Accurate analytic model for the weak lensing convergence one-point probability distribution function and its autocovariance

Leander Thiele[1,2,*] J. Colin Hill[3,4] and Kendrick M. Smith[1]

[1]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*
[2]*Department of Physics, Princeton University, Princeton, New Jersey 08544, USA*
[3]*Department of Physics, Columbia University, 538 West 120th Street, New York, New York 10027, USA*
[4]*Center for Computational Astrophysics, Flatiron Institute, New York, New York 10003, USA*

The one-point probability distribution function (PDF) is a powerful summary statistic for non-Gaussian cosmological fields, such as the weak lensing (WL) convergence reconstructed from galaxy shapes or cosmic microwave background (CMB) maps. Thus far, no analytic model has been developed that successfully describes the high-convergence tail of the WL convergence PDF for small smoothing scales from first principles. Here, we present a halo-model formalism to compute the WL convergence PDF, building upon our previous results for the thermal Sunyaev-Zel'dovich field. Furthermore, we extend our formalism to analytically compute the covariance matrix of the convergence PDF. Comparisons to numerical simulations generally confirm the validity of our formalism in the non-Gaussian, positive tail of the WL convergence PDF, but also reveal the convergence PDF's strong sensitivity to small-scale systematic effects in the simulations (e.g., due to finite resolution). Finally, we present a simple Fisher forecast for a Rubin-Observatory-like survey, based on our new analytic model. Considering the $\{A_s, \Omega_m, \Sigma m_\nu\}$ parameter space and assuming a *Planck* CMB prior on A_s only, we forecast a marginalized constraint $\sigma(\Sigma m_\nu) \approx 0.08$ eV from the WL convergence PDF alone, even after marginalizing over parameters describing the halo concentration-mass relation. This error bar on the neutrino mass sum is comparable to the minimum value allowed in the normal hierarchy, illustrating the strong constraining power of the WL convergence PDF. We make our code publicly available at https://github.com/leanderthiele/hmpdf.

I. INTRODUCTION

While the cosmic microwave background (CMB) historically has been the driving force in cosmological parameter inference, we are now experiencing a proliferation in high-quality data from the late-time matter distribution. In contrast to the primary CMB, late-time fields are described by nonlinear clustering of matter, rendering the distribution of many relevant observables highly non-Gaussian. For such non-Gaussian fields, the problem of extracting all information contained therein is unsolved; while for Gaussian fields, such as the primary CMB, the power spectrum is an optimal summary statistic containing all the information, no such summary statistic is known in the non-Gaussian case.

One late-time field of interest is the weak lensing (WL) convergence. Weak gravitational lensing describes the deflection of light by the matter distribution, imparting a shear and magnification on the images of observed background galaxies or CMB fluctuations. The WL convergence is a redshift-weighted measure of the integrated matter density along the line of sight; thus, it is a powerful probe of the matter distribution.

Since in the course of the nonlinear gravitational clustering the matter distribution departs significantly from Gaussianity, appreciable amounts of information leak from the power spectrum into higher-order statistics. This motivated previous studies to consider parameter inference from such measures of non-Gaussianity, such as the WL skewness and bispectrum [1–8]. An alternative summary statistic is the one-point probability distribution function (PDF), which simply constitutes the histogram of WL convergence pixel values. Originally considered in the context of peak statistics [9,10], more recent studies have demonstrated that the WL convergence PDF can add significant constraining power in parameter inference, not only in the σ_8-Ω_m plane, but also on the neutrino mass sum (e.g., Refs. [11–13]).

In some respects WL shares similarities with the thermal Sunyaev-Zel'dovich (tSZ) effect, which describes the scattering of CMB photons by hot electrons residing mostly in massive halos. Because the tSZ signal is approximately

proportional to $M_{\text{halo}}^{5/3}$, the halo model allows an excellent description of the tSZ PDF. This fact was recently utilized in order to construct a semianalytic model for the tSZ PDF [14,15] (here "semi" accounts for the fact that the model contains some functions that are most accurately fixed by fitting to numerical simulations).

In this work, we demonstrate that the halo-model formalism developed in the tSZ context can be applied to the WL convergence PDF as well, with small modifications due to two complications: 1) in contrast to tSZ, the WL convergence signal does not include the additional $M_{\text{halo}}^{2/3}$ temperature bias, which brings the distribution closer to Gaussian and renders the halo model slightly less accurate; 2) furthermore, while the tSZ signal is strictly positive, the WL convergence receives negative contributions from underdense regions (voids). In contrast to previous works on the subject [6,16–18], our formalism is better suited to describe relatively large positive values of the convergence PDF (which are sourced by massive halos), while performing less well in the only mildly nonlinear regime and especially at negative convergences. As we will show, our formalism is also not very accurate when the convergence field is smoothed on large angular scales. Perturbative methods [1,2,19–23], and the large deviation statistics formalism developed in Refs. [24,25] are better suited in such a situation. In terms of physical input, our formalism is quite similar to the stochastic numerical method developed in Refs. [26–28].

Besides a theoretical model for the expected form of the WL PDF, in order to do parameter inference we also require a prescription for its statistical distribution. While this distribution is non-Gaussian and difficult to compute, a first step is the computation of the covariance matrix. In terms of practical applications, the covariance matrix can be useful if the PDF is sufficiently downsampled such that the likelihood can be computed in the Gaussian approximation, as was done in Ref. [14]; alternatively, non-Gaussian inference methods such as likelihood-free inference [29] can benefit from the covariance matrix as a starting point. In view of these potential applications, we generalize the halo-model formalism to compute not only the one- but also the two-point PDF, the latter being sufficient for computation of the covariance matrix.

The remainder of this paper is structured as follows. In Sec. II, we present the theoretical part of this work, starting from the general theory of weak gravitational lensing and proceeding to our halo-model formalism for the one- and two-point PDFs. There, we also discuss the modifications to the formalism in comparison to the tSZ case. In Sec. III, we present various results obtained with our formalism for the one-point PDF: a number of calculations intended to build up intuition on the WL convergence PDF, and comparisons to two independent sets of numerical simulations. In Sec. IV, we turn to the two-point PDF and the covariance matrix of the one-point PDF. We perform a null

test and compare the analytic covariance matrix to a large N-body simulation. In Sec. V, we utilize the previous results to produce a simple Fisher parameter forecast. We conclude in Sec. VI. Further analytic calculations useful in building up intuition are presented in Appendix A, some details on the numerical evaluation of the formalism are collected in Appendix B, and in Appendix C we discuss the validity of several approximations.

II. THEORY

A. Background

Gravitational lensing distorts and magnifies the shapes of distant sources (e.g., galaxies or CMB fluctuations) as a result of the projected gravitational potential of matter along the line of sight (LOS), including dark and baryonic matter. In the weak lensing limit, these effects are encoded in the lensing convergence field, $\kappa(\hat{\boldsymbol{n}})$:

$$\kappa(\hat{\boldsymbol{n}}) = \int_0^\infty dz \delta(\boldsymbol{x}(\chi(z)\hat{\boldsymbol{n}}, z)) W^\kappa(z), \quad (1)$$

where $\chi(z)$ is the comoving distance to redshift z, δ is the matter density fluctuation, $\delta(\boldsymbol{x}) \equiv (\rho(\boldsymbol{x}) - \bar{\rho})/\bar{\rho}$, and the lensing projection kernel is given by

$$W^\kappa(z) = \frac{3}{2}\Omega_m H_0^2 \frac{(1+z)}{H(z)}\frac{\chi(z)}{c}\int_z^\infty dz_s \frac{dn}{dz_s}\frac{(\chi(z_s) - \chi(z))}{\chi(z_s)}, \quad (2)$$

where dn/dz_s is the distribution of sources, normalized such that $\int dz dn/dz = 1$. Note that for CMB lensing, $dn/dz = \delta_D(z - z_*)$, where δ_D is the Dirac δ function and $z_* \approx 1100$ is the redshift of last scattering. For weak lensing due to galaxies, dn/dz is generally a more complicated function. Note that we have specialized to the case of a flat universe in Eqs. (1) and (2). For reference, the lensing convergence is related to the lensing potential, $\phi(\hat{\boldsymbol{n}})$, via $\kappa(\hat{\boldsymbol{n}}) = -\nabla^2\phi(\hat{\boldsymbol{n}})/2$ (where ∇^2 is the two-dimensional Laplacian on the sky), or $\kappa_\ell = \ell(\ell+1)\phi_\ell/2$ in harmonic space.

Given a 3D halo density profile [e.g., the Navarro-Frenk-White (NFW) profile [30]], we can define the lensing convergence profile, $\underline{\kappa}(\boldsymbol{\theta}, M, z)$ for a halo of mass M at redshift z:

$$\underline{\kappa}(\boldsymbol{\theta}, M, z) = \Sigma_{\text{crit}}^{-1}(z)\int_{\text{LOS}} \rho\left(\sqrt{l^2 + d_A^2|\boldsymbol{\theta}|^2}, M, z\right)dl, \quad (3)$$

where $\rho(\boldsymbol{r}, M, z)$ is the halo density profile, $d_A(z)$ is the angular diameter distance to redshift z, and $\Sigma_{\text{crit}}(z)$ is the critical surface density (in physical units here) for lensing at redshift z assuming a source distribution dn/dz,

$$\Sigma_{\mathrm{crit}}^{-1}(z) = \frac{4\pi G\chi(z)}{c^2(1+z)} \int_z^\infty dz' \frac{(\chi(z') - \chi(z))}{\chi(z')} \frac{dn}{dz'} \quad (4)$$

$$= \frac{8\pi G}{3\Omega_m H_0^2} \frac{H(z)}{c(1+z)^2} W^\kappa(z). \quad (5)$$

For a spherically symmetric density profile, the convergence profile is azimuthally symmetric, i.e., $\underline{\kappa}(\boldsymbol{\theta}, M, z) = \underline{\kappa}(\theta, M, z)$. For the NFW density profile, analytic forms exist for the convergence profile. One subtlety, however, is the nonconvergence of the enclosed mass in the NFW profile as $r \to \infty$, which thus necessitates a radial cutoff in calculations using this profile.

B. WL PDF in the halo model

In Ref. [15], an analytic approach based on the halo model was constructed to describe the one-point PDF of the tSZ field, building on a simpler model presented in Ref. [14]. In particular, the effects of halo overlaps along the LOS and halo clustering, which were neglected in Ref. [14], were included in Ref. [15]. However, the expressions derived in Ref. [15] are more broadly applicable to the one-point PDF of *any* (projected) cosmic field that can be modeled in a halo-based approach. The halo model approach is very accurate for the tSZ field as this field is heavily dominated by contributions from massive halos (e.g., Refs. [31–33]), due to the temperature dependence of the tSZ signal. A primary goal of this paper is to assess the accuracy of this model for other cosmic fields, in particular the WL convergence field. Thus, as a first step to model the WL convergence PDF, we can simply use the expressions from Ref. [15], but with the y (tSZ) profile replaced by the κ profile defined in Eq. (3). The rest of the formalism derived in that work then goes through unchanged.

For completeness and ease of reference in later sections, we include the derivation of the one-point κ PDF in this formalism here. In some places, algebraic manipulations are omitted for brevity; we refer the interested reader to Ref. [15] for full details.

1. One-halo term

We refer to the (differential) κ one-point PDF as $P(\kappa)$. Considering a bin spanning $[\kappa_i, \kappa_{i+1}]$, we define the binned version of the PDF as

$$p_i = \int_{\kappa_i}^{\kappa_{i+1}} d\kappa P(\kappa). \quad (6)$$

The fundamental concept underlying the model developed in Refs. [14,15] is that p_i quantifies the sky fraction subtended by κ values in the range $[\kappa_i, \kappa_{i+1}]$. For an individual spherically symmetric halo with an azimuthally symmetric projected κ profile $\underline{\kappa}(\theta)$, this sky fraction is

simply the area in the annulus between $\theta(\kappa_i)$ and $\theta(\kappa_{i+1})$, where $\theta(\kappa_i)$ is the angular distance from the center of the profile to the radius where $\kappa(\theta) = \kappa_i$. If we then assume that halos are sufficiently rare that they never overlap on the sky, the one-point PDF is simply given by adding up the annular area contributions from all halos:

$$p_i = \int dz dM \frac{\chi^2}{H} \frac{dn}{dM} \pi(\underline{\theta}^2(\kappa_i) - \underline{\theta}^2(\kappa_{i+1})) + \delta_i(1 - F_{\mathrm{halos}}),$$
$$(7)$$

where $dn(M, z)/dM$ is the halo mass function (i.e., the number of halos of mass M at redshift z per unit mass and comoving volume), $\underline{\theta}(\kappa, M, z)$ is the inverse function of $\underline{\kappa}(\theta, M, z)$, F_{halos} is the total sky area subtended by all halos (assuming some radial cutoff for the halo profiles), δ_i is unity if $\kappa = 0$ lies in the bin and zero otherwise, and redshift and mass dependences have been suppressed in the equation for compactness. Equation (7) is only accurate in the limit in which halos do not overlap on the sky; moreover, it neglects effects due to the clustering of halos. While these assumptions are (moderately) accurate for the tSZ field, they are not accurate for the WL convergence field.

We thus seek a more general approach, in which these limiting assumptions are discarded. The basic ideas and results for our improved formalism were presented for the tSZ field in Ref. [15]; here, we adapt the formalism to the WL convergence field and introduce a more compact notation. Our goal is to compute the one- and two-point PDFs, $P(\kappa_a)$ and $P(\kappa_a, \kappa_b; \phi)$, where ϕ is the angular separation between two sky locations at which we measure the two convergence values κ_a and κ_b. It will be convenient to work in Fourier space, introducing

$$P_a \equiv \int d\kappa_a e^{i\lambda_a \kappa_a} P(\kappa_a), \quad (8)$$

$$P_{ab} \equiv \int d\kappa_a d\kappa_b e^{i(\lambda_a \kappa_a + \lambda_b \kappa_b)} P(\kappa_a, \kappa_b; \phi), \quad (9)$$

where we have abbreviated the notation for conciseness.

We will separate the PDFs into a one- and a two-halo term, writing $P = P^{1\mathrm{h}} P^{2\mathrm{h}}$. In this section we compute the one-halo term, i.e., we ignore the clustering of halos for the moment. We introduce two further pieces of notation: we denote the projected halo mass function

$$n \equiv n(M, z) = \frac{\chi^2(z)}{H(z)} \frac{dn(M, z)}{dM}, \quad (10)$$

which gives the expected number of halos per unit mass and redshift interval in a unit solid angle; furthermore we introduce the quantities

$$\hat{K}_a^{(\theta)} \equiv e^{i\underline{\kappa}(\theta)\lambda_a} - 1, \tag{11}$$

$$K_a^{(\ell)} \equiv \int_\theta \hat{K}_a^{(\theta)} J_0(\ell\theta), \tag{12}$$

where $\underline{\kappa}$ are the convergence profiles, $\int_\theta \equiv \int 2\pi\theta d\theta$ and we have again suppressed the mass and redshift dependences.

First, we consider a narrow bin of width $dMdz$ in mass-redshift space, such that halo overlaps can be neglected for the infinitesimal number of halos in this bin. For a given realization of the halo distribution, we have at the arbitrarily chosen origin $\mathbf{0}$

$$e^{i\lambda_a\kappa(\mathbf{0})} = 1 + \sum_h \hat{K}_a^{(\theta_h)}, \tag{13}$$

where the sum runs over all halos in the given mass-redshift bin and θ_h is their separation from the origin. Thus, we find for the one-point PDF

$$\begin{aligned} P_a^{\mathrm{1h}} &= \langle e^{i\lambda_a\kappa(\mathbf{0})}\rangle_h \\ &= 1 + dMdzn \int_{\hat{\mathbf{n}}} \hat{K}_a^{(\hat{\mathbf{n}})} \\ &= 1 + dMdznK_a^{(0)}, \end{aligned} \tag{14}$$

where the subscript h indicates that we are averaging over realizations of the halo distribution. Likewise, we find for the two-point PDF

$$\begin{aligned} P_{ab}^{\mathrm{1h}} &= \langle e^{i\lambda_a\kappa(\mathbf{0})} e^{i\lambda_b\kappa(\boldsymbol{\phi})}\rangle_h \\ &= 1 + dMdzn\left[K_a^{(0)} + K_b^{(0)} + \int_{\hat{\mathbf{n}}} \hat{K}_a^{(\hat{\mathbf{n}})} \hat{K}_b^{(\hat{\mathbf{n}}-\boldsymbol{\phi})}\right] \\ &= P_a^{\mathrm{1h}} P_b^{\mathrm{1h}}\left[1 + dMdzn \int_\ell K_a^{(\ell)} K_b^{(\ell)} J_0(\ell\phi)\right], \end{aligned} \tag{15}$$

where we have introduced $\int_\ell \equiv \int \ell d\ell/2\pi$. Note that to the order considered so far, terms of the form $1 + dMdzA$ can equally well be written as $\exp dMdzA$. Under the Born approximation, the convergence is an additive quantity. Thus, the complete PDFs can be obtained by convolution, which is equivalent to multiplication in Fourier space:

$$P_a^{\mathrm{1h}} = \exp \int_{M,z} K_a^{(0)}, \tag{16}$$

$$\frac{P_{ab}^{\mathrm{1h}}}{P_a^{\mathrm{1h}} P_b^{\mathrm{1h}}} = \exp \int_{M,z,\ell} K_a^{(\ell)} K_b^{(\ell)} J_0(\ell\phi), \tag{17}$$

where we have introduced $\int_M \equiv \int dMn$, $\int_z \equiv \int dz$ for brevity. As we have demonstrated in Ref. [15], expanding the exponentials to first order leads to the approximate model from Ref. [14]. In this sense, terms of order n^p in the

Taylor expansion of P^{1h} can be interpreted as describing overlaps of p halos along the line of sight.

2. Two-halo term

The two-halo term arises from the dependence of halo density on the underlying long-wavelength linear density field, which at sky location $\hat{\mathbf{n}}$ and redshift z we denote by

$$\delta(\hat{\mathbf{n}}) \equiv \delta_{\mathrm{lin}}(\hat{\mathbf{n}}, z). \tag{18}$$

The change in the halo density can to a first approximation be written as

$$n \to n(\hat{\mathbf{n}}) = n[1 + b\delta(\hat{\mathbf{n}})], \tag{19}$$

where $b \equiv b(M, z)$ is the linear halo bias. In order to compute the two-halo term in the PDFs, we proceed in two steps: first, we compute the correction to the one-halo term in a given realization δ, obtaining P^δ, and then we perform the average over realizations. We note that in contrast to Ref. [15] we denote by P^δ only the multiplicative correction factor to the PDF. Using the substitution in Eq. (19), the required correction factors can be written down immediately:

$$P_a^\delta(\mathbf{0}) = \exp \int_z \alpha_a\delta(\mathbf{0}), \tag{20}$$

$$P_{ab}^\delta(\mathbf{0}) = \exp \int_z \alpha_a\delta\left(-\frac{\boldsymbol{\phi}}{2}\right) + \alpha_b\delta\left(\frac{\boldsymbol{\phi}}{2}\right) + \beta_{ab}\delta(\mathbf{0}), \tag{21}$$

where we have introduced

$$\alpha_a \equiv \int_M bK_a^{(0)}, \tag{22}$$

$$\beta_{ab} \equiv \int_{M,\ell} bK_a^{(\ell)} K_b^{(\ell)} J_0(\ell\phi). \tag{23}$$

We take the opportunity to point out a subtlety here: because we are working with a fixed realization δ at the moment, isotropy is broken and the PDFs depend explicitly on sky location. Thus, we need to assume that the linear density field δ varies sufficiently slowly that the halo model formalism we have been assuming still makes sense. As we will explicitly demonstrate in Appendix A 1, this assumption is equivalent to the statement that the linear matter correlation function vanishes on scales similar to typical halo radii. Now all that is left to do is to perform the average over realizations of δ, giving the clustering corrections

$$P^{\mathrm{2h}} = \langle P^\delta\rangle_\delta. \tag{24}$$

We remind the reader that for light-cone integrals $f_i(\hat{\mathbf{n}}) = \int_z W_i\delta(\hat{\mathbf{n}})$ the Limber approximation allows us to write

$$\langle f_i(\mathbf{0}) f_j(\boldsymbol{\phi}) \rangle_\delta = \int_z H \zeta(\phi) W_i W_j, \qquad (25)$$

where $\zeta(\phi)$ is the redshift-dependent line-of-sight projected matter correlation function, which in terms of the linear matter power spectrum can be written as

$$\zeta(\phi) = \int \frac{k \, dk}{2\pi} P_{\text{lin}}(k, z) J_0(k\phi\chi(z)). \qquad (26)$$

Utilizing the Limber approximation and the identity $\langle e^x \rangle = e^{\langle x^2 \rangle / 2}$ valid for Gaussian distributed x, we obtain

$$P_a^{2\text{h}} = \exp \int_z \frac{H\zeta(0)}{2} \alpha_a^2, \qquad (27)$$

$$\frac{P_{ab}^{2\text{h}}}{P_a^{2\text{h}} P_b^{2\text{h}}} = \exp \int_z H \left[\alpha_a \alpha_b \zeta(\phi) + \frac{1}{2} \beta_{ab}^2 \zeta(0) \right. $$
$$\left. + \beta_{ab}(\alpha_a + \alpha_b) \zeta\left(\frac{\phi}{2}\right) \right]. \qquad (28)$$

This concludes the main theoretical part of this work. We would like to point out two features of the formalism presented here. 1) Although we have assumed that halos are only described by their mass and redshift, one could consider additional labels c (e.g., related to halo environment or formation history). This would introduce a c dependence in the halo mass function n and add further integrations over c on equal footing with the mass integrations (the redshift integrations are special because of the simplification introduced by the Limber approximation). 2) The general formalism applies to any N-point PDF. For example, for the three-point PDF one would have to compute three-index objects γ_{ijk} that are analogous to α_i and β_{ij}. However, the "momentum" labels ℓ would turn into vectorial quantities now, which presumably complicates the required integrations considerably.

C. WL PDF contributions from nonvirialized matter and voids

The expressions above only account for the contributions to the WL κ PDF due to matter in halos ("virialized matter"). For the Compton-y field, this approximation was very accurate due to the temperature dependence of the tSZ signal, which strongly biases the y field toward electrons in massive halos. For κ, this approximation is less accurate, as the WL convergence field is an unbiased tracer of the matter distribution. Moreover, in contrast to Compton-y, there are negative-signal regions in the κ field (i.e., projected underdensities in the matter distribution). We thus require some method to treat both the "matter outside of halos" and "voids."

With regard to the matter outside of halos, one option would be the following. We assume that the rest of the κ map (not accounted for by the halo-based model) is purely

a Gaussian random field (GRF) that is uncorrelated with the virialized-halo part of the κ map. We can compute the variance of this GRF (call it the "residual variance") by simply using the halo model to compute the angular power spectrum $C_\ell^{\kappa\kappa}$ and truncating the halo model integrals at the M values above which the explicit profile-based calculation is used (so that the variance contributed by those objects is not double counted). Alternatively, we could compare the variance of the halo-model PDF with the variance obtained from the "Halofit" fitting function [34,35], and extract the residual variance from the difference of these quantities. After computing the residual variance, a Gaussian PDF of this width can be convolved with the halo-based PDF to obtain the final κ PDF.

We implement both approaches described above. We find the changes in the PDF with respect to the unmodified halo-model-only result to be extremely minor. The first approach suffers from the problem that the halo model becomes ill defined in the low-mass regime relevant to this calculation, and thus the small residual variance is quite uncertain. On the other hand, we frequently find the variance computed from Halofit to be smaller than the variance deduced from the halo-model PDF, invalidating the basic assumption. Thus, in this work we do not include either of these ideas for incorporating convergence contributions from matter outside of halos, instead using only the halo model described in the previous section. If exact results are necessary, we recommend testing stability with respect to the lower limit in mass integrations.

With regard to the negative-κ voids, a simple prescription is based on the fact that the mean $\langle \kappa \rangle = 0$ by construction. Thus, to a first approximation, we simply compute the halo-model PDF as described above and then shift it such that the physical constraint is enforced. This idea clearly fails to provide an accurate description of the negative-κ tail of the PDF. It also is not immediately obvious that it leads to good predictions for the positive-κ tail, primarily because it does not take into account void-halo correlations. Thus, comparison to numerical simulations will be crucial in assessing the accuracy of this simple approximation.

III. RESULTS: ONE-POINT PDF

Before discussing various results obtained with the formalism developed in the previous section, we mention several choices for fitting functions and numerical settings. For the halo mass function dn/dM and the linear bias b, we use the fitting functions of Ref. [36]. We describe halos with an NFW profile, using the concentration-mass relation of Ref. [37]. We use the Colossus package [38] for calculations in the halo model, and CAMB [39] and CLASS [40,41] (with Halofit corrections from Refs. [34,35]) for matter and WL convergence power spectra. As mentioned before, the NFW profile necessitates a radial cutoff; we find the WL convergence PDF to be nearly independent of this cutoff and choose it at $r_{\text{max}} = 1.6 r_{\text{vir}}$, where r_{vir} is computed according

to Ref. [42] and the prefactor of 1.6 is chosen to obtain good agreement between the halo-model-computed and Halofit-computed WL convergence power spectra. Unless otherwise stated, all halo masses are given in terms of $M \equiv M_{200c}$, and we choose integration limits $11 \leq \log_{10} M/h^{-1} \, M_\odot \leq 16$, so that the PDF is very well converged (as will be shown in Sec. III C). For simplicity, we specialize to a Dirac-δ distribution of source galaxies, $dn/dz = \delta_D(z - z_s)$, at a single source redshift z_s. In order to incorporate pixelization effects, we convolve the convergence profiles with a window function

$$W_{1\mathrm{pt},\ell}^{\mathrm{pix}} = \frac{4}{\pi} \int_0^{\pi/4} d\varphi \, \mathrm{sinc}(\cos(\varphi)\ell a) \mathrm{sinc}(\sin(\varphi)\ell a), \quad (29)$$

where a is half the pixel side length. This prescription is only approximate: there is no precise method to incorporate quadratic pixels while keeping the convergence profiles azimuthally symmetric. However, the error incurred is negligible for the purposes of this work. Note that it would also be irrelevant in any real parameter inference, since for realistic shape noise levels the Wiener filter one would apply to the map (as described in Sec. V) cuts off harmonic-space modes before the pixelization effect becomes relevant.

Having developed the analytic formalism in the previous section, we now proceed to discuss various results in the following subsections. In Sec. III A, we examine the effect of corrections (overlaps and clustering) on the WL one-point PDF. In Sec. III B, we compare our model's predictions to results from two sets of cosmological N-body simulations. In Sec. III C, we disentangle the contributions from different halo mass and redshift intervals to the PDF. In Sec. III D, we discuss the dependence of the WL PDF on cosmological and concentration model parameters.

A. Impact of overlaps and clustering

As discussed before, our formalism utilizes a halo-model-based framework similar to the tSZ PDF calculation in Ref. [14], with the crucial difference that we incorporate corrections arising from halo clustering and overlaps along the line of sight (as developed in Ref. [15]). Naturally, we should examine the size of these corrections. In Fig. 1, we plot the exact result from our formalism in red, while the result neglecting halo clustering is represented in green. As noted in Ref. [14], the result neglecting overlaps is only applicable if the minimum halo mass contribution is relatively large; thus, we plot two versions of the PDF neglecting both overlaps and clustering with different minimum masses in solid/dashed blue. These results are shown for four different choices of source redshift ranging from $z_s = 1$ to 2.5.

We see that both clustering and overlaps constitute substantial corrections, of order a few 10%. This is in marked contrast to the conclusion we drew in the tSZ case, where the clustering effect was subdominant and did not
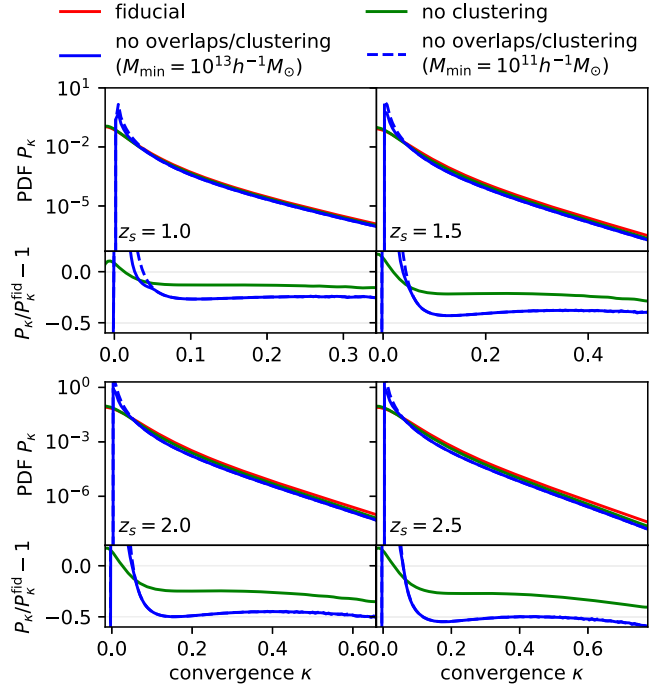

FIG. 1. The effects of clustering and overlaps for four different source redshifts z_s. The PDF neglecting clustering (green) seems to show similar behavior to what was observed in the tSZ case [15], but the clustering effect is much more pronounced. Note that the shift to $\langle \kappa \rangle = 0$ does not make sense in the no-overlaps formalism: by assumption, the sky is infinitely large so the condition is automatically satisfied (this corresponds to the divergence at $\kappa = 0$).

exceed a few percent. As we shall see in Sec. III C, the convergence PDF receives larger contributions from low-mass halos at lower redshifts in comparison to the tSZ PDF, which explains the more pronounced clustering contribution. We note the unphysical divergence of the results neglecting overlaps near $\kappa = 0$, which is removed by the improved model presented in this work.

B. Comparison to numerical simulations

We compare results from our formalism to WL convergence PDFs extracted from two different sets of N-body simulations, namely MassiveNuS[1] (Sec. III B 1) [43] and Takahashi et al.[2] (hereafter T17) (Sec. III B 2) [44]. Both of these simulations provide ray-traced WL convergence maps. We provide further details on each simulation analysis below.

1. Comparison to MassiveNuS

We analyze a set of 10^4 ray-traced weak lensing convergence maps from the MassiveNuS simulation suite,

---

[1]http://astronomy.nmsu.edu/aklypin/SUsimulations/MassiveNuS/.
[2]http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/.

which are derived from a set of N-body simulations that include dark matter and an approximate treatment of massive neutrinos via a linear response method [43]. Each convergence map is 3.5×3.5 deg^2 with 512^2 square pixels, corresponding to a pixel side length of 0.41 arcmin. The effect of the pixel window is treated in our analytic calculations via Eq. (29). The simulations include maps for a wide range of cosmological parameters, but we consider only the fiducial simulated cosmology, with parameters given by $\Omega_m = 0.3$, $\Omega_b = 0.046$, $h = 0.7$, $A_s = 2.1 \times 10^{-9}$, $n_s = 0.97$, and zero neutrino mass ($\sigma_8 = 0.8523$ is a derived parameter). We use these parameters in all analytic calculations that compare to MassiveNuS.

We analyze convergence maps constructed with δ-function source planes at various redshifts. We consider κ values ranging from $[-5\sigma_\kappa, 20\sigma_\kappa]$, where σ_κ is the variance of the κ field measured from the full simulation set for each source redshift option. The bins are linearly spaced with width $\sigma_\kappa/5$. In all PDF measurements, we enforce the constraint that $\langle \kappa \rangle = 0$.

To ensure that the simulation results are robust to cosmic variance fluctuations resulting from the small map size, we also analyze a set of 10^5 convergence maps that were produced for covariance matrix estimation at the fiducial cosmology, using additional, independent N-body simulations. We subdivide this large set into ten subsets of 10^4 maps each, and verify that any fluctuations in the measured κ PDFs across the subsets are negligible.

In Fig. 2, we plot a comparison between the fiducial analytic one-point PDF (solid red) and the PDF measured in MassiveNuS (black), for four different source redshifts, $z_s = 1, 1.5, 2, 2.5$. While the discrepancies at negative κ are entirely expected, the large differences in the positive-κ tail are not expected given our intuition that the halo model should perform very well in this regime. In order to explain these discrepancies, we plot WL convergence power spectra in Fig. 3. We observe good agreement between the Halofit result (blue) and the fiducial result (solid red) computed using the standard halo model expressions (e.g., Ref. [45]). On the other hand, MassiveNuS lacks power for $\ell \gtrsim 10^3$. This is likely related to small-scale resolution effects in the simulation, presumably a combination of finite mass resolution and force softening (e.g., Refs. [46,47]). As a simple test of whether these resolution effects can explain the discrepancies seen in the one-point PDFs, we calibrate a k-space filter with which we smooth the NFW density profiles such that the resulting convergence power spectra match the MassiveNuS results. We find the filter

$$W(k) = [1 + (kR)^2]^{-0.7}, \qquad (30)$$

where $R = 0.17 h^{-1}$ Mpc comoving, to yield relatively good agreement. The resulting power spectra are plotted
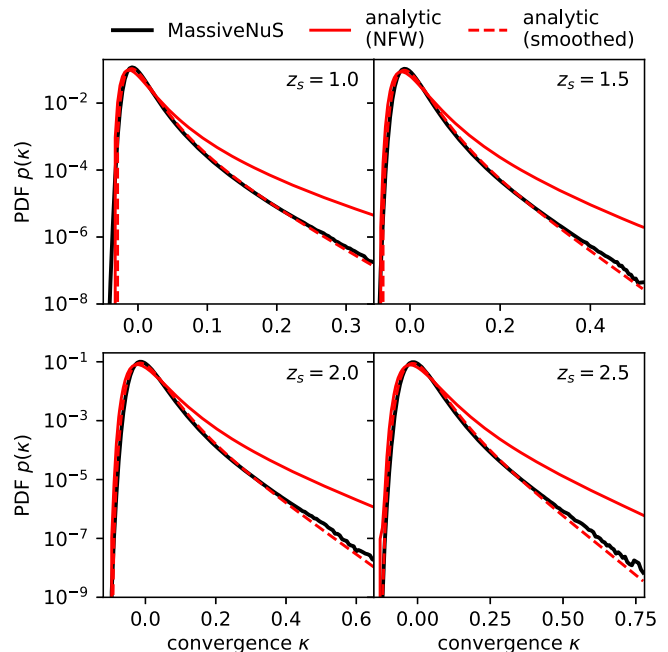

FIG. 2. Comparison between the MassiveNuS simulations results (black) [43] and our analytic model (red) for the WL convergence one-point PDF, for four different source redshifts. Solid red: fiducial result of our model; dashed red: the result obtained by smoothing the NFW density profiles with a k-space filter calibrated on the convergence power spectra measured in MassiveNuS, as described in the text and illustrated in Fig. 3. This filter captures the non-negligible effects due to the finite resolution of the simulation. Here, as well as in the other plots in which we show simulation data, the error bars would be invisible by eye.

in dashed red in Fig. 3. Having calibrated the filter $W(k)$ on the power spectra, we then compute the resulting one-point PDF, plotted in dashed red in Fig. 2. We observe much better agreement now. In the part of the PDF that is relatively close to Gaussian the analytic result matches the simulations almost exactly. Small discrepancies remain in the high-κ tail, which is not surprising since our smoothing filter was calibrated solely on the two-point correlation function, while the one-point PDF in the tail depends strongly on higher-order correlation functions. Thus, we conclude that the resolution effects leading to a lack of power at high ℓ are likely responsible for the high-κ discrepancy between our model and the simulation result, rather than a deficiency of our halo model formalism.

2. Comparison to T17

We analyze a set of 108 full-sky, ray-traced weak lensing convergence maps from T17 [44], which are derived from a suite of large dark-matter-only N-body simulations. The maps are provided in HEALPix [48] format at resolution $N_{\rm side} = 8192$, corresponding to an approximate pixel scale
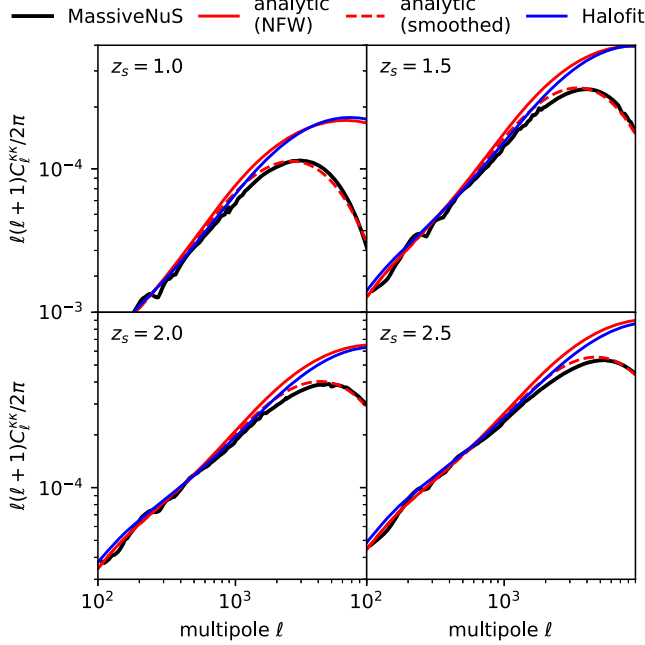
FIG. 3. Convergence power spectra. Blue is the Halofit result, while solid red is our fiducial analytic result from the halo model. The MassiveNuS power spectra (black) show a deficiency in power at $\ell \gtrsim 10^3$, likely due to resolution effects. The dashed red lines are analytic power spectra obtained by smoothing the NFW density profiles, as described in the text, in order to mimic the resolution effect.
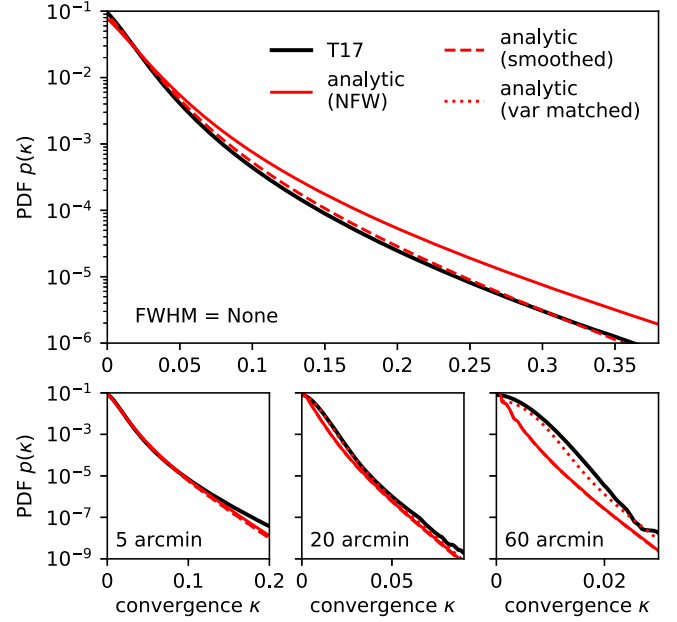

FIG. 4. Comparison between the T17 simulation results (black) [44] and the analytic result (red) for the convergence one-point PDF, for four different Gaussian smoothing scales labeled by their FWHM in the panels. Solid red: fiducial result of our model; dashed red: the result obtained by smoothing the NFW density profiles; dotted red: the result obtained by convolving with a Gaussian such that the final variance matches the variance measured in the simulation. Note that in the top and bottom left panels the analytic variance is slightly larger than the one measured in the simulations, and thus no dotted line is plotted.

of 0.43 arcmin. The parameters used in the simulations are $\Omega_m = 0.279$, $\Omega_b = 0.046$, $h = 0.7$, $\sigma_8 = 0.82$, $n_s = 0.97$, and zero neutrino mass. We match these parameters in all analytic calculations that compare to T17.

In our comparison to the T17 simulations we focus on the effect of smoothing the convergence maps. Thus, we only produce results for a single source redshift, $z_s = 1.0334$, but apply Gaussian filters of varying full-width half-maximum (FWHM) values to the convergence maps. We consider κ values ranging from $[-5\sigma_\kappa, 20\sigma_\kappa]$, where σ_κ is the variance of the κ field measured from the full simulation set for each choice of smoothing filter. The bins are linearly spaced with width $\sigma_\kappa/5$. In all PDF measurements, we enforce the constraint that $\langle\kappa\rangle = 0$.

The results are plotted in Fig. 4. Black curves are simulation results, while red is the analytic result. Focusing on the upper panel [where no additional smoothing has been applied to the maps apart from the inherent pixelization, which is treated via Eq. (29)], we again observe a discrepancy between the fiducial analytic result (solid red) and the simulation. As a further test to our hypothesis that the discrepancy can be explained with simulation resolution effects, we again construct a k-space smoothing filter for the NFW profiles. We choose the redshift-dependent filter

$$W(k; z) = \frac{1}{1 + kR(z)},$$
$$R(z) = 0.055h^{-1} \text{ Mpc} \times [\log(1 + z) + 0.07]. \quad (31)$$

The function $R(z)$ was chosen to give a good fit to the softening lengths employed in the T17 simulation, with some adjustments of the prefactor (our R is about 10% smaller than the softening length). We find the resulting one-point PDF to be largely independent of the precise functional form chosen for $W(k; z)$, as long as it decreases steadily to ~ 0.5 when $kR \sim 1$. The natural correspondence between the smoothing scale R and the softening length is a further indication that simulation resolution effects are responsible for the observed discrepancies in the one-point PDF.

A second purpose of this section is to evaluate how well our formalism can describe the PDF of convergence maps smoothed with a Gaussian filter. As we explicitly show in Appendix A 2, as the smoothing scale increases the PDF becomes closer to Gaussian (as is physically expected) and receives larger contributions from the two-halo term. These facts imply that our formalism, which is most accurate for the non-Gaussian parts of the PDF that are dominated by massive halos, is expected to perform worse.

Thus, comparison to simulated maps is a useful test of the domain of validity. This is plotted in the lower three panels of Fig. 4. We observe that, as should be expected, the difference between solid and dashed red becomes negligible as the smoothing scale increases. However, our formalism does not recover the simulation PDFs very well. As the smoothing scale increases, the PDF receives more and more contributions from nonvirialized matter, which is not included in our halo-model formalism. One attempt to solve this problem is to convolve the analytic PDF with a Gaussian such that the resulting variance $\langle \kappa^2 \rangle$ is equal to the variance measured in the simulation. The results of this procedure are plotted as dotted red lines. Although the agreement (naturally) gets better, it is still far from perfect. One possible explanation is the fact that we do not describe the negative-κ regime accurately enough in our formalism, and upon convolution with a relatively broad Gaussian this inaccuracy leaks into the positive-κ part.

Thus, we draw two conclusions from our comparison with the T17 simulations. 1) We have presented further evidence that convergence one-point PDFs measured in simulations are susceptible to large errors due to small-scale resolution effects. Thus, the discrepancies observed in Figs. 2 and 4 do not invalidate our analytic formalism. 2) Our approach seems inadequate to generate accurate predictions for the PDF of convergence maps smoothed over scales larger than a few arcminutes. Perturbative methods are likely better suited to compute theoretical predictions in this regime.

C. Halo mass and redshift contributions

We now proceed to build up some physical intuition on the dominant contributions to the convergence PDF. Since we label halos solely by their mass and redshift, we disentangle the contributions that different mass and redshift intervals give to the PDF. In Fig. 5, we plot heat maps of the mass and redshift contributions to the PDF for source redshifts $z_s = 1$ and 2.5. Each of the three rows represents a different bin of the PDF at comparable values of the convergence κ in units of the respective standard deviation σ_κ. Each pixel in the individual heat maps corresponds to the fraction of the final value of the PDF if all masses and redshifts smaller than or equal to the one corresponding to the pixel are included (thus, the upper right corner has by definition a value of one in each heat map). Note that overlaps make the interpretation of these heat maps somewhat complicated, in particular for low values of κ. In terms of redshift evolution, we can clearly see the growth of structure modulated by the lensing kernel. In terms of mass contributions, the intuitive picture that higher values of κ are sourced by more massive objects is confirmed.

D. Parameter dependence

In this section, we discuss the dependence of the convergence one-point PDF on the cosmological model
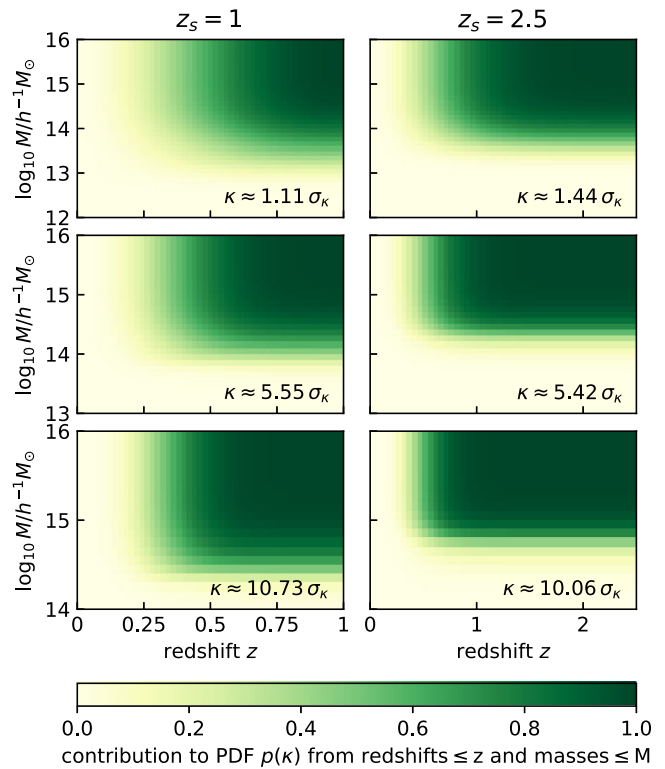

FIG. 5. Cumulative mass and redshift contributions to the PDF for source redshifts 1 and 2.5 (columns). The rows represent different values of the convergence in units of the standard deviation. Note the different vertical scales in different rows.

as well as the halo concentration-mass relation. The results presented here are a prerequisite for the Fisher forecast in Sec. V, but are also useful as a means to build up intuition. We show results for a single source redshift $z_s = 1$.

We choose our fiducial cosmology as $h = 0.7$, $\Omega_m = 0.3$, $\Omega_b = 0.046$, $A_s = 2.1 \times 10^{-9}$, $n_s = 0.97$, and $\Sigma m_\nu = 0.06$ eV. We assume the normal hierarchy for the neutrino masses. Following Ref. [37], we write the concentration-mass relation as

$$c(M, z) = A \left(\frac{M}{M_0} \right)^B \left(\frac{1+z}{1+z_0} \right)^C, \quad (32)$$

where we choose $M_0 = 10^{14.5} h^{-1} M_\odot$, $z_0 = 0.35$, so as to break the leading degeneracy between the three parameters A, B, C (cf. Fig. 5).

Our results are shown in Fig. 6, with varied cosmology in the left panel and varied concentration model in the right panel. The solid/dashed lines generally represent parameter variations by $\pm 10\%$, except for the neutrino mass sum, where solid corresponds to 0.12 eV and dashed to massless neutrinos. Note that the residual curves for Ω_m and A were shrunk by the stated factors to increase readability. With regard to varied cosmological parameters, with a fixed fractional change Ω_m has by far the strongest influence on
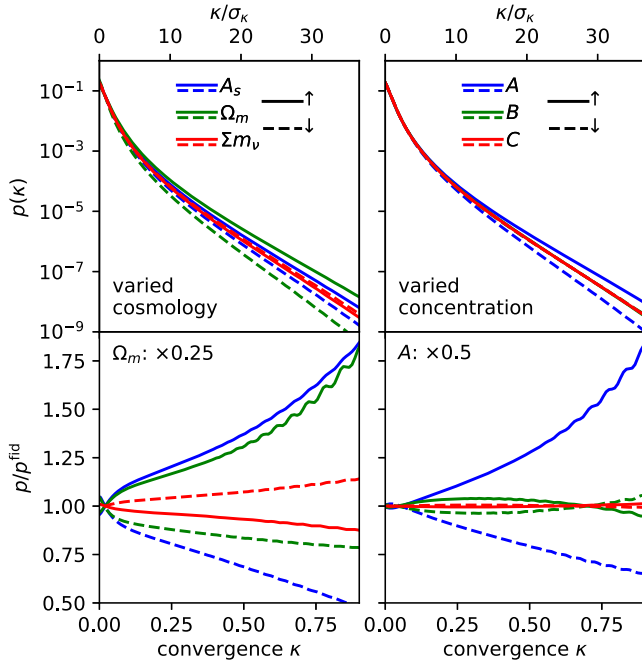
FIG. 6. Dependence of the convergence one-point PDF on cosmology (left panel) and concentration model (right panel). For an explanation of the parameters A, B, C see the text. Solid/dashed lines represent in-/decreases of the respective parameter by 10%, except for the neutrino mass sum Σm_ν, for which the fiducial model is 0.06 eV while solid/dashed represent 0.12 and 0 eV respectively.
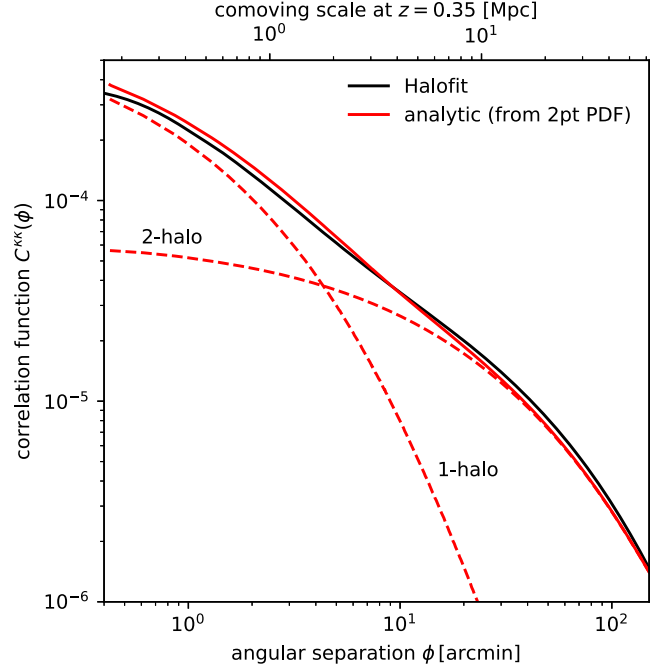

FIG. 7. Comparison between the two-point correlation function reduced from our formalism for the two-point PDF and the Halofit computation, for sources at $z_s = 1$. Small discrepancies on large scales are related to the choice of radial cutoff or minimum halo mass. The discrepancies on small scales are likely due to inaccuracies of Halofit in the very nonlinear regime.

the PDF. From the slight shape difference in the residual curves, we can hope that the degeneracy between neutrino mass sum and other parameters is not too large. With regard to the concentration model, the amplitude has by far the largest effect, while the variation with halo mass and redshift are of minor importance. This finding rests on the fact that the mass and redshift integrands are relatively sharply peaked around M_0, z_0 for our simple Dirac-δ source distribution (cf. Fig. 5), a condition that may not be met in the case of a more realistic source distribution.

IV. RESULTS: TWO-POINT PDF

As numerical evaluation of the two-point PDF is somewhat involved, we collect some useful simplifications in Appendix B. In order to validate the analytic formalism for the two-point PDF, and by extension the covariance matrix, we perform two tests, detailed in the following.

A. Two-point correlation function

In Fig. 7, we plot a comparison between the (isotropic) two-point correlation function obtained in our formalism with the result predicted by the Halofit fitting function (for sources at $z_s = 1$). Note that while the one-halo term can be directly computed in our formalism, the two-halo term is simply inferred as the difference of the other two curves.

The top x axis in the figure indicates the comoving scale at the approximate redshift where the main contribution to the one-point PDF is sourced. The small discrepancies on large scales are due to the relatively large minimum halo mass chosen in the computation. We confirmed that the precise choice of this cutoff has no effect on the covariance matrix. Note that the relatively good agreement on large angular scales is a good indication that we are treating the halo clustering effect correctly (in fact, this is the first direct validation, since the comparison of the one-point PDF to numerical simulations is complicated by simulation resolution effects). However, as we show in Appendix A 1, the correlation function is not sensitive to all terms in the two-point PDF. The discrepancies on small scales are likely stemming from the Halofit rather than the halo model side, since we do not expect Halofit to be accurate on these rather nonlinear scales. (Another possible explanation would be deviations from the NFW profile on small scales in the numerical simulations used to calibrate Halofit.) We point out that this type of plot is useful to identify numerical instability in a calculation of the covariance matrix.

B. Covariance matrix

Given the one- and two-point PDFs, the covariance matrix of the one-point PDF can be computed as

$$\mathrm{Cov}_{ab} = \frac{1}{N_{\mathrm{pix}}} \sum_{\phi} [P_{ab}(\phi) - P_a P_b]. \qquad (33)$$

Here, the indices label the κ bins, and the sum runs over all pixel separations in a given map. N_{pix} is the number of pixels in the map, i.e., related to the sky coverage. In practice, it is accurate enough to explicitly perform the sum over pixels for the smallest separations and approximate the remaining summation by an integral. Note that $P_{ab}(\phi = 0) = P_a \delta_{ab}$.

We measure the covariance matrix using the 108 full-sky T17 simulations and compare it to our analytic result. As discussed in Sec. III B 2, resolution issues appear to cause a discrepancy between the analytic result and the T17 (and also MassiveNuS) simulations. Thus, in order to make the comparison as direct as possible, we compute the covariance matrix with the smoothed NFW profiles described before, corresponding to the dashed red line in the top panel of Fig. 4.

In the bottom panel of Fig. 8, we show a comparison of the correlation matrices (i.e., $\mathrm{Cov}_{ab}/\sqrt{\mathrm{Cov}_{aa}\mathrm{Cov}_{bb}}$). We observe good agreement in the general structure. The analytic model is able to recover the transition to anti-correlation at low κ (albeit displaced by about one bin).
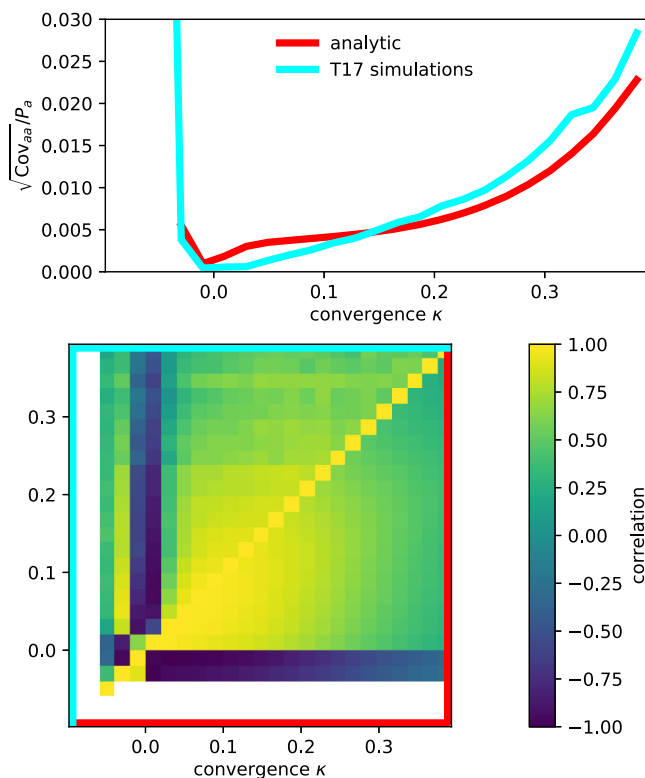

FIG. 8. Comparison of the analytic and T17 covariance matrices. Top panel: diagonal elements, rescaled by the respective one-point PDFs. Bottom panel: correlation matrices (T17 in upper left triangle, analytic in lower right triangle; the color codes are as in the top panel).

The simulations appear to have more correlations in the high-κ bins, although these bins are noisy and dominated by rare events in the simulations. There appears to be a step-like feature in the simulation covariance matrix at $\kappa \sim 0.23$, transitioning rather suddenly from high to low correlation. The analytic model does not predict such a feature, and since the halo model should work well at this relatively large convergence, we are inclined to trust the model more than the simulations.

In the top panel of Fig. 8, we show a comparison between the diagonal elements (we divide out the respective one-point PDFs). While the model recovers the diagonal elements to better than an order of magnitude (and to within $\sim 10\%$ for $\kappa \gtrsim 0.1$), there are systematic differences that cannot be explained by noise in the simulations. Given the limitations of our model with regard to nonvirialized matter, we concede that the discrepancies at low κ are likely due to failure of the model. However, for $\kappa \gtrsim 0.1$, we would expect the halo model to perform well, while, as we have already seen in Sec. III B 2, the simulations are susceptible to resolution issues. Although we have tried to take these into account by smoothing the NFW density profiles with a filter calibrated on the one-point PDF, it is likely that the covariance matrix depends differently on this filter and thus systematic discrepancies are still to be expected. This also highlights the possible dangers in purely simulation-based inference procedures, which could be biased by such resolution effects.

V. FISHER FORECAST

In this section, we present a simple Fisher matrix parameter forecast using the WL convergence PDF. We assume a Rubin-Observatory-like survey with sky coverage of $20\,000$ deg^2, pixel size $\Omega_{\mathrm{pix}} = (0.41\,\mathrm{arcmin})^2$ and source density $n_{\mathrm{gal}} = 45$ arcmin^{-2}. The latter number is taken from Ref. [11]; in contrast to this work we make the simplifying assumption of a Dirac-δ source distribution at $z_s = 1$. (Note that the inclusion of tomographic information would further improve the forecast error bars computed here.) Our fiducial cosmological model is the same as in Sec. III D.

In order to upweight the cosmological signal with respect to the shape noise, we convolve the convergence profiles with a Wiener filter, constructed as

$$F_\ell = C_\ell^{\kappa\kappa}/(C_\ell^{\kappa\kappa} + N_\ell), \qquad (34)$$

where we compute the convergence power spectrum using Halofit and the noise power spectrum $N_\ell = 0.3^2/n_{\mathrm{gal}}$ is flat. Applying the Wiener filter is crucial in optimizing the sensitivity of the one-point PDF, since it is a real-space statistic.

The shape noise in the filtered maps has the correlation function

$$\zeta_{\kappa,\text{sn}}(\phi) = \sum_{\ell} \frac{\ell + 1/2}{2\pi} N_{\ell} F_{\ell}^2 W_{2\text{pt},\ell}^{\text{pix}} P_{\ell}(\cos\phi), \qquad (35)$$

where $W_{2\text{pt},\ell}^{\text{pix}}$ is the pixel window function,

$$W_{2\text{pt},\ell}^{\text{pix}} = \frac{4}{\pi} \int_0^{\pi/4} d\varphi [\text{sinc}(\cos(\varphi)\ell a)\text{sinc}(\sin(\varphi)\ell a)]^2, \qquad (36)$$

where a is half the pixel side length. The "noisy" one-point PDF and covariance matrix are computed as

$$p_{\text{sn},a} = (\mathcal{G}_1[P_i])_a, \qquad (37)$$

$$\text{Cov}_{\text{sn},ab} = \sum_{\phi} [(\mathcal{G}_2^{(\phi)}[P_{ij}(\phi)])_{ab} - (\mathcal{G}_1[P_i])_a (\mathcal{G}_1[P_j])_b], \qquad (38)$$

where $\mathcal{G}_n[\cdot]$ indicates convolution with an n-dimensional Gaussian. Here, $\mathcal{G}_1$ has variance $\zeta_{\kappa,\text{sn}}(0)$ and $\mathcal{G}_2^{(\phi)}$ has the covariance matrix

$$C_{\text{sn}}^{(\phi)} = \begin{pmatrix} \zeta_{\kappa,\text{sn}}(0) & \zeta_{\kappa,\text{sn}}(\phi) \\ \zeta_{\kappa,\text{sn}}(\phi) & \zeta_{\kappa,\text{sn}}(0) \end{pmatrix}. \qquad (39)$$

In Fig. 9, we plot the same PDFs already shown in Fig. 6, this time including the shape noise contribution. We bin the PDFs into bins of width $\Delta\kappa = 1.65 \times 10^{-3}$, and quote the number of pixels expected in the model survey. Again, the parameter variations are 10% for all parameters except the neutrino mass sum, for which the variation is 100% (dashed is massless neutrinos, solid is $\Sigma m_\nu = 0.12$ eV). Note that, for clarity, we rescaled some of the curves in the bottom row of Fig. 9 by the given amounts. It is interesting to see that the Wiener filter not only serves the purpose of minimizing the noise contribution, but also increases the sensitivity of the PDF on cosmology in comparison to the concentration model (compare Fig. 6, which has the same parameter variations). This can be understood as a consequence of the suppression of small-scale power which makes the exact shape of the convergence profiles (and thus the concentration model) less relevant, while the main dependence on the cosmological model comes through the halo mass function. This will be crucial for the parameter forecast.

We compute the Fisher matrix as

$$F_{ab} = \frac{\partial \boldsymbol{p}^T}{\partial \theta^a} \text{Cov}^{-1} \frac{\partial \boldsymbol{p}}{\partial \theta^b} + \frac{1}{2} \text{trCov}^{-1} \frac{\partial \text{Cov}}{\partial \theta^a} \text{Cov}^{-1} \frac{\partial \text{Cov}}{\partial \theta^b}, \qquad (40)$$

where $\boldsymbol{p}$ is the binned one-point PDF and θ is the parameter vector (indexed by a, b). Contrary to conventional wisdom,
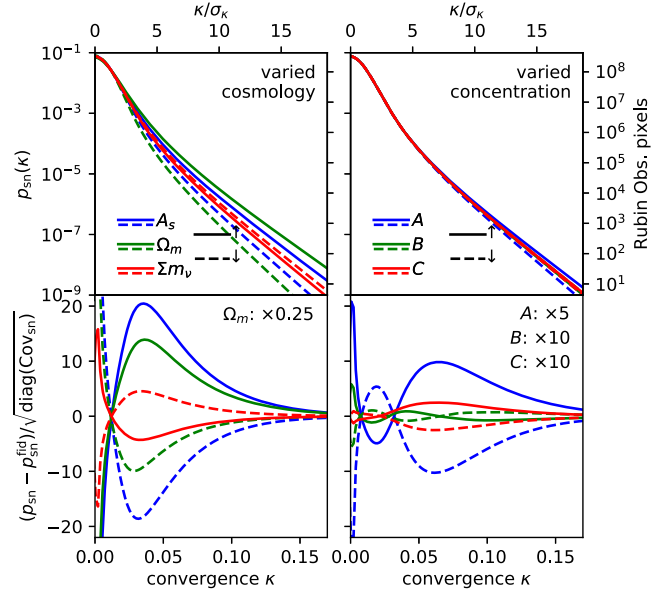

FIG. 9. The convergence one-point PDFs, as a function of parameter variations, including shape noise of $\sigma_{\kappa,\text{sn}} = 0.2$. The plotting conventions are the same as in Fig. 6; however, since in this plot we applied a Wiener filter to the convergence profiles the convergence values on the κ axis are not directly comparable. As in Fig. 6, all parameter variations are 10%, except for Σm_ν for which it is 100%. The bottom row shows the ratio of the residuals to the square-root of the diagonal elements of the covariance matrix.

the second term in the Fisher matrix is not always negligible and changes parameter constraints by a few 10%. In the data vector, we include the PDF in the interval $\kappa \in [-0.03, 0.17]$; we find the constraints to be relatively independent of the choice of binning. It is worth noting that we are including some negative-κ part of the PDF, which is not well described by our model. However, due to the Wiener filter and noise convolution it is rather challenging to cleanly exclude this regime, and as we will argue below we do not believe that keeping the uncertain values invalidates the major conclusions from our forecast.

We consider six free parameters, namely $\{A_s, \Omega_m, \Sigma m_\nu\}$ on the cosmology side and $\{A, B, C\}$ in the parametrization of the concentration model from Ref. [37]. For the concentration model, we choose the mass and redshift pivots described at the end of Sec. III D (in the end we rescale A to the original value).

In Fig. 10, we plot four different Fisher forecasts, differing solely by the priors we place on the concentration model and the scalar fluctuation amplitude A_s. Black includes no priors at all, magenta includes the error bars on A, B, C from Ref. [37] as diagonal Gaussian priors, green includes the CMB prior from Ref. [49] on A_s, and orange includes both priors. For clarity, in Table I we quote the fractional 1σ constraints (in percent) on the three cosmological parameters as well as the concentration
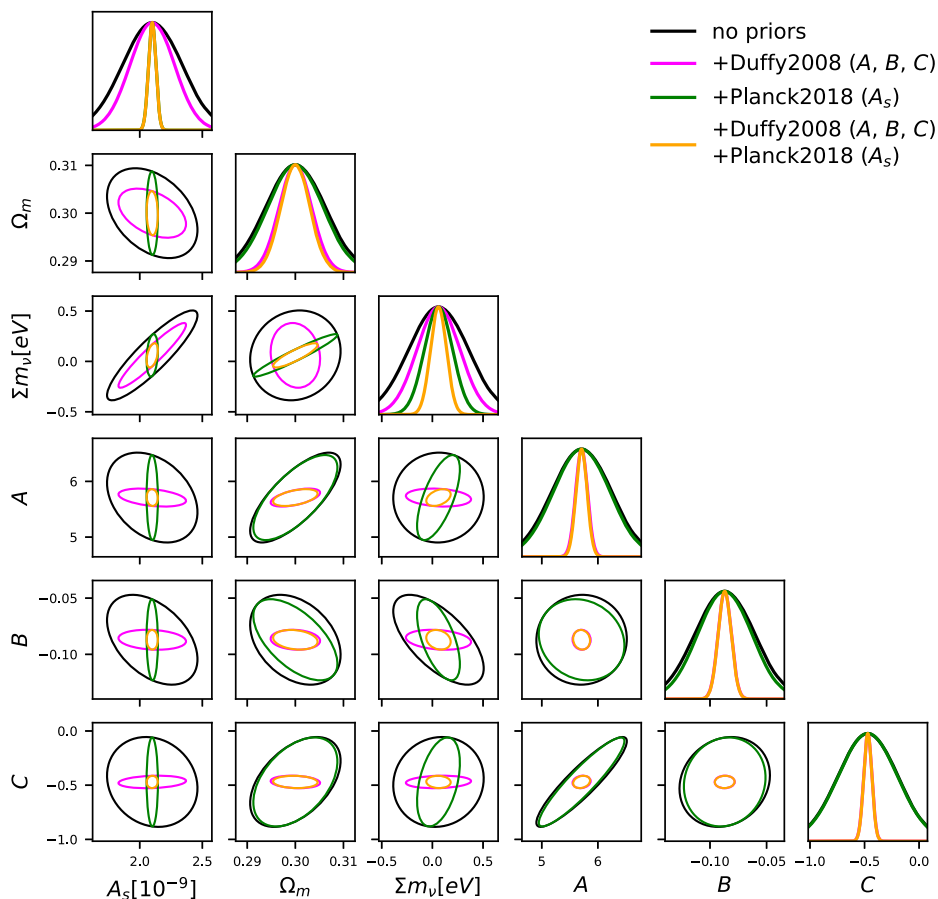
FIG. 10. Four Fisher forecasts for parameter constraints from the weak lensing one-point PDF. We are assuming Gaussian shape noise as described in the text, Rubin Observatory sky coverage of $20\,000$ deg^2, and a source density of 45 arcmin^{-2} in a Dirac-δ distribution at redshift $z_s = 1$. The ellipses are 1σ confidence intervals. Black: both cosmology and concentration model are completely inferred from the data; magenta: including simulation priors from Ref. [37] on the concentration model; green: including CMB prior from Ref. [49] on A_s; orange: including both priors on the concentration model and A_s.

model parametrization for the different choices of priors. We observe that if priors on the concentration model as well as A_s are included, our results suggest that the WL convergence PDF alone can provide an error bar on Σm_ν comparable to the minimum possible neutrino mass sum. Including tomographic information and/or the full WL convergence power spectrum would only improve these constraints.

TABLE I. 1σ constraints on cosmological and concentration model parameters, for four different choices of priors. The numbers are relative to the fiducial value, in percent.

	$\sigma(X)/X_{\rm fid}[\%]$					
	A_s	Ω_m	Σm_ν	A	B	C
No priors	11	2.1	490	9.3	30	58
+Duffy2008	8.5	1.1	350	1.9	6.7	8.1
+*Planck*2018 (A_s)	1.4	1.9	230	8.8	27	58
+Duffy2008 + *Planck*2018 (A_s)	1.4	1.0	132	1.7	6.6	8.0

Our simple Fisher forecast has several limitations:

(1) We are assuming a Gaussian likelihood even though we know that the one-point PDF is a non-Gaussian statistic. Unfortunately, the full likelihood for this observable is not yet known. Evidence of a small bias when assuming a Gaussian likelihood was seen for the tSZ PDF in Ref. [14]. However, a Gaussian likelihood was used and shown to be unbiased for the WL PDF in Ref. [11], albeit with the caveat that high-κ bins were removed, which Gaussianizes the statistic. The PDF observable could be a useful opportunity to apply new methods in likelihood-free inference (e.g., Refs. [29,50]).

(2) We work in the Fisher approximation; however, Markov chain Monte Carlo results from Ref. [11] indicate that this is not a bad approximation.

(3) Our analytic covariance matrix is likely not exact for small values of κ; however, our results from Sec. IV B indicate that the formalism tends to overestimate the covariance matrix in this regime,

which implies that in this respect our forecast parameter constraints are conservative.

(4) Our formalism is unable to make exact predictions for the negative-κ tail. However, this loss of detail also is likely to imply that our forecast is conservative. Including full information from the negative-κ region would only improve the constraints found here.

(5) We make the simplifying assumption of a Dirac-δ source distribution in a single bin. Accounting for a realistic spread in the source distribution would likely not significantly affect the constraints; moreover, Ref. [11] has demonstrated that tomography with multiple source bins has the potential to significantly improve the constraints.

(6) It is not clear that the Wiener filter employed is optimal. One possibility would be to consider the PDFs of maps smoothed on various scales simultaneously. This approach should recover some of the scale-dependent information contained in the full N-point functions that is lost when compressing to the zero-lag one-point PDF. The cross-covariance between the PDFs should be easy to compute in a simple extension of the formalism presented above (all terms in our model for the covariance matrix are symmetric under interchange of two convergence profiles, the modification should amount to breaking that symmetry by filtering them with two different kernels).

(7) We do not consider any cosmology dependence of the concentration model (although this is likely small), and our prior on the concentration model parameters may be optimistic, particularly given baryonic effects on the small-scale matter distribution.

The last point implies that a more robust understanding of the concentration-mass relation and halo density profiles in general would be extremely beneficial for parameter inference from the WL PDF, similar to the WL power spectrum.

Our forecast is similar in setup to the simulation-based one presented in Ref. [11]. However, there are a few key differences: 1) they do not make the simplifying assumption of a Dirac-δ source distribution; 2) they filter the convergence maps with a different ℓ-space filter; 3) they are able to use more of the negative-κ regime; 4) they (naturally) cannot marginalize over small-scale modeling uncertainties; 5) they do not include noise correlations, which we find to have an appreciable effect. As a consequence of these differences, we find the agreement between our results and theirs to be satisfactory. For a rough qualitative comparison, we consider our result with the concentration-model prior (i.e., the second row in Table I), where we have $\sigma(10^9 A_s) = 0.18$, $\sigma(\Omega_m) = 0.003$, $\sigma(\Sigma m_\nu) = 0.21$ eV. We find that including the effect of noise correlations gives about a factor of 2 degradation in constraints, which is approximately offset by increasing the

source number density by a factor of 4. Thus, for the purposes of this rough comparison, we choose the turquoise contours in Fig. 3 of Ref. [11] as reference (these do not include noise correlations and have a source density of 13.25 arcmin^{-2}). Approximating the posterior as Gaussian, we read off $\sigma(10^9 A_s) = 0.1$, $\sigma(\Omega_m) = 0.006$, $\sigma(\Sigma m_\nu) = 0.12$ eV. Thus, our halo-model-only forecast reproduces these simulation-derived constraints to within a factor of 2. Note, however, the difference in orientation of ellipses involving Ω_m.

As a final result of this section, we explore the constraints' dependence on the choice of the minimum and maximum convergence values in the data vector $\boldsymbol{p}$. We consider the three-parameter figure of merit,

$$\text{FOM} = (\det F^{-1}_{\text{cosmo}})^{-1/3}, \qquad (41)$$

where F_{cosmo} is the sub-block of the Fisher matrix corresponding to the cosmological parameters $\{A_s, \Omega_m, \Sigma m_\nu\}$. We work with the maximum set of priors, corresponding to the orange lines in Fig. 10. A heat map of this quantity as a function of minimum and maximum cutoff is shown in Fig. 11. Again, we emphasize that the "convergence" values quoted there are related through the Wiener filter to the physical convergence, making the interpretation somewhat difficult. The first conclusion from Fig. 11 is that the minimum cutoff κ_{min} is rather important for the constraining power. Second, for the minimum cutoff chosen in our analysis above, $\kappa_{\text{min}} = -0.03$, the information content of the positive-κ tail is essentially saturated
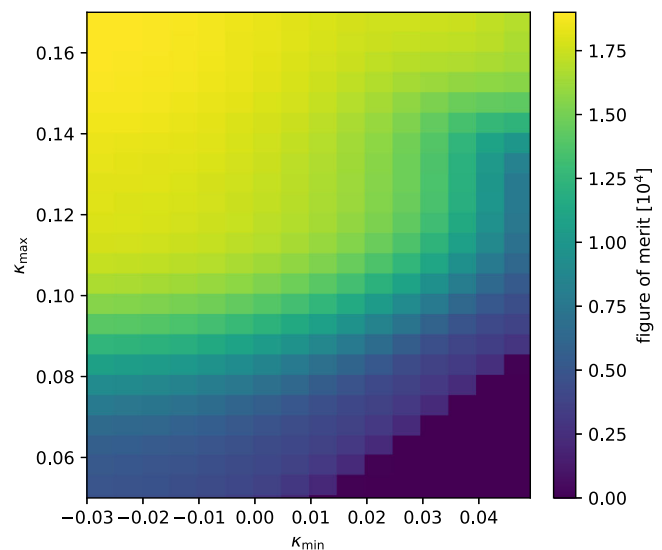

FIG. 11. Dependence of the normalized three-parameter figure of merit, as defined in Eq. (41), on the minimum/maximum κ cutoff. We observe that the negative-κ part of the PDF contains substantial information, and that our fiducial choice of $\kappa_{\text{max}} = 0.17$ extracts essentially all of the information content (assuming $\kappa_{\text{min}} = -0.03$).

with our choice of $\kappa_{max} = 0.17$. In fact, using a somewhat smaller value of κ_{max} could serve to Gaussianize the likelihood for the WL PDF, evidently without a significant loss in constraining power.

VI. CONCLUSIONS

We have developed a halo-model-based formalism for the weak lensing convergence one-point and two-point PDFs (and, by extension, the covariance matrix of the one-point PDF). The strengths of our model lie in its superior speed compared to simulations, the ability to explicitly marginalize over small-scale uncertainties (parametrized through the concentration model), and its interpretability.

As expected on physical grounds, the accuracy of our model is highest in the positive-convergence tail. We have shown that, in this regime, discrepancies in the one-point PDF with respect to numerical simulations are explained by simulation resolution issues and do not invalidate our method. It may be argued that as soon as the convergence map is smoothed on a sufficient scale, the simulation resolution effects would be less severe. However, at the wave numbers at which the MassiveNuS simulations start to show appreciable power deficiencies, the Wiener filter employed in this work still assumes values of order 0.1; thus, smoothing on a reasonable scale appears not sufficient to neutralize the small-scale issues in the simulations.

On the other hand, in the negative convergence regime and for large smoothing scales our formalism is less accurate. Alternative approaches are likely better suited for accurate theoretical predictions in these regimes.

Validation of the covariance matrix was found to be challenging, and discrepancies with respect to the simulations remain over a range of convergence values. However, discrepancies at low κ are likely irrelevant in any real analysis due to the dominance of the shape noise contributions there. The smaller discrepancies at high κ could simply be due to resolution effects in the simulations, as we already demonstrated for the one-point PDF itself.

Using our formalism, we have performed a Fisher forecast in the $\{A_s, \Omega_m, \Sigma m_\nu\}$ parameter space for a Rubin-Observatory-like survey. We have found that the convergence PDF alone could provide a 1σ error bar on Σm_ν that is comparable to the minimum neutrino mass sum allowed from oscillation experiments, if a CMB prior on A_s and simulation priors on the concentration-mass relation are included. Our results are in good agreement with previous simulation-based forecasts, and could be generated in a fraction of the time. We have also presented arguments why the several limitations of our simple forecast are likely to render it conservative, i.e., a more sophisticated analysis would probably find a further improvement in constraints (although this gain could be negated by the inclusion and marginalization of systematic errors in the measurement of convergence values). Our work demonstrates that an analytic approach to non-Gaussian WL statistics is feasible for upcoming surveys, at least in terms of the statistical constraining power. Tests for biases will necessitate end-to-end simulation analyses, which are beyond the scope of this work.

We believe that a comprehensive model for the weak lensing convergence one-point PDF and its covariance matrix would be most accurate if it combines different approaches. For example, one could imagine taking simulation results for the negative-convergence and Gaussian part of the PDF, while the positive-convergence tail is generated with our formalism. A desirable side effect of this method could be that smaller simulation volumes are required in order to sample the quasilinear part of the density field.

Possible extensions of our model could involve the use of compensated density profiles (e.g., Ref. [51]) (which could help improve accuracy near the Gaussian peak of the PDF), the effective halo model approach from Ref. [52], and the inclusion of voids.

APPENDIX A: SOME ANALYTIC CALCULATIONS

1. Proof that we recover the power spectrum

We can show quite easily that the two-point function is given by

$$C(\phi) = [\partial_{ab}^2 \log P_{ab}]_{\lambda_a = \lambda_b = 0}, \tag{A1}$$

where $\partial_a \equiv -i\partial/\partial\lambda_a$. Note that the one-point factors in the two-point PDF give no contribution, since their logarithm is a sum of functions that depend only on λ_a or λ_b. Since the Fourier-space PDF P_{ab} is a product of one- and two-halo terms, the correlation function becomes a sum.

The one-halo term is given by Eq. (17):

$$\log P_{ab}^{1h} \supset \int_{M,z,\ell} K_a^{(\ell)} K_b^{(\ell)} J_0(\ell\phi). \tag{A2}$$

Performing the differentiation, we obtain

$$C^{\text{1h}}(\phi) = \int_{\ell} J_0(\ell\phi) \int_{M,z} \left[\int_{\theta} \underline{\kappa}(\theta) J_0(\ell\theta) \right]^2, \quad \text{(A3)}$$

which is equivalent to the standard expression

$$C^{\text{1h}}(\ell) = \int_{M,z} |\tilde{\kappa}_{\ell}|^2. \quad \text{(A4)}$$

The two-halo term is given by Eq. (28):

$$\log P_{ab}^{\text{2h}} \supset \int_z H \left[\alpha_a \alpha_b \zeta(\phi) + \frac{1}{2}\beta_{ab}^2 \zeta(0) \right. $$
$$\left. + \beta_{ab}(\alpha_a + \alpha_b)\zeta\left(\frac{\phi}{2}\right) \right]. \quad \text{(A5)}$$

Since $\alpha_a = 0$, $\beta_{ab} = \partial_a \beta_{ab} = 0$ when $\lambda_a = \lambda_b = 0$, only the first term in the square brackets survives, and we get

$$C^{\text{2h}}(\phi) = \int_z H\zeta(\phi) \left[\int_{M,\theta} b\underline{\kappa}(\theta) \right]^2. \quad \text{(A6)}$$

Transforming to conjugate space, this gives

$$C^{\text{2h}}(\ell) = \int_z \frac{H}{\chi^2} P_{\text{lin}}(\ell/\chi, z) \left[\int_{M,\theta} b\underline{\kappa}(\theta) \right]^2. \quad \text{(A7)}$$

We see that this is not exactly equal to the usual halo model calculation, which would have an extra Bessel function in the square brackets. The difference, as remarked above, arises from the fact that we approximate the linear overdensity field as approximately constant over the typical size of a halo, so that the linear power spectrum becomes negligible whenever the argument $\ell\theta$ of the neglected Bessel function would become appreciable.

2. Large smoothing limit

In this section we briefly discuss how smoothing the convergence field with a Gaussian filter of large aperture affects the one-point PDF. Denoting the smoothing scale by σ, we have for the smoothed convergence profiles

$$\kappa_{\sigma}(\theta) = \text{Gaussian}_{\sigma}(\theta) * \underline{\kappa}(\theta), \quad \text{(A8)}$$

which, after inserting the conjugate space expressions, leads to

$$\kappa_{\sigma}(\theta) = \int_{\theta',\ell} \underline{\kappa}(\theta') J_0(\ell\theta) J_0(\ell\theta') e^{-\ell^2\sigma^2/2}. \quad \text{(A9)}$$

In general, the ratio θ'/σ will attain its maximum when θ' is comparable to the projected halo radius. Of course, this varies with halo mass and redshift, but it is still reasonable to formally introduce a scale $\hat{\theta}$ that characterizes typical halo extents. Since we assume σ to be large, we can expand in $\hat{\theta}/\sigma$.

The zero-order term in the expansion parameter $\hat{\theta}/\sigma$ is given by

$$\kappa_{\sigma}^{(0)}(\theta) = \sigma^{-2} e^{-\theta^2/2\sigma^2} \bar{\underline{\kappa}}, \quad \text{(A10)}$$

where we have introduced

$$\bar{\underline{\kappa}} \equiv \int_{\theta} \underline{\kappa}(\theta). \quad \text{(A11)}$$

It simply measures the total amount of signal in a single halo, which is then smoothed into a Gaussian by the σ filter.

We now compute the nth-order cumulants k_n. Similarly to the power-spectrum calculation performed in the previous section, the cumulants are related to derivatives of $\log P$, so that the cumulants naturally split into one- and two-halo terms:

$$k_n^{\text{1h}} = [\partial_a^n \log P_a^{\text{1h}}]_{\lambda_a=0}$$
$$= \frac{2\pi}{n} \sigma^{2-2n} \int_{M,z} \bar{\underline{\kappa}}^n, \quad \text{(A12)}$$

$$k_n^{\text{2h}} = [\partial_a^n \log P_a^{\text{2h}}]_{\lambda_a=0}$$
$$= \sigma^{4-2n} \int_z \frac{H\zeta(0)}{2} \sum_{m=1}^{n-1} \binom{n}{m} \frac{2\pi}{m} \frac{2\pi}{n-m}$$
$$\times \left[\int_M b\bar{\underline{\kappa}}^{n-m} \right] \left[\int_M b\bar{\underline{\kappa}}^m \right]. \quad \text{(A13)}$$

Thus, we see that the scalings with the expansion parameter are

$$k_n^{\text{1h}} \propto (\hat{\theta}/\sigma)^{2n-2}; \qquad k_n^{\text{2h}} \propto (\hat{\theta}/\sigma)^{2n-4}. \quad \text{(A14)}$$

From this, we can draw two conclusions:
(1) Increasing the order of a cumulant by one introduces two powers of the expansion parameter. Thus, in the limit where the expansion parameter is small, we converge at a Gaussian distribution.
(2) The cumulants arising from the clustering term are larger by two powers of the expansion parameter, so that for large enough smoothing scales the two-halo term eventually takes over (despite being only a relatively small correction in the unsmoothed case).

APPENDIX B: NUMERICAL EVALUATION

For efficient computation of the one- and especially the two-point PDF, two observations can be made. First, integrals of the form

$$\int_{M,\theta} e^{i\underline{\kappa}(\theta)\lambda} \tag{B1}$$

can be transformed to

$$\int d\kappa e^{i\kappa\lambda} \int_M \pi \frac{d\underline{\theta}^2(\kappa)}{d\kappa}, \tag{B2}$$

where $\underline{\theta}(\kappa)$ is the inverse function of the convergence profile $\underline{\kappa}(\theta)$. Thus, we can use the FFT. Note that this requires that the convergence profiles are invertible (monotonic). Application of ℓ-space filters (such as the pixel window function or the Wiener filter) can occasionally lead to nonmonotonic profiles; in that case we split the integral into segments in which the convergence profiles are monotonic.

Second, in the two-point PDF it is not necessary to compute the $K_a^{(\ell)}$ from Eq. (12) explicitly, since the integral over ℓ reduces to

$$\int_\ell J_0(\ell\theta_a) J_0(\ell\theta_b) J_0(\ell\phi), \tag{B3}$$

which has the analytic form [Eq. (3) on p. 411 of Ref. [53]]

$$[4\pi^2\Delta(\theta_a,\theta_b,\phi)]^{-1}. \tag{B4}$$

Here, Δ denotes the area of the triangle with the arguments as its sides. If no triangle can be formed, the integral vanishes. The case $\phi = 0$ is relevant for this work; then the integral is a multiple of $\delta_D(\theta_a - \theta_b)$ and this propagates through in such a way that the zero-separation two-point PDF simplifies to

$$P(\kappa_a,\kappa_b;\phi=0) = P(\kappa_a)\delta_D(\kappa_a - \kappa_b), \tag{B5}$$

as it should.

APPENDIX C: VALIDATING ASSUMPTIONS OF THE HALO MODEL APPROACH

There are two distinct classes of assumptions made in this work. The first class comprises the basic underpinnings of the halo model. In comparing to two different sets of simulations, we have seen that we can describe the convergence PDF accurately well into the non-Gaussian tail, while our model has deficiencies for $\kappa \lesssim 0$ and for large smoothing scales. These problems are consistent with intuitive expectation: we know that any halo model formalism would face these problems. On the other hand, the second class of assumptions refers to certain technical choices made in the formalism that could in principle be dropped without leaving the realm of the halo model.

1. Born approximation

The first of these is the Born approximation. Our formalism crucially requires this approximation, since it relies on the additivity of the convergence signal. While we do not examine the Born approximation in any detail here, we note two reasons why we believe it to constitute only a minor correction to the PDF. 1) We performed a simple numerical test in which we placed a single halo of mass $10^{14}h^{-1}\,M_\odot$ at $z = 0.5$ and ray traced a parallel beam through its potential. The resulting deflection angles are in extremely good agreement with the Born approximation. Although this does not test for higher-order effects such as lens-lens coupling, it still constitutes a simple test demonstrating the smallness of post-Born terms. 2) Some tests have already been performed in the literature [54,55], seemingly coming to the conclusion that post-Born terms are quite negligible for galaxy lensing.

2. Triaxiality

The second technical assumption is the neglect of halo triaxiality. Our formalism could in principle be adapted quite easily to allow for triaxiality, introducing more integrations over a shape-distribution function and halo orientations. However, these additional integrations, combined with the fact that the projected convergence profiles would no longer be azimuthally symmetric, renders the computation substantially more complex. Thus, we explicitly test for the influence of triaxiality on the convergence PDF. To this end, we measure the distribution of principal axes ratios in the MassiveNuS halo catalog. This distribution is plotted in Fig. 12. We find that the relatively coarse
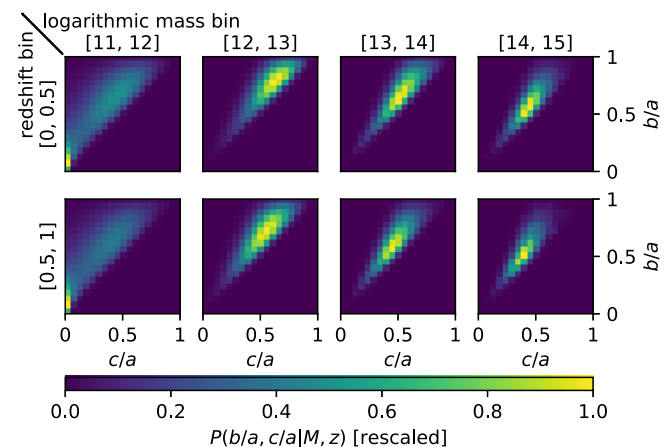

FIG. 12. Measured axes ratio distribution from the MassiveNuS halo catalog, used to construct the simplified simulations that we utilize to assess the effect of ellipsoidal vs spherical halos (cf. Fig. 13). Note that the lower triangular part of these matrices is necessarily zero by the definition of $a > b > c$. We note the peaks near $b = c = 0$ in the lowest mass bin, which are due to imprecise axes ratio measurements and spurious halo identifications close to the simulation's resolution limit.
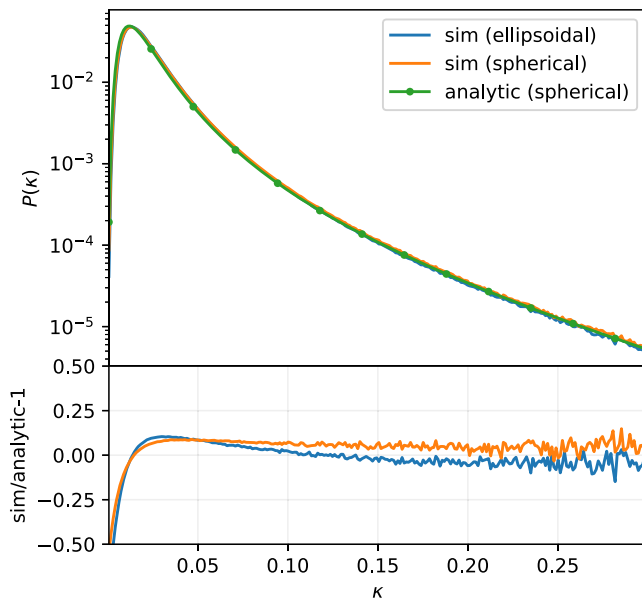
FIG. 13. Here we explore the effect of including nonspherical halos on the PDF (here source redshift $z_s = 1$). We see that the discrepancy between the complete ellipsoidal simplified simulations and the analytic result is relatively small, in particular in comparison to the discrepancy between our model and the MassiveNuS and T17 simulations. Note that in this exercise we have set a minimum mass cutoff at $\log M_{vir}/M_\odot = 12.5$ (i.e., higher than the fiducial choice of $\log M_{200m}/M_\odot = 11$). This is done purely in order to keep runtime of the simplified simulations reasonable, but since it is consistently implemented on the analytic side no issues arise from this choice.

binning in both mass and redshift is sufficient to capture the variation of the shape distribution. The extreme principal axes ratios for the lowest mass bin are likely driven by nonvirialized objects erroneously included in the halo catalog by the halo finder. However, this mass bin gives only a negligible contribution to the PDF (cf. Sec. III C).

Then, we perform simplified simulations as described in more detail in Ref. [15]. In short, these simulations randomly populate maps with convergence profiles drawn from a given distribution and measure the PDF in the end. By construction, the simplified simulations do not include the clustering effect, but this is immaterial for the purposes of the test we want to perform here. For the convergence

profiles, we proceed as follows. First, for a halo of given mass and redshift, we compute the NFW density profile. Then, we deform this spherically symmetric profile according to ratios of principal axes drawn from the distribution described before. Finally, we perform a random rotation of the resulting ellipsoid and then integrate the density along the line of sight to obtain the convergence. The result of this procedure is plotted in Fig. 13. The green line with round markers is the analytic result assuming spherically symmetric profiles, while the blue line is the described set of simplified simulations with ellipsoidal halos. As a consistency test, we also run the simplified simulations with the principal axes distribution set to a delta function at $b/a = c/a = 1$ (i.e., all halos are spherically symmetrical); the result from this test is represented by the orange line. First, we observe some discrepancies between the green and orange lines, which in principle we should expect to coincide. However, there are two reasons why perfect agreement is not reached: first, we remind the reader that our treatment of the quadratic pixels is not entirely correct; second, we expect some systematic errors in the numerical integration through the deformed NFW profiles. Thus, we consider this code test as passed. Keeping this in mind, the discrepancies between the analytic result and the simplified simulations incorporating triaxial halos are relatively small. Thus, triaxiality can certainly not account for the discrepancies we observed between our model and the MassiveNuS as well as the T17 simulations. We conclude that more reliable modeling of small scale matter clustering is of much greater importance than incorporating the small corrections from triaxiality.

3. Substructure

A final assumption is the lack of substructure in the density profiles. Considering our results concerning triaxiality, it is reasonable to assume that, given the map pixelization, the error incurred by ignoring substructure is relatively small as well. Given the dramatic effect the Wiener filter in Sec. V had in suppressing the one-point PDF's sensitivity on the concentration model, it is unlikely that parameter inference would require incorporation of substructure corrections.

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

[2] P. Schneider, L. Van Waerbeke, B. Jain, and G. Kruse, Mon. Not. R. Astron. Soc. **296**, 873 (1998).

[3] L. Van Waerbeke, T. Hamana, R. Scoccimarro, S. Colombi, and F. Bernardeau, Mon. Not. R. Astron. Soc. **322**, 918 (2001).

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

[5] M. Kilbinger and P. Schneider, Astron. Astrophys. **442,** 69 (2005).

[6] A. Cooray and W. Hu, Astrophys. J. **548,** 7 (2001).

[7] D. Huterer, M. Takada, G. Bernstein, and B. Jain, Mon. Not. R. Astron. Soc. **366,** 101 (2006).

[8] E. Semboloni, H. Hoekstra, and J. Schaye, Mon. Not. R. Astron. Soc. **434,** 148 (2013).

[9] K. Reblinsky, G. Kruse, B. Jain, and P. Schneider, Astron. Astrophys. **351,** 815 (1999).

[10] B. Jain and L. Van Waerbeke, Astrophys. J. **530,** L1 (2000).

[11] J. Liu and M. S. Madhavacheril, Phys. Rev. D **99,** 083508 (2019).

[12] J. Liu, J. C. Hill, B. D. Sherwin, A. Petri, V. Böhm, and Z. Haiman, Phys. Rev. D **94,** 103501 (2016).

[13] K. Patton, J. Blazek, K. Honscheid, E. Huff, P. Melchior, A. J. Ross, and E. Suchyta, Mon. Not. R. Astron. Soc. **472,** 439 (2017).

[14] J. C. Hill *et al.*, arXiv:1411.8004.

[15] L. Thiele, J. C. Hill, and K. M. Smith, Phys. Rev. D **99,** 103511 (2019).

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

[17] Y. Wang, D. E. Holz, and D. Munshi, Astrophys. J. **572,** L15 (2002).

[18] S. Das and J. P. Ostriker, Astrophys. J. **645,** 1 (2006).

[19] A. Stebbins, arXiv:astro-ph/9609149.

[20] J. V. Villumsen, Mon. Not. R. Astron. Soc. **281,** 369 (1996).

[21] B. Jain and U. Seljak, Astrophys. J. **484,** 560 (1997).

[22] N. Kaiser, Astrophys. J. **498,** 26 (1998).

[23] L. van Waerbeke, F. Bernardeau, and Y. Mellier, Astron. Astrophys. **342,** 15 (1999).

[24] C. Uhlemann, C. Pichon, S. Codis, B. L'Huillier, J. Kim, F. Bernardeau, C. Park, and S. Prunet, Mon. Not. R. Astron. Soc. **477,** 2772 (2018).

[25] A. Barthelemy, S. Codis, C. Uhlemann, F. Bernardeau, and R. Gavazzi, Mon. Not. R. Astron. Soc. **492,** 3420 (2020).

[26] K. Kainulainen and V. Marra, Phys. Rev. D **80,** 123020 (2009).

[27] K. Kainulainen and V. Marra, Phys. Rev. D **84,** 063004 (2011).

[28] K. Kainulainen and V. Marra, Phys. Rev. D **83,** 023009 (2011).

[29] J. Alsing, B. Wandelt, and S. Feeney, Mon. Not. R. Astron. Soc. **477,** 2874 (2018).

[30] J. F. Navarro, C. S. Frenk, and S. D. M. White, Astrophys. J. **490,** 493 (1997).

[31] E. Komatsu and U. Seljak, Mon. Not. R. Astron. Soc. **336,** 1256 (2002).

[32] N. Battaglia, J. R. Bond, C. Pfrommer, and J. L. Sievers, Astrophys. J. **758,** 75 (2012).

[33] J. C. Hill and E. Pajer, Phys. Rev. D **88,** 063526 (2013).

[34] R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce, P. A. Thomas, G. Efstathiou, and H. M. P. Couchman, Mon. Not. R. Astron. Soc. **341,** 1311 (2003).

[35] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya, and M. Oguri, Astrophys. J. **761,** 152 (2012).

[36] J. L. Tinker, B. E. Robertson, A. V. Kravtsov, A. Klypin, M. S. Warren, G. Yepes, and S. Gottlöber, Astrophys. J. **724,** 878 (2010).

[37] A. R. Duffy, J. Schaye, S. T. Kay, and C. D. Vecchia, Mon. Not. R. Astron. Soc. Lett. **390,** L64 (2008).

[38] B. Diemer, Astrophys. J. Suppl. Ser. **239,** 35 (2018).

[39] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538,** 473 (2000).

[40] J. Lesgourgues, arXiv:1104.2932.

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

[42] G. L. Bryan and M. L. Norman, Astrophys. J. **495,** 80 (1998).

[43] J. Liu, S. Bird, J. M. Zorrilla Matilla, J. C. Hill, Z. Haiman, M. S. Madhavacheril, A. Petri, and D. N. Spergel, J. Cosmol. Astropart. Phys. 03 (2018) 049.

[44] R. Takahashi, T. Hamana, M. Shirasaki, T. Namikawa, T. Nishimichi, K. Osato, and K. Shiroyama, Astrophys. J. **850,** 24 (2017).

[45] A. Cooray, W. Hu, and J. Miralda-Escudé, Astrophys. J. **535,** L9 (2000).

[46] J. M. Zorrilla Matilla, S. Waterval, and Z. Haiman, Astron. J. **159,** 284 (2020).

[47] M. Joyce, L. Garrison, and D. Eisenstein, arXiv:2004.07256 [Mon. Not. R. Astron. Soc. (to be published)].

[48] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, Astrophys. J. **622,** 759 (2005).

[49] Planck Collaboration, Astron. Astrophys. **641,** A6 (2020).

[50] J. Alsing, T. Charnock, S. Feeney, and B. Wand elt, Mon. Not. R. Astron. Soc. **488,** 4440 (2019).

[51] A. Y. Chen and N. Afshordi, Phys. Rev. D **101,** 103522 (2020).

[52] O. H. E. Philcox, D. N. Spergel, and F. Villaescusa-Navarro, Phys. Rev. D **101,** 123520 (2020).

[53] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Cambridge University Press, Cambridge, England, 1922).

[54] A. Petri, Z. Haiman, and M. May, Phys. Rev. D **95,** 123503 (2017).

[55] A. Barthelemy, S. Codis, and F. Bernardeau, Mon. Not. R. Astron. Soc. **494,** 3368 (2020).