrophys. **358**, 30 (2000).

[14] N. Kaiser, G. Wilson, and G. A. Luppino, arXiv:astro-ph/0003338.

[15] D. M. Wittman, J. A. Tyson, D. Kirkman, I. Dell'Antonio, and G. Bernstein, Nature (London) **405**, 143 (2000).

[16] N. Scoville et al., Astrophys. J. Suppl. Ser. **172**, 38 (2007).

[17] C. Heymans et al., Mon. Not. R. Astron. Soc. **427**, 146 (2012).

[18] J. T. A. de Jong et al., The Messenger **154**, 44 (2013).

[19] J. A. Tyson, D. Wittman, I. Dell'Antonio, and A. Becker, V. Margoniner (DLS Team), Bull. Am. Astron. Soc. **33**, 1464 (2001).

[20] M. Jarvis et al., Mon. Not. R. Astron. Soc. **460**, 2245 (2016).

[21] L. Miller, T. D. Kitching, C. Heymans, A. F. Heavens, and L. van Waerbeke, Mon. Not. R. Astron. Soc. **382**, 315 (2007).

[22] T. D. Kitching, L. Miller, C. E. Heymans, L. van Waerbeke, and A. F. Heavens, Mon. Not. R. Astron. Soc. **390**, 149 (2008).

[23] L. Miller, C. Heymans, T. D. Kitching, L. van Waerbeke, T. Erben, H. Hildebrandt, H. Hoekstra, Y. Mellier, B. T. P. Rowe, J. Coupon, J. P. Dietrich, L. Fu, J. Harnois-Déraps, M. J. Hudson, M. Kilbinger, K. Kuijken, T. Schrabback, E. Semboloni, S. Vafaei, and M. Velander, Mon. Not. R. Astron. Soc. **429**, 2858 (2013).

[24] J. Zuntz, T. Kacprzak, L. Voigt, M. Hirsch, B. Rowe, and S. Bridle, Mon. Not. R. Astron. Soc. **434**, 1604 (2013).

[25] M. Gentile, F. Courbin, and G. Meylan, arXiv:1211.4847.

[26] M. D. Schneider, D. W. Hogg, P. J. Marshall, W. A. Dawson, J. Meyers, D. J. Bard, and D. Lang, Astrophys. J. **807**, 87 (2015).

[27] J. Benjamin, L. Van Waerbeke, C. Heymans, M. Kilbinger, T. Erben, H. Hildebrandt, H. Hoekstra, T. D. Kitching, Y. Mellier, L. Miller, B. Rowe, T. Schrabback, F. Simpson, J. Coupon, L. Fu, J. Harnois-Déraps, M. J. Hudson, K. Kuijken, E. Semboloni, S. Vafaei, and M. Velander, Mon. Not. R. Astron. Soc. **431**, 1547 (2013).

[28] C. Heymans and A. Heavens, arXiv:astro-ph/0310495.

[29] L. J. King and P. Schneider, Astron. Astrophys. **398**, 23 (2003).

[30] H. Hildebrandt, T. Erben, K. Kuijken, L. van Waerbeke, C. Heymans, J. Coupon, J. Benjamin, C. Bonnett, L. Fu, H. Hoekstra, T. D. Kitching, Y. Mellier, L. Miller, M. Velander, M. J. Hudson, B. T. P. Rowe, T. Schrabback, E. Semboloni, and N. Benítez, Mon. Not. R. Astron. Soc. **421**, 2355 (2012).

[31] R. Massey, J. Rhodes, A. Leauthaud, P. Capak, R. Ellis, A. Koekemoer, A. Réfrégier, N. Scoville, J. E. Taylor, J. Albert, J. Bergé, C. Heymans, D. Johnston, J.-P. Kneib, Y. Mellier, B. Mobasher, E. Semboloni, P. Shopbell, L. Tasca, and L. Van Waerbeke, Astrophys. J. Suppl. Ser. **172**, 239 (2007).

[32] M. J. Jee, J. A. Tyson, M. D. Schneider, D. Wittman, S. Schmidt, and S. Hilbert, Astrophys. J. **765**, 74 (2013).

[33] J. Benjamin, C. Heymans, E. Semboloni, L. Van Waerbeke, H. Hoekstra, T. Erben, M. D. Gladders, M. Hetterscheidt, Y. Mellier, and H. K. C. Yee, Mon. Not. R. Astron. Soc. **381**, 702 (2007).

[34] M. Kilbinger et al., Mon. Not. R. Astron. Soc. **430**, 2200 (2013).

[35] T. Schrabback et al., Astron. Astrophys. **516**, A63 (2010).

[36] H. Hildebrandt et al., Mon. Not. R. Astron. Soc. **465**, 1454 (2017).

[37] M. J. Jee, J. A. Tyson, S. Hilbert, M. D. Schneider, S. Schmidt, and D. Wittman, Astrophys. J. **824**, 77 (2016).

[38] C. Heymans et al., Mon. Not. R. Astron. Soc. **432**, 2433 (2013).

[39] T. Abbott et al. (Dark Energy Survey Collaboration), Phys. Rev. D **94**, 022001 (2016).

[40] M. Takada and B. Jain, Mon. Not. R. Astron. Soc. **348**, 897 (2004).

[41] E. Semboloni, T. Schrabback, L. van Waerbeke, S. Vafaei, J. Hartlap, and S. Hilbert, Mon. Not. R. Astron. Soc. **410**, 143 (2011).

[42] J. P. Dietrich and J. Hartlap, Mon. Not. R. Astron. Soc. **402**, 1049 (2010).

[43] L. Marian, S. Hilbert, R. E. Smith, P. Schneider, and V. Desjacques, Astrophys. J. **728**, L13 (2011).

[44] L. Marian, R. E. Smith, S. Hilbert, and P. Schneider, Mon. Not. R. Astron. Soc. **423**, 1711 (2012).

[45] L. Marian, R. E. Smith, S. Hilbert, and P. Schneider, Mon. Not. R. Astron. Soc. **432**, 1338 (2013).

[46] S. Hilbert, L. Marian, R. E. Smith, and V. Desjacques, Mon. Not. R. Astron. Soc. **426**, 2870 (2012).

[47] J. Liu, A. Petri, Z. Haiman, L. Hui, J. M. Kratochvil, and M. May, Phys. Rev. D **91**, 063507 (2015).

[48] T. Kacprzak et al. (DES Collaboration), Mon. Not. R. Astron. Soc. **463**, 3653 (2016).

[49] A. Heavens, Mon. Not. R. Astron. Soc. **343**, 1327 (2003).

[50] P. G. Castro, A. F. Heavens, and T. D. Kitching, Phys. Rev. D **72**, 023516 (2005).

[51] T. D. Kitching, A. F. Heavens, and L. Miller, Mon. Not. R. Astron. Soc. **413**, 2923 (2011).

[52] T. D. Kitching et al., Mon. Not. R. Astron. Soc. **442**, 1326 (2014).

[53] N. Kaiser and G. Squires, Astrophys. J. **404**, 441 (1993).

[54] W. Hu and C. R. Keeton, Phys. Rev. D **66**, 063506 (2002).

[55] A. N. Taylor, arXiv:astro-ph/0111605.

[56] P. Simon, Astron. Astrophys. **560**, A33 (2013).

[57] R. M. Szepietowski, D. J. Bacon, J. P. Dietrich, M. Busha, R. Wechsler, and P. Melchior, Mon. Not. R. Astron. Soc. **440**, 2191 (2014).

[58] T. A. Enßlin, M. Frommert, and F. S. Kitaura, Phys. Rev. D **80**, 105005 (2009).

[59] M. Greiner, D. H. F. M. Schnitzeler, and T. A. Enßlin, Astron. Astrophys. **590**, A59 (2016).

[60] E. Hubble, Astrophys. J. **79**, 8 (1934).

[61] P. Coles and B. Jones, Mon. Not. R. Astron. Soc. **248**, 1 (1991).

[62] M. C. Neyrinck, I. Szapudi, and A. S. Szalay, Astrophys. J. **698**, L90 (2009).

[63] M. Greiner and T. A. Enßlin, Astron. Astrophys. **574**, A86 (2015).

[64] F.-S. Kitaura, J. Jasche, and R. B. Metcalf, Mon. Not. R. Astron. Soc. **403**, 589 (2010).

[65] J. Jasche, F. S. Kitaura, C. Li, and T. A. Enßlin, Mon. Not. R. Astron. Soc. **409**, 355 (2010).

[66] A. Taruya, M. Takada, T. Hamana, I. Kayo, and T. Futamase, Astrophys. J. **571**, 638 (2002).

[67] L. Clerkin et al., Mon. Not. R. Astron. Soc. **466**, 1444 (2017).

[68] S. Hilbert, J. Hartlap, and P. Schneider, Astron. Astrophys. **536**, A85 (2011).

[69] G. M. Bernstein and R. Armstrong, Mon. Not. R. Astron. Soc. **438**, 1880 (2014).

[70] J. Alsing, A. Heavens, A. H. Jaffe, A. Kiessling, B. Wandelt, and T. Hoffmann, Mon. Not. R. Astron. Soc. **455**, 4452 (2016).

[71] A. Heavens, J. Alsing, A. Jaffe, T. Hoffmann, A. Kiessling, and B. Wandelt, arXiv:1602.05345.

[72] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Phys. Rep. **367**, 1 (2002).

[73] E. V. Linder and A. Jenkins, Mon. Not. R. Astron. Soc. **346**, 573 (2003).

[74] R. D. Blandford, A. B. Saust, T. G. Brainerd, and J. V. Villumsen, Mon. Not. R. Astron. Soc. **251**, 600 (1991).

[75] J. E. Gunn, Astrophys. J. **150**, 737 (1967).

[76] C. S. Kochanek, Mon. Not. R. Astron. Soc. **247**, 135 (1990).

[77] S. Seitz, P. Schneider, and J. Ehlers, Classical Quantum Gravity **11**, 2345 (1994).

[78] N. Kaiser, Astrophys. J. **439**, L1 (1995).

[79] C. Seitz and P. Schneider, Astron. Astrophys. **318**, 687 (1997).

[80] E. Krause and C. M. Hirata, Astron. Astrophys. **523**, A28 (2010).

[81] C. Shapiro and A. Cooray, J. Cosmol. Astropart. Phys. 03 (2006) 007.

[82] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Phys. Rep. **215**, 203 (1992).

[83] M. Selig, M. R. Bell, H. Junklewitz, N. Oppermann, M. Reinecke, M. Greiner, C. Pachajoa, and T. A. Enßlin, Astron. Astrophys. **554**, A26 (2013).

[84] D. Blas, J. Lesgourgues, and T. Tram, J. Cosmol. Astropart. Phys. 07 (2011) 034.

[85] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, SIAM J. Sci. Stat. Comput. **16**, 5 (1995).

[86] R. E. Angulo, V. Springel, S. D. M. White, A. Jenkins, C. M. Baugh, and C. S. Frenk, Mon. Not. R. Astron. Soc. **426**, 2046 (2012).

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

[88] Y. P. Jing, Astrophys. J. **620**, 559 (2005).