

# Numerical complexity of the joint nulled weak-lensing probability distribution function

Alexandre Barthelemy<sup>\*</sup>

*CNRS & Sorbonne Université, UMR 7095, Institut d'Astrophysique de Paris,  
75014 Paris, France*

Francis Bernardeau

*Institut de Physique Théorique, Université Paris-Saclay,  
CEA, CNRS, UMR 3681, 91191 Gif-sur-Yvette, France*

Sandrine Codis

*AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot,  
Sorbonne Paris Cité, 91191 Gif-sur-Yvette, France*

Cora Uhlemann

*School of Mathematics, Statistics and Physics, Newcastle University,  
Herschel Building, NE1 7RU Newcastle-upon-Tyne, United Kingdom*



(Received 22 June 2021; accepted 15 February 2022; published 28 February 2022)

In the context of tomographic cosmic shear surveys, there exists a nulling transformation of weak lensing observations (also called BNT transform) that allows us to simplify the correlation structure of tomographic cosmic shear observations, as well as to build observables that depend only on a localized range of redshifts and thus independent from the low-redshift/small-scale modes. This procedure renders possible accurate, and from-first-principles, predictions of the convergence and aperture mass one-point distributions (PDF). We here explore other consequences of this transformation on the (reduced) numerical complexity of the estimation of the joint PDF between nulled bins and demonstrate how to use these results to make theoretical predictions.

DOI: [10.1103/PhysRevD.105.043537](https://doi.org/10.1103/PhysRevD.105.043537)

## I. INTRODUCTION

The statistics of weak-lensing fields provide us with a powerful tool for precision cosmology (see, e.g., a review in [1]) and motivated the build up of new generation large galaxy surveys such as the Legacy Survey of Space and Time (LSST) [2] or Euclid [3] which will collect data of unprecedented quality in the coming years.

Among all the studied statistics of those fields, and especially among probes of their non-Gaussian features, the probability distribution function (PDF) has received some attention. Although some theoretical efforts were engaged long before, some formal reasons can for example be found in [4] who demonstrated that the weak-lensing convergence PDF provides information complementary to the cosmic shear two-point correlation. Note that this complementarity between two-point and non-Gaussian statistics is even more relevant in the presence of shot and shape noise. Still making use of a Fisher analysis but this time using a “from-first-principles” theoretical model of the convergence PDF,

[5] also found that it yields tighter constraints for the equation of state of dark energy, the amplitude of fluctuations, the total matter fraction and the sum of neutrino masses, especially when performing a multiscale analysis and in addition to the two-point correlation function.

As for the theoretical modeling, although there exists numerous works based on numerical simulations or more recently the halo model as in [6], we will focus here on the works whose approach could be qualified as from-first-principles.

Already in [7] and their following papers, some hierarchical models for the statistics of the matter field were used to construct the PDF of the aperture mass, and these works were later reinterpreted in terms of a large deviation principle and opening the way to extensions to more realistic observables in [8]. Note that the general formalism involving large deviation theory was before [8] used in the context of the clustering of the matter field itself by the same team, see for example [9] or [10] for an introduction. Roughly at the same time as [8], the works of [7] also inspired [11] for their modeling of the PDF of tangential shear profiles. The most recent works on the PDF of the

<sup>\*</sup>alexandre.barthelemy@iap.fr

convergence and the aperture mass within this large deviation framework are [12] and [13] which take into account the full geometry of the light cone and probe the validity regime of the formalism with the help of numerical simulations. These papers also make use of a nulling strategy initially developed in [14] and not at all specific to the PDF since directly implemented at the level of the lensing maps.

This nulling strategy, sometimes referred to as the BNT transform, is built on the fact that the lensing window function, that traces the matter density field along the line of sight and gives rise to the observed lensing quantity such as the convergence and the shear within the Born approximation, depends on the cosmological background evolution and in particular, does not depend on the dynamical properties of the density fluctuations. Given this fact and in the context of a tomographic analysis where the lensing fields are observed for sources over a wide range of redshifts, [14] defines a linear transformation—which thus only depends on the cosmological parameters through two moments only of the lens distance distribution—that can be applied to the observed maps. It then gives rise to a new set of maps which are effectively the result of a new set of lensing window functions. These *nullled* window functions now vanish over all redshifts comprised between the closest source and the observer, effectively *nulling* the effect of nearby structures. This transformation has multiple consequences: it allows us to separate linear or quasilinear scales from very nonlinear scales thus making theoretical modeling easier to control; it also makes nullled lensing maps partially uncorrelated, a property we will use to our advantage here. These features are illustrated in Fig. 1 and further discussed hereafter.

Until now, the nulling strategy has been advocated in the context of multiple nullled maps for the 2-point correlators,

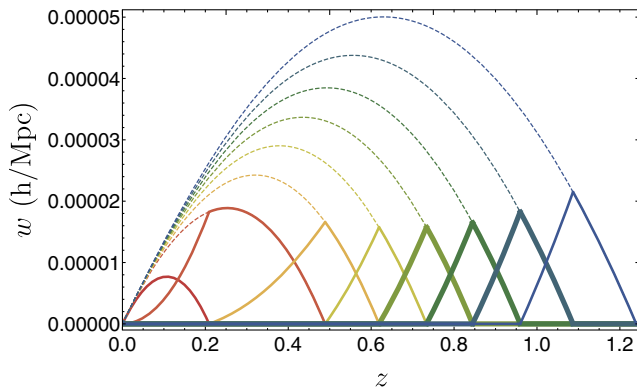


FIG. 1. The radial selection functions, prior to nulling (dashed lines) and after nulling (thick lines). The source redshifts are taken as the mean redshift of the first 8 equally populated redshift bins of the Euclid experiment. The three thicker nullled selection functions are the ones we use for the practical implementation in Sec. V.

angular power spectra or correlation functions, in the context of galaxy and CMB lensing but also line-intensity mapping lensing [15–17], but not for more complex observables and in particular not for the joint density PDFs. The goal of this paper is thus now to explore some consequences of the correlation structure of the nullled convergence and aperture mass fields for different source redshifts along the line of sight on the computation of their joint PDF. Note that this makes the derivations we present much more general than this cosmological context since it presents consequences on the computation of the joint PDF of any series of random variables with “blocks” of nonzero intercorrelations, i.e., variables that are chain-correlated 2 by 2 or 3 by 3. Section II briefly recalls the definition of the convergence and the aperture mass, Secs. III and IV derive the main results of this paper on the numerical complexity of the joint cumulant generating function (CGF) and PDF, Sec. V illustrates how one could use these derivations to compute the joint PDF of the nullled convergence field for different sources treating the underlying matter field as in [12] (large deviation theory), and finally Sec. VI concludes.

## II. CONVERGENCE AND APERTURE MASS

The convergence  $\kappa$  is traditionally interpreted as a line-of-sight projection of the matter density contrast distribution between the observer and the source. Neglecting lens-lens couplings and within the Born approximation it is written as [18]

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_s} d\chi w(\chi, \chi_s) \delta(\chi, \mathcal{D}\boldsymbol{\theta}), \quad (1)$$

where  $\chi$  is the comoving radial distance— $\chi_s$  the radial distance of the source—that depends on the cosmological model, and  $\mathcal{D}$  is the comoving angular distance. The lensing kernel  $w$  is expressed as

$$w(\chi, \chi_s) = \frac{3\Omega_m H_0^2}{2c^2} \frac{\mathcal{D}(\chi)\mathcal{D}(\chi_s - \chi)}{\mathcal{D}(\chi_s)} (1 + z(\chi)). \quad (2)$$

The aperture mass  $M_{\text{ap}}$  is defined as a geometrical average of the local convergence within a window of vanishing average

$$M_{\text{ap}}(\boldsymbol{\theta}) = \int d^2\boldsymbol{\theta}' U_{\boldsymbol{\theta}}(\boldsymbol{\theta}') \kappa(\boldsymbol{\theta}' - \boldsymbol{\theta}) \quad (3)$$

with the compensated filter obeying

$$\int d^2\boldsymbol{\theta}' U_{\boldsymbol{\theta}}(\boldsymbol{\theta}') = 0. \quad (4)$$

The aperture mass can moreover be interestingly expressed as a function of the tangential component  $\gamma_t$  of the shear [19,20], thus rendering the aperture mass a direct

observable up to a reduced shear correction that can also be accounted for as discussed in [13].

Note that the calculation we present in this paper and especially in Sec. IV, our main result, are very general and in particular their application in the cosmological context does not depend on the choice of a specific filtering scheme for the convergence field. For practical implementation purposes and graphical representation, we will however adopt a simple prescription where the convergence  $\kappa_{<\theta}$  is filtered within a top-hat window of angular radius  $\theta$  and the associated aperture mass is given by

$$M_{\text{ap}}(\boldsymbol{\theta}) = \kappa_{<\theta_2}(\boldsymbol{\theta}) - \kappa_{<\theta_1}(\boldsymbol{\theta}). \quad (5)$$

In the following we will drop both the mention of  $\boldsymbol{\theta}$ , thanks to isotropy, and  $<\theta$  to simplify the expressions.

### III. THE CUMULANT GENERATING FUNCTIONS AND THE PDFS

In order to predict the joint statistics in multiple redshift bins, the starting point is the expression of the cumulant generating functions of projected densities which will, in our case, correspond to the convergence or the aperture mass. Following [12] this cumulant generating function can be shown to take the form,

$$\phi(\{\lambda_i\}) = \int d\mathbf{r} \phi_{\text{cyl/slope}} \left( \sum_i \lambda_i w_i(r); r\theta_0; z(r) \right) \quad (6)$$

where  $\theta_0$  is the filtering angular size,  $w_i(r)$  is the line-of-sight profile which leads to the observable  $\kappa_i$  map,  $\phi_{\text{cyl}}(\lambda; R; z)$  is the cumulant generating function of the density in a cylinder of transverse size  $R$ , at redshift  $z$ , written as a function of  $\lambda$ , and  $\phi_{\text{slope}}(\lambda; R; z)$  is the cumulant generating function of the difference between the densities in two concentric cylinders of transverse sizes  $R \equiv R_1$  and  $R_2 > R_1$  at redshift  $z$ . The subscripts cyl and slope respectively lead to the convergence and the aperture mass and we write  $\phi_{\text{cyl/slope}}$  when either or the other can be used. For simplicity, we will present the formalism for the convergence  $\kappa$  but the same equations hold for the aperture mass field albeit replacing cyl by slope.

The joint probability density function (PDF) of the maps  $\kappa_i$  corresponding to redshift bins indexed by  $i$  can then be constructed through an inverse Laplace transformation of the cumulant generating function,

$$\mathcal{P}(\{\kappa_i\}) = \int \prod_{i=1}^{n_t} \frac{d\lambda_i}{2\pi i} \exp \left[ - \sum_{i=1}^{n_t} \lambda_i \kappa_i + \phi(\{\lambda_i\}) \right] \quad (7)$$

where  $n_t$  is the total number of source planes and where the integrals run along the imaginary axis. The properties of such a transformation have been described in detail in many papers (see for instance [21] and references therein).

We simply note here that marginalizing over one variable  $\kappa_p$  is obtained by setting the corresponding variable  $\lambda_p$  to 0,

$$\begin{aligned} & \int d\kappa_p \mathcal{P}(\{\kappa_i\}) \\ &= \int \frac{d\lambda_1}{2\pi i} \cdots \frac{d\lambda_{p-1}}{2\pi i} \frac{d\lambda_{p+1}}{2\pi i} \cdots \frac{d\lambda_{n_t}}{2\pi i} \\ & \times \exp \left[ - \sum_{i=1}^{p-1} \lambda_i \kappa_i - \sum_{i=p+1}^{n_t} \lambda_i \kappa_i + \phi(\{\lambda_i\}|_{\lambda_p=0}) \right]. \quad (8) \end{aligned}$$

This property will be very useful in the following.

For simplification we have assumed until now that all sources are located on discrete source planes which makes the selection functions  $w_i(r)$  specific functions of the radial distance, itself only dependent on the background geometry when working within the Born approximation. As was previously shown in [14], it is then possible to define a lower triangular matrix  $p_{ij}$  [22] that depends solely on the distance-redshift dependence that linearly transforms the selection functions  $w_i(r)$  into

$$\hat{w}_i(r) = \sum_j p_{ij} w_j(r) \quad (9)$$

in such a way that the functions  $\hat{w}_i(r)$  vanish over all redshifts comprised between the observer and the closest sources. We illustrate this transformation in Fig. 1 where for a set of  $\{\kappa_i\}$  and in the case of discrete sources the transformation is possible for all planes except the first two. The *nulled* selection functions are displayed with solid lines whereas the dashed lines of the same colors represent the corresponding regular selection functions.

Let us now reconsider Eq. (6) for the variable  $\hat{\kappa}_i(r)$ ,

$$\hat{\kappa}_i(r) = \sum_j p_{ij} \kappa_j, \quad (10)$$

which only changes  $w_i(r)$  into  $\hat{w}_i(r)$  in expression (6). The main integral that appears in this expression can now be split in  $n_t$  different parts—in between the locations of the discrete source planes  $\{r_i\}$ . This nulled generating function then reads,

$$\hat{\phi}(\{\lambda_i\}) = \sum_{n=1}^{n_t} \int_{r_{n-1}}^{r_n} dr \phi_{\text{cyl/slope}} \left( \sum_{i=n}^{n+1} \lambda_i \hat{w}_i(r); r\theta_0 \right) \quad (11)$$

where  $\lambda_{n_t+1} = 0$  and it is now clear that for each part, only two distinct subsequent  $\lambda$  variables of each nulled bin appear. In other words the cumulant generating function takes the form

$$\hat{\phi}(\{\lambda_i\}) = \sum_{i=1}^{n_t-1} \tilde{\phi}_i(\lambda_i, \lambda_{i+1}) + \tilde{\phi}_{n_t}(\lambda_{n_t}). \quad (12)$$

Note that the functions  $\tilde{\phi}_i$  are not cumulant generating functions on their own but are closely related to them. By identification and using the fact that  $\hat{\phi}(\lambda_i) = \hat{\phi}(\{\lambda_j = \delta_{ij}\lambda_i\})$ , it can then be shown that we have the following structure,

$$\begin{aligned} \hat{\phi}(\{\lambda_i\}) &= \hat{\phi}(\lambda_1, \lambda_2) - \hat{\phi}(\lambda_2) + \hat{\phi}(\lambda_2, \lambda_3) \\ &\quad - \cdots - \hat{\phi}(\lambda_{n_i-1}) + \hat{\phi}(\lambda_{n_i-1}, \lambda_{n_i}), \end{aligned} \quad (13)$$

where  $\hat{\phi}$  are now all genuine cumulant generating functions of the corresponding variables.

It thus remains that, with this choice of nulled variables, the joint cumulant generating function has a specific functional form: it is composed of a sum of functions that depend on two variables only, which implies that the full joint PDF can be obtained from bivariate CGFs (and hence PDFs) of neighboring bins. We will thus now explore the consequences of such a feature on the computations and properties of the joint PDF.

One can first infer a number of general properties: it is clear that the two sets of variables  $\{\hat{\kappa}_i\}_{i=1,\dots,p-1}$  and  $\{\hat{\kappa}_i\}_{i=p+1,\dots,n_i}$  are correlated only through the variable  $\hat{\kappa}_p$ . In other words, once one marginalises over  $\hat{\kappa}_p$ , the two sets of variables are independent as there are no common structures that would contribute to both sets.

At the level of the cumulant generating function, one can notice that we have

$$\hat{\phi}(\{\lambda_i\}|_{\lambda_p=0}) = \hat{\phi}(\{\lambda_i\}_{i=1,\dots,p-1}) + \hat{\phi}(\{\lambda_i\}_{i=p+1,\dots,n_i}) \quad (14)$$

so that the cumulant generating function is split in two separate functions.

The application of Eq. (8) then readily shows that we have,

$$\begin{aligned} \int d\hat{\kappa}_p \mathcal{P}(\{\hat{\kappa}_i\}_{i=1,\dots,n_i}) \\ = \mathcal{P}(\{\hat{\kappa}_i\}_{i=1,\dots,p-1})\mathcal{P}(\{\hat{\kappa}_i\}_{i=p+1,\dots,n_i}). \end{aligned} \quad (15)$$

This independence property for physically separated bins is specific to the nulled variables.

Finally, realistic tomographic cases are not made of a collection of discrete source planes but instead source galaxies are split in redshift bins of finite width. The resulting nulling selection functions are then slightly more complex as illustrated in Fig. 2. One can indeed see here that the nulled profiles overlap with the first and second neighbors—and not only with the first. As a result, the functional form of the cumulant generating function that we obtain is

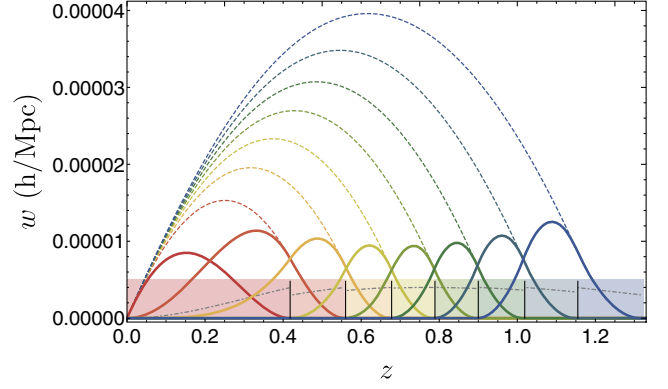


FIG. 2. The realistic radial selection functions of the first eight equally populated redshift bins (shown in colored rectangles) of the Euclid experiment, prior to nulling (dashed lines) and after nulling (thick lines). The dot-dashed gray lines inside the colored rectangles show the shape of the normalized distribution of sources inside each bin.

$$\begin{aligned} \hat{\phi}(\{\lambda_i\}) &= \tilde{\phi}(\lambda_1, \lambda_2, \lambda_3) + \cdots + \tilde{\phi}(\lambda_{n_i-2}, \lambda_{n_i-1}, \lambda_{n_i}) \\ &\quad + \tilde{\phi}(\lambda_{n_i-1}, \lambda_{n_i}) + \tilde{\phi}(\lambda_{n_i}). \end{aligned} \quad (16)$$

In this case, a property similar to Eq. (15) holds except that an integration over two consecutive variables is required.

### A. Impact of shape noise and intrinsic alignment

The inclusion of a shape noise—which comes from the variance of the intrinsic ellipticities of galaxies—to the joint nulled CGF would lead to the addition of a term  $\varphi_{\text{noise}}(\{\lambda_i\})$  (such that  $\hat{\phi}_{\text{tot}} = \hat{\phi} + \varphi$ ) that would follow the same functional form (16) whether realistic source bins or discrete source planes are considered, with

$$\varphi_{\text{noise}}(\{\lambda_i\}) = \sum_{jii'} p_{ij} p_{i'j} \frac{\sigma_S^2(j)}{2} \lambda_i \lambda_{i'} \quad (17)$$

where  $\sigma_S^2(j)$  is the amplitude of the shape noise in the original source bin  $j$  (and with respect to the shape noise, bins are all independent) assuming it induces a Gaussian noise. This amplitude for the Euclid experiment can for example be found in [23]. The only nonvanishing terms in (17) are those for which there is a  $j$  for which  $j-2 \leq i \leq j$  and  $j-2 \leq i' \leq j$  that implies that  $|i-i'| \leq 2$  thus reproducing the form (16).

Intrinsic alignments (IA)—originating from the correlation between the orientation of galaxies and their environment—represent a contamination for cosmic shear measurements [24]. When the observed shape of a galaxy is decomposed into an intrinsic term and a contribution coming from weak lensing along the line-of-sight, two sorts of correlators arise, the autocorrelation spectra of intrinsic shapes (II in the case of the two-point spectrum) and the cross-correlation spectra between foreground intrinsic shapes and

background shear (GI still in the case of the two-point spectrum). There cannot be any contribution coming from foreground shear and background intrinsic shapes, or more generally, the shear-induced ellipticity of a galaxy can only be correlated to objects located within its radial selection function, as those correlations occur on smaller scales (typically below a hundred megaparsecs, therefore not between distant redshift bins). Thus in the case of nulling and similarly to shape noise, intrinsic alignments would also induce correlations that span over three nulled bins only, although with *a priori* a more complex dependence with  $\lambda$  since intrinsic alignments may not induce Gaussian noise. Still for the Euclid experiment, a modeling of the impact of intrinsic alignments on the lensing power spectrum and bispectra can be found in [23], and used to obtain the 2nd and 3rd cumulants in the equivalent of Eq. (17) for the inclusion of galaxies alignments. However, there is no guarantee that such linear or quasilinear models are valid in the regime we are probing here and more works is therefore necessary to model this effect properly, including small-scale baryonic effects [25].

This thus implies that the correlation structure we presented for nulled variables is still valid when such effects are taken into account. Note however that corrective terms coming from reduced shear, magnification bias or postborn corrections formally break this structure, although weakly, as they change the functional form of Eq. (1) and introduce couplings between lenses which are not considered independent anymore.

Finally note that one obvious but strong limitation of this procedure is the precision to which the redshift of sources inside tomographic bins can be determined (photometric redshift errors). Indeed, in the worst case where the precision of, for example and in the context of a large area survey, photometric redshifts is not very good with a significant number of catastrophic events, the “true” redshift distribution of galaxies inside each redshift bin can become strongly overlapping as is the case for example for the KIDS survey [see Fig. 2 of [26]]. As a result, the nulling procedure cannot be performed perfectly. We expect that for future surveys like Euclid, photometric redshift errors can be reduced as suggested by the treatment of those errors in Euclid collaboration prelaunch papers [see for example Eq. (46) of [23]]. In any case, the quantification of the impact of photometric redshift errors on the nulling procedure is something which can be done rather straightforwardly in the formalism that we present—computing BNT weights with a given galaxy distribution while having lensing kernel computed from a different one accounting for uncertainties and checking the residual correlations and their impact on say, cosmological parameters estimation. Note also that, as a last-resort backup plan, we could also divide-and-conquer the problem by modeling the photo- $z$  error on the univariate/bivariate/trivariate PDFs and capturing the additional correlations caused by redshift errors in a

Gaussian-style copula in the same fashion as what we do in Appendix B. Finally, another option could also come from the reliance on some numerical lensing simulation as in [27] to generate systematics-infused control samples and use them to model the impact of photometric redshift and shear calibration uncertainty. This is however an important work on its own and is left for the future as beyond the scope of the results we present here.

The quantification of all of these effects on the nulling strategy will be performed in future works in preparation.

#### IV. NUMERICAL COMPLEXITY FOR PDF COMPUTATIONS

Let us now explore the resulting numerical complexity of the computation of the joint PDF of nulled maps,  $\hat{\kappa}_i$ . Note that this is not just a numerical trick to reduce the computing time, this is the simpler form the joint PDF can take using the correlation structures that we have. In particular, Eq. (7) for nulled bins can as a result be written as some function of joint PDFs between two or three bins but not in a simpler form than the one we present here, see for example Appendix A.

Let us start our analysis in the case of discrete source planes, that is taking advantage of the functional form (12) for the cumulant generating function and thus assuming that this generating function can be written as a sum of functions of two consecutive variables.

If we assume that the number of operations for a given integral over  $\lambda_i$  is  $N$ , then, and in the absence of factorization properties, the *a priori* complexity for the computation of the joint  $\kappa$ -PDF, as given by Eq. (7), is  $N^n$ . The purpose of this section is to show how this complexity, and hence the computational time, can be reduced when changing variables from  $\kappa$  to  $\hat{\kappa}$ . To reach this goal, let us define the function  $\zeta_2(n_i)$ , supposedly an integer, that gives an estimate of the number of operations to be done as  $N^{\zeta_2(n_i)}$ . Similarly, we also define a function  $\zeta_3(n_i)$  for the functional form (16) of the cumulant generating function corresponding not to successive localized source planes but wide redshift bins.

Let us start with a simple case. For a joint analysis of three source planes, the cumulant generating function takes the form,

$$\hat{\phi}(\lambda_1, \lambda_2, \lambda_3) = \hat{\phi}(\lambda_1, \lambda_2) - \hat{\phi}(\lambda_2) + \hat{\phi}(\lambda_2, \lambda_3) \quad (18)$$

so that we can write,

$$\mathcal{P}(\hat{\kappa}_1, \hat{\kappa}_2, \hat{\kappa}_3) = \int \frac{d\lambda_2}{2\pi i} e^{-\lambda_2 \hat{\kappa}_2 + \hat{\phi}_{\hat{\kappa}_1}(\lambda_2) - \hat{\phi}(\lambda_2) + \hat{\phi}_{\hat{\kappa}_3}(\lambda_2)} \quad (19)$$

where

$$e^{\hat{\phi}_{\hat{\kappa}_i}(\lambda_j)} = \int \frac{d\lambda_i}{2\pi i} \exp(-\lambda_i \hat{\kappa}_i + \hat{\phi}(\lambda_i, \lambda_j)). \quad (20)$$

The computation of each function  $e^{\hat{\phi}_{\kappa_i}(\lambda_j)}$  simply scales like  $N^1$ . This thus implies that the estimated number of operations for the computation of expression (19) is  $N \times (N + N)$  which scales like  $N^2$  so that  $\zeta_2(3) = 2$ .

Let us now explore the general case and define the quantity

$$e^{\hat{\phi}_{\kappa_p}(\lambda_{p-1}, \lambda_{p+1})} = \int \frac{d\lambda_p}{2\pi i} \exp[-\lambda_p \hat{\kappa}_p + \hat{\phi}(\lambda_{p-1}, \lambda_p) - \hat{\phi}(\lambda_p) + \hat{\phi}(\lambda_p, \lambda_{p+1})]. \quad (21)$$

The number of operations for the computation of this function is  $N$ . The idea is now, as exemplified in the simple 3-bin case, to decimate the variables, 1 out of 2, as illustrated on Fig. 3.

Let us first assume that the number of planes,  $n_t$ , is an odd number,  $n_t = 2m + 1$ . From the structure of the cumulant generating function, and the previous definition, follows one of the main results of this paper

$$\mathcal{P}(\{\hat{\kappa}_i\}) = \int \prod_{i=1}^m \frac{d\lambda_{2i}}{2\pi i} e^{-\sum_{i=1}^m \lambda_{2i} \hat{\kappa}_{2i}} \times \exp \left[ \sum_{i=0}^m \hat{\phi}_{\kappa_{2i+1}}(\lambda_{2i}, \lambda_{2(i+1)}) - \hat{\phi}(\lambda_{2i}) \right] \quad (22)$$

with  $\lambda_0 = \lambda_{n_t+1} = 0$ . The number of operations required to integrate the odd variables is of the order of  $(m + 1)N$ . What is then remarkable is that the structure of the resulting  $m$ -inverse Laplace transform is the same as before: the effective cumulant generating function is a sum of function of two adjacent variables only. As a result the expected number of operations to be performed is  $(m + 1)N \times N^{\zeta_2(m)}$ . To a logarithmic correction we will ignore in the following we thus have  $\zeta_2(2m + 1) = 1 + \zeta_2(m)$ . Let us finish this evaluation by noting that there is no gain in reduced numerical complexity going from  $2m + 1$  to  $2m$  source planes. We thus have  $\zeta_2(2m) = \zeta_2(2m + 1) = 1 + \zeta_2(m)$ .

Applying this simple rule to the first few integers, we find,

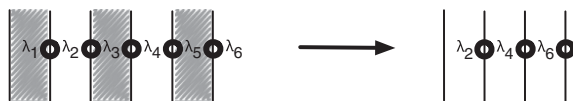


FIG. 3. Schematic representation of the decimation procedure for the case of discrete source planes. The generating function is a sum of functions of adjacent variables. Each of these terms is depicted with a dark circle. The integrations over 1 variable out of 2, (grey areas) can then be made independently. The resulting structure is again a sum of functions of adjacent variables and one can thus apply this decimation strategy recursively.

$$\begin{aligned} \zeta_2(1) &= 1 \\ \zeta_2(2) &= \zeta_2(3) = 2 \\ \zeta_2(4) &= \dots = \zeta_2(7) = 3 \\ \zeta_2(8) &= \dots = \zeta_2(15) = 4. \end{aligned} \quad (23)$$

Finally note again that the reason we need to go through this somewhat complicated numerical integration scheme comes from the fact that there is no simpler form for the joint  $\hat{\kappa}$ -PDF in terms of bivariate PDFs of neighboring bins as is the case for the CGFs. To explicit this, we show in Appendix A how the PDF of 3 neighboring bins is expressed explicitly as a function of bivariate PDFs.

We now move to the case of Eq. (16). We can use the same trick and decimate the variables. We however identified here two possible strategies which are depicted in Fig. 4.

The first one is to decimate the variables as before: 1 variable out of 2. The difference here is that this first set of variables are not uncorrelated, they are actually correlated 2 by 2. To be more precise let us again assume that  $n_t = 2m + 1$  and the complexity of this operation is thus  $N^{\zeta_2(m+1)}$  defined in the previous section. We are then left with  $m$  variables that are open again correlated 2 by 2 and therefore have

$$\zeta_3(2m + 1) = \zeta_2(m + 1) + \zeta_2(m) \quad (24)$$

and similarly

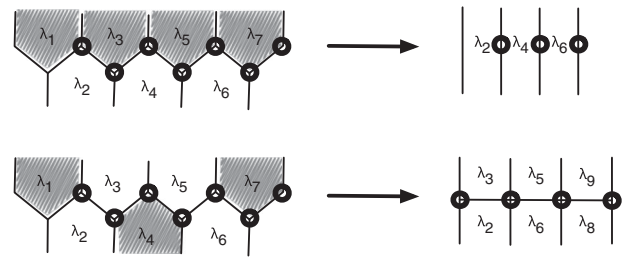


FIG. 4. Schematic representation of the decimation procedures for the case of extended source distributions. The generating function is a sum of functions of three adjacent variables. Each of these terms is depicted with a dark circle. In the top panel the proposed scheme is to first consider the integration of 1 variable out of 2 (shaded odd variables). These are correlated 2 by 2 and we thus apply the first step depicted in Fig. 3. We then consider the even variables which are also correlated 2 by 2. This allows to once again perform the first step of Fig. 3 and we then go back to the odd variables and apply this scheme recursively. In the bottom panel the procedure is to integrate 1 variable out of 3 where each of those integrations can be performed independently. This leads to an effective cumulant generating function for which variables are correlated via adjacent pairs which we again know how to treat.

TABLE I. Numerical complexity  $N^{\zeta(n_t)}$  for the calculation of the joint  $\kappa$ -PDF in case of discrete source planes,  $\zeta_2(n_t)$ , and extended source planes,  $\zeta_3(n_t)$  and  $\zeta'_3(n_t)$ , as a function of the number of source planes. The latter two results correspond to two different strategies as described in the main text.

$n_t$	$\zeta_2(n_t)$	$\zeta_3(n_t)$	$\zeta'_3(n_t)$
1	1	1	1
2	2	2	2
3	2	3	3
4	3	4	3
5	3	4	5
6	3	4	5
7	3	5	5
8	4	6	5
9	4	6	5
10	4	6	5
11	4	6	7
12	4	6	7
13	4	6	7
14	4	6	7
15	4	7	7

$$\zeta_3(2m) = 2\zeta_2(m). \quad (25)$$

The second strategy (the resulting complexity being denoted with a new function  $\zeta'_3$ ) is to decimate 1 variable out of 3, so that the first set of variables are uncorrelated. For this operation, the number of operation scales like  $N$ . To be more precise let us assume  $n_t = 3m + 1$ . After decimation we are left with  $2m$  variables. These variables then follow a specific structure. They form pairs that are correlated to their nearest neighbors, which is a situation we know how to handle. Each pair introduces a  $N^2$  factor so that we eventually have,

$$\zeta'_3(3m + 1) = 1 + 2\zeta_2(m). \quad (26)$$

With such a strategy we have  $\zeta'_3(3m - 1) = \zeta'_3(3m) = \zeta_3(3m + 1)$  (except for  $\zeta'_3(1)$  and  $\zeta'_3(2)$ ).

It turns out that these two strategies have competing performances, as illustrated in Table I.

## V. PRACTICAL IMPLEMENTATION

We now move to the actual implementation of the decimation strategy presented in the last section. We will thus start with the computation of  $\phi_{\text{cyl}}$  and  $\phi_{\text{slope}}$  and their analytical continuation in the complex plane. The computation of  $\hat{\phi}(\{\lambda_i\})$  is then performed on a fixed grid of imaginary values  $\{\lambda_i\}$  and we will finally illustrate the procedure described in the previous section with several plots.

### A. Computation of $\hat{\phi}(\{\lambda_i\})$

The computation of this quantity have already been described in several papers, most recently in [12] for the convergence and [13] for the aperture mass using a from-first-principles formalism inspired by Large deviation theory, though some equations were already known but not interpreted in this framework as in [7].

In this context, the joint cumulant generating function of the density filtered in concentric long cylinders of transverse sizes  $\{R_i\}$  at redshift  $z$  is given by

$$\phi_{\{\delta_i\}}(\{\lambda_i\}) = \sum_i \lambda_i \delta_i - \Psi_{\{\delta_i\}}(\{\delta_i\}), \quad (27)$$

where  $\{\delta_i\}$  are functions of  $\{\lambda_i\}$  through the stationary conditions

$$\lambda_k = \frac{\partial \Psi_{\{\delta_i\}}(\{\delta_i\})}{\partial \delta_k}, \quad \forall k. \quad (28)$$

The rate function  $\Psi_{\{\delta_i\}}(\{\delta_i\})$  is given by

$$\Psi_{\{\delta_i\}}(\{\delta_i\}) = \frac{1}{2} \sum_{k,j} \Xi_{kj} \mathcal{C}^{-1}(\delta_k) \mathcal{C}^{-1}(\delta_j), \quad (29)$$

where  $\Xi_{kj}$  is the inverse of the nonlinear covariance matrix between cylinders of transverse radii  $\{R_i(1 + \delta_i)^{1/2}\}$ , and the inverse of the cylindrical collapse dynamics  $\mathcal{C}$  can be approximately written as

$$\mathcal{C}^{-1}(\delta_k) = \nu - \nu(1 + \delta_k)^{-1/\nu}, \quad \nu = 1.4. \quad (30)$$

Then, from Eq. (27) and the generic properties of cumulant generating functions, one can finally define

$$\phi_{\text{cyl}}(\lambda) = \phi_{\delta}(\lambda) \quad \text{and} \quad (31)$$

$$\phi_{\text{slope}}(\lambda) = \phi_{\delta_1, \delta_2}(-\lambda, \lambda). \quad (32)$$

The previous set of equations thus allows one to numerically compute  $\phi_{\text{cyl}}(\lambda)$  and  $\phi_{\text{slope}}(\lambda)$  for any real value of  $\lambda$ . Unfortunately, since the computation of the PDF of projected densities requires an integration in the complex plane, we need a prescription for the analytic continuation of  $\phi_{\text{cyl/slope}}$ . As previously explained in [7] and [13], a good strategy to keep the analytical properties of the construction (27) consists in fitting an effective  $\mathcal{C}$  function rewriting Eq. (27) for  $\phi_{\text{cyl/slope}}$  as

$$\phi_{\text{cyl/slope}}(\lambda) = \lambda \mathcal{C}(\tau_{\text{eff}}) - \frac{1}{2} \tau_{\text{eff}}^2, \quad (33)$$

with the effective stationary condition then written as

$$\lambda = \frac{d}{d\mathcal{C}} \frac{\tau_{\text{eff}}^2}{2} = \tau_{\text{eff}} \left( \frac{d\mathcal{C}(\tau_{\text{eff}})}{d\tau_{\text{eff}}} \right)^{-1}. \quad (34)$$

Note then that since

$$\frac{d\phi_{\text{cyl/slope}}(\lambda)}{d\lambda} = \mathcal{C}(\tau_{\text{eff}}), \quad (35)$$

and hence

$$\frac{1}{2}\tau_{\text{eff}}^2 = \lambda \frac{d\phi_{\text{cyl/slope}}(\lambda)}{d\lambda} - \phi_{\text{cyl/slope}}(\lambda), \quad (36)$$

one can thus fit both the values of  $\tau_{\text{eff}}$  and  $\mathcal{C}(\tau_{\text{eff}})$  from the generating function computed for real values of  $\lambda$ . An expansion in series of  $\mathcal{C}$ ,

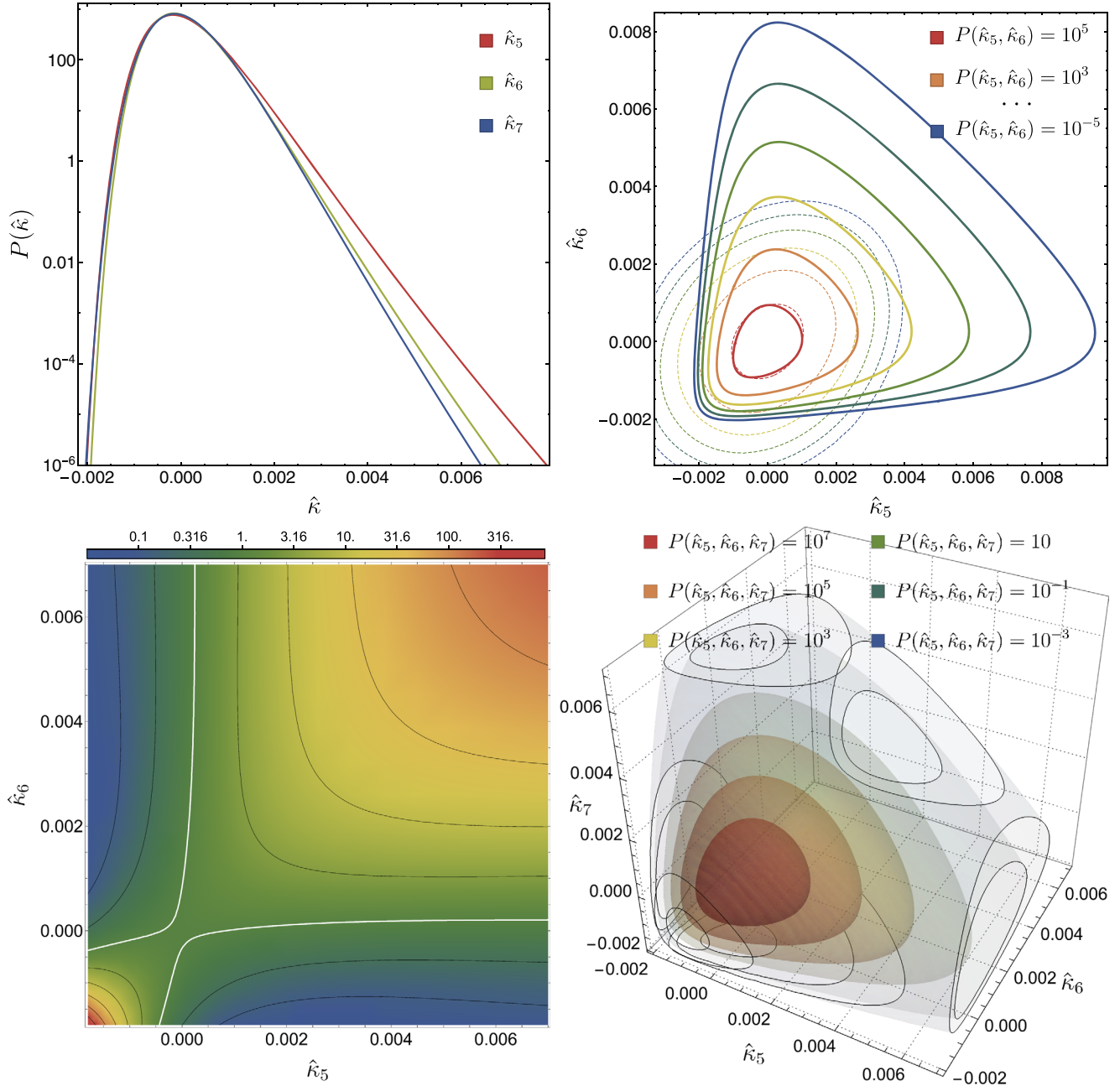


FIG. 5. (joint)-PDFs between the three nulled bins filtered by a top-hat window function of radius 15 arcmin. Top-left-hand panel: One-point  $\hat{\kappa}$ -PDF within the three nulled bins. Top-right-hand panel: Isoprobability density contours of the joint  $\hat{\kappa}$ -PDF between the two first nulled bins. The nulled fields are filtered by a top-hat window function of radius 15 arcmin. The dashed thin lines correspond to a 2-dimensional normal distribution with the same covariance. Bottom-left-hand-panel: Density plot of  $P(\hat{\kappa}_5, \hat{\kappa}_6)/P(\hat{\kappa}_5)P(\hat{\kappa}_6)$  to explicit the shape of the correlation between  $\hat{\kappa}_5$  and  $\hat{\kappa}_6$ . The black lines correspond to isocontours of values the ticks of the color bar and the white line is for the tick equal to 1. They serve as guide for the eye. Bottom-right-hand panel: Joint  $\hat{\kappa}$ -PDF between the three nulled bins.



$$\mathcal{C}(\tau_{\text{eff}}) = \sum_{k=0}^n \frac{\mu_k}{k!} \tau_{\text{eff}}^k, \quad (37)$$

typically stopping at  $n = 5$ , allows to fit the  $\mu_k$  coefficients which are closely related to the cumulants of  $\phi_{\text{cyl/slope}}$  as noted in Appendix D of [13].

This construction finally enables one to successfully compute  $\phi_{\text{cyl/slope}}$  and thus the joint cumulant generating function of projected densities  $\hat{\phi}(\{\lambda_i\})$  in Eq. (11) for any tuples of complex  $\{\lambda_i\}$ .

## B. Results

For the final part of this paper, we implement the decimation strategy presented in Sec. IV coupled with the computation of  $\hat{\phi}(\{\lambda_i\})$  of the last subsection. For simplicity and visualization purposes we restrict ourselves to the computation of the joint convergence between three successive nulled bins taken from Fig. 1, the 5th, 6th, and 7th, and which are respectively the linear combination of source redshifts located at  $z_s = 0.62 - 0.73 - 0.85$ ,  $0.73 - 0.85 - 0.96$ , and  $0.85 - 0.96 - 1.1$ . In practice, the joint PDF is computed from the CGF in a few seconds. Thus denoting by respectively  $\hat{\kappa}_5$ ,  $\hat{\kappa}_6$ , and  $\hat{\kappa}_7$  the nulled convergence in each of those bins and noticing that their joint PDF is given by Eq. (19) we obtain the results given in Fig. 5 for 1 and 2D marginals and for the full 3D PDF. As expected from the formalism, the exponential decay of the

high- $\hat{\kappa}$  tails are visible on all these plots and the highly skewed (asymmetric) PDFs hints toward the fact that the information content of the projected densities fields is not entirely probed by its 2-point correlation function which is sufficient for Gaussian random fields only.

Figure 5 also displays the shape of the correlation between  $\hat{\kappa}_5$  and  $\hat{\kappa}_6$  seen as the residual between the 2D PDF and the product of the two 1D PDFs. This residual would be exactly equal to 1 if the two nulled variables were independent. Note that this modulation of the independent-variables case is quite important and understandable as follows:  $\hat{\kappa}_5$  and  $\hat{\kappa}_6$  result from the overlapping contributions of lenses along the line of sight. As a consequence, if  $\hat{\kappa}_5$  were to take an improbable very high (resp. low) value, then the modulation would have to take that into account by raising the probability of finding a very high (resp. low) value for  $\hat{\kappa}_6$  compared to the independent-variables case.

We also show on Fig. 6 the correlation induced between  $\hat{\kappa}_5$  and  $\hat{\kappa}_7$  if a constraint on  $\hat{\kappa}_6$ , here  $\hat{\kappa}_6 > 0$ , is imposed. In the absence of constraint, these two variables are independent (since their nulled lensing kernels do not overlap) but the constraint breaks this property. In this case, one needs to compute the full 3-variable PDF before integrating over the constraint. The residual between the 2D constrained PDF and the product of the two 1D constrained PDFs is relatively small compared to the previous case (typically between 50% and 150%) and points out an overprobability, compared to the independent assumption, for  $\hat{\kappa}_5$  and  $\hat{\kappa}_7$  to

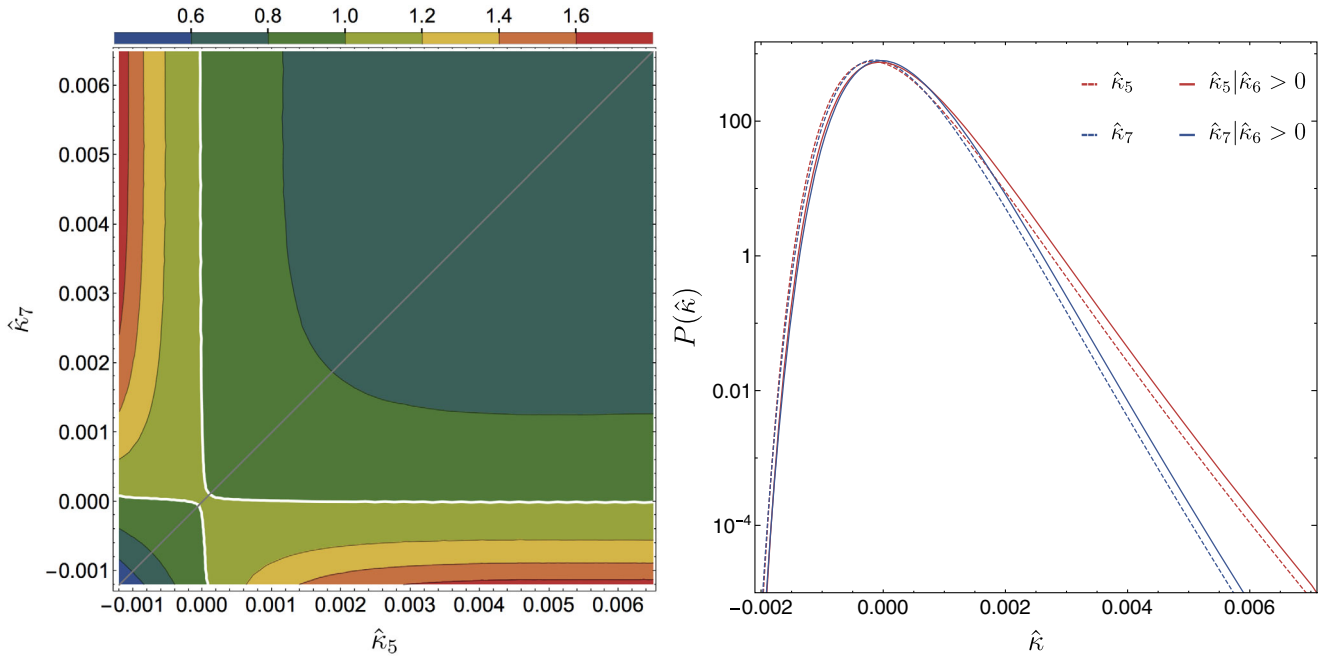


FIG. 6. Left-hand panel: Contour plot of  $P(\hat{\kappa}_5, \hat{\kappa}_7 | \hat{\kappa}_6 > 0) / P(\hat{\kappa}_5 | \hat{\kappa}_6 > 0) / P(\hat{\kappa}_7 | \hat{\kappa}_6 > 0)$  to explicit the shape of the correlation between  $\hat{\kappa}_5$  and  $\hat{\kappa}_7$  if a constraint on  $\hat{\kappa}_6$  is imposed. The white line is for the tick equal to 1. The nulled fields are filtered by a top-hat window of radius 15 arcmin. Right-hand panel: The solid lines represent  $P(\hat{\kappa}_5 | \hat{\kappa}_6 > 0)$  and  $P(\hat{\kappa}_7 | \hat{\kappa}_6 > 0)$  while the dashed lines show the same PDFs as the top-left panel of Fig. 5, that is without any constraint. Together the two panels give  $P(\hat{\kappa}_5, \hat{\kappa}_7 | \hat{\kappa}_6 > 0)$ .

have opposite rare values while same sign rare events are less likely. Note that since  $\hat{\kappa}_6$  has its own distribution the constraint  $\hat{\kappa}_6 > 0$  is more likely to give values of  $\hat{\kappa}_6$  close to zero and thus the constraint imposes, e.g., that for a large positive value of  $\hat{\kappa}_5$ ,  $\hat{\kappa}_7$  should compensate in the opposite direction to realise it. Hence the  $\hat{\kappa}_7$  distribution would be shifted toward more negative values compared to its marginal distribution. For  $\hat{\kappa}_6 = 0$ ,  $\hat{\kappa}_5$  and  $\hat{\kappa}_7$  would be exactly anticorrelated but this is not exactly the case here as seen from the fact that Fig. 6 is not exactly symmetric with respect to the  $\hat{\kappa}_5 = \hat{\kappa}_7$  axis.

## VI. CONCLUSION

This paper explores some properties of the computational complexity of the joint probability distribution function of a set of random variables only correlated by blocks. Building on this specific property we presented different strategies to effectively reduce by several orders of magnitude the number of operations needed to compute this joint probability distribution function.

In the field of cosmology and in the context of tomographic cosmic shear surveys, such a setup appears when one linearly transforms the set of observed projected densities (convergence, aperture mass) for different source redshifts into their nulled counterparts—which boils down to rearranging the information present in the survey. As such, the nulled projected densities can be seen as a new basis of pseudo-eigenvectors that render the covariance matrix block-diagonal—even in the presence of shape noise and intrinsic alignments—but which also holds the advantage of localising along the line of sight the contribution of structures that make the total signal. To these two main desirable aspects, we now finally add the reduced numerical complexity of the computation of theoretical predictions for the joint PDF of nulled variables.

Even further reduced numerical complexity could also be achieved using appropriate approximation schemes such as the one we present in Appendix B and that relies on mathematical copulas. This comes at the price of losing the exact result coming from the theory but drastically reduces the computation cost since this scheme only relies in our case on the knowledge of the univariate PDFs and the correlation coefficients, that are strongly linked to the overlap between the lensing kernels.

Regardless of the model—(semi)analytical or derived from numerical simulations—used to infer the statistical properties of the underlying matter field, the nulling formalism ought to be implemented in the context of cosmological analysis of weak lensing surveys. Indeed, while it preserves the information content of the observables, it allows us to better identify the origin of this information whether it is in terms of physical scales, redshift, and combinations of those. In this context, our results show that the actual numerical derivation of joint

PDFs is tractable even if one wants to exploit a large number of bins in order to be as precise as possible while preserving the amount of information available. Also note that the method presented in this paper can also be used to obtain the joint PDF of the regular fields via a simple change of variables once the joint PDF of nulled bins has been computed.

For the specific case of non-Gaussian observables such as the (joint) PDF of projected densities, it was shown in previous works (see [12] and [13]) that, for the desired filtering angular scales, shape noise in individual nulled bins makes the detection of non-Gaussian features virtually impossible in a single nulled map. One of the reasons of this failure is that each nulled map collects noise from three of the original maps it is built from, see [13] for details. However, the exploitation of joint PDFs now offers a potential solution to this issue as more useful signal can then be combined together while the shape noise does not increase in the same proportion since it is shared between nulled bins. This thus makes the detection of sought-after non-Gaussian features more likely.

At this stage we can sketch a possible strategy: the key to a successful exploitation of the data is our ability to select the physical scales that are contributing the most to the signal and reject the smallest scales, in practice the nearby lenses, for which we miss reliable modeling and which incidentally are likely to be strongly affected by large supersample covariance effects. The idea would then be to exploit joined nulled bins, built out of a large set of  $n_l$  bins, but *a priori* ignoring the two closest for which the physical scales are still strongly mingled. Hopefully the  $n_l - 2$  farthest bins would be able to provide us with significant detection of the nonlinear couplings in a regime where they are well understood. And if not, it would point to the fact that the non-Gaussian information that we observe in regular bins is coming from not very well understood nor modelled nonlinear scales, which would hence call to extreme caution when using them to extract information. In any case, the nulling formalism and the joint PDF we presented would play a key role in answering these questions. Those studies are however beyond the scope of this paper and will be performed elsewhere.

## ACKNOWLEDGMENTS

This work is partially supported by the SPHERES Grant No. ANR-18-CE31-0009 of the French *Agence Nationale de la Recherche* and by Fondation MERAC. A. B.'s work is supported by a fellowship from CNES. This work has made use of the Horizon Cluster hosted by Institut d'Astrophysique de Paris. We thank Stéphane Rouberol for running smoothly this cluster for us. We also thank Oliver Friedrich, Aoife Boyle, Alex Gough and Raphaël Gavazzi for fruitful discussions and comments on the manuscript.

## APPENDIX A: REWRITING THE TRIVARIATE PDF AS FUNCTIONAL OF BIVARIATE PDFS

We aim here at expressing  $\mathcal{P}(\hat{\kappa}_1, \hat{\kappa}_2, \hat{\kappa}_3)$  purely as a function of  $\mathcal{P}(\hat{\kappa}_1, \hat{\kappa}_2)$  and  $\mathcal{P}(\hat{\kappa}_2, \hat{\kappa}_3)$ . We know this is technically possible given Eq. (18).

We start by writing the CGFs used as building blocks for the joint multivariate PDF in terms of univariate and bivariate PDFs using Laplace transformations as

$$e^{\hat{\phi}(\lambda_2)} = \int d\hat{\kappa}'_2 \exp(\hat{\kappa}'_2 \lambda_2) \mathcal{P}(\hat{\kappa}'_2), \quad (\text{A1})$$

$$e^{\hat{\phi}_{\kappa_1}(\lambda_2)} = \int d\hat{\kappa}'_2 \exp(\hat{\kappa}'_2 \lambda_2) \mathcal{P}(\hat{\kappa}_1, \hat{\kappa}'_2), \quad (\text{A2})$$

and similarly for  $\exp(\hat{\phi}_{\kappa_3}(\lambda_2))$ . We insert these into the expression for the trivariate PDF built from pairwise CGFs from Eq. (19) and thus finally obtain

$$\begin{aligned} & \mathcal{P}(\hat{\kappa}_1, \hat{\kappa}_2, \hat{\kappa}_3) \\ &= \int \frac{d\lambda_2}{2\pi i} e^{-\hat{\kappa}_2 \lambda_2} \frac{\int d\hat{\kappa}'_2 e^{\hat{\kappa}'_2 \lambda_2} [\mathcal{P}(\hat{\kappa}_1, \cdot) * \mathcal{P}(\cdot, \hat{\kappa}_3)](\hat{\kappa}'_2)}{\int d\hat{\kappa}''_2 e^{\hat{\kappa}''_2 \lambda_2} \mathcal{P}(\hat{\kappa}'_2)} \end{aligned} \quad (\text{A3})$$

where  $\mathcal{P}(\hat{\kappa}_1, \cdot) * \mathcal{P}(\cdot, \hat{\kappa}_3)$  denotes the convolution of  $\mathcal{P}(\hat{\kappa}_1, \hat{\kappa}_2)$  and  $\mathcal{P}(\hat{\kappa}_2, \hat{\kappa}_3)$  with respect to  $\hat{\kappa}_2$ . This is a result of the convolution theorem stating that the product of the Laplace transforms of two functions can be written as the Laplace transform of their convolution. As expected from the somewhat complicated relationship between PDFs and CGFs, Eq. (A3) is unfortunately not nearly as simple as Eq. (18) though they both depict the exact same result without any loss of information.

## APPENDIX B: JOINT PDFS IN TERMS OF COPULAS

### 1. Dissecting joint PDFs into marginals and copulas

According to Sklar's theorem [28], any multivariate joint PDF  $P(\{\hat{\kappa}_i\})$  can be written in terms of univariate marginal PDFs  $P(\hat{\kappa}_i)$  and a copula which describes the dependence structure between the variables. Copulas are well suited for high-dimensional statistical applications because they allow one to model and estimate joint probability distributions by estimating marginals and copulas separately. Copulas can be obtained from a range of parametric models that encode the dependence structure using just a few parameters and then compress the information of a high-dimensional multivariate PDF down to a set of uni- or bivariate marginal PDFs and a set of correlation parameters.

To fully disentangle the marginals from the correlation structure, one uses the values of the marginal cumulative distribution function (CDF)  $u_i = C(\hat{\kappa}_i) \in [0, 1]$  as variables. Those so-called filling factors  $u_i$  are simply obtained from

$$u_i(\hat{\kappa}_i) = \mathcal{C}(\hat{\kappa}_i) = \int^{\hat{\kappa}_i} d\hat{\kappa}'_i \mathcal{P}(\hat{\kappa}'_i). \quad (\text{B1})$$

The marginal CDFs provide a ranking of variables in terms of cumulative probability and the filling factors  $u_i$  allow to identify quantiles easily. The multivariate copula  $C$  is now defined as the joint CDF  $\mathcal{C}$  expressed in the new ranked variables  $u_i$

$$\mathcal{C}(\{u_i\}) = \mathcal{C}(\{\hat{\kappa}_i = C^{-1}(u_i)\}), \quad (\text{B2})$$

Similarly as the PDF is obtained as derivative of the CDF, the copula density  $c$  is obtained from a derivative of the copula

$$c(u_1, \dots, u_n) = \frac{\partial \mathcal{C}(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n} \quad (\text{B3})$$

$$= \frac{\mathcal{P}(\hat{\kappa}_1, \dots, \hat{\kappa}_n)}{\mathcal{P}(\hat{\kappa}_1) \dots \mathcal{P}(\hat{\kappa}_n)} \Big|_{\hat{\kappa}_i = C^{-1}(u_i)}, \quad (\text{B4})$$

which by the chain rule corresponds to the ratio of the joint PDF and the would-be joint PDF if the variables  $\hat{\kappa}_i$  were independent. The joint PDF can be reconstructed from the set of univariate marginals  $\{\mathcal{P}(\hat{\kappa}_i)\}$  and the multivariate copula.

For a bivariate joint PDF, one can write

$$\mathcal{P}(\hat{\kappa}_i, \hat{\kappa}_j) = \mathcal{P}(\hat{\kappa}_i) \mathcal{P}(\hat{\kappa}_j) c_{ij}(\mathcal{C}(\hat{\kappa}_i), \mathcal{C}(\hat{\kappa}_j)), \quad (\text{B5})$$

where  $c_{ij}$  is the copula density and if  $\hat{\kappa}_i$  and  $\hat{\kappa}_j$  are independent, then  $c_{ij} = 1$ . There are many parametric copula families available to encode the dependence between  $\mathcal{C}(\hat{\kappa}_i)$  and  $\mathcal{C}(\hat{\kappa}_j)$  using a correlation parameter controlling the strength of dependence. A particularly simple case is the *Gaussian bivariate copula*, which is controlled by the pairwise correlation coefficient  $r_{ij}$

$$\mathcal{P}(\hat{\kappa}_i, \hat{\kappa}_j) = \mathcal{P}(\hat{\kappa}_i) \mathcal{P}(\hat{\kappa}_j) c_{ij}^{\text{Gauss}}(r_{ij})(\mathcal{C}(\hat{\kappa}_i), \mathcal{C}(\hat{\kappa}_j)). \quad (\text{B6})$$

We demonstrate a reconstruction of the joint PDF  $\mathcal{P}(\hat{\kappa}_5, \hat{\kappa}_6)$  using the marginals  $\mathcal{P}(\hat{\kappa}_{5/6})$  and a Gaussian copula density in Fig. 7. It can be seen overall that the Gaussian copula approximation agrees very well with the exact LDT calculation in the bulk of the PDF and a bit less in the high and low convergence tails but with still a much better agreement than if we were to consider that the two  $\hat{\kappa}$  bins were independent. A more rigorous assessment of the validity of the approximation would first require to assess the validity of the LDT/copula approaches with real or simulated data and then quantify the cosmological information content present in the bulk and tails of the PDFs, for example pursuing a fisher analysis as in [5].

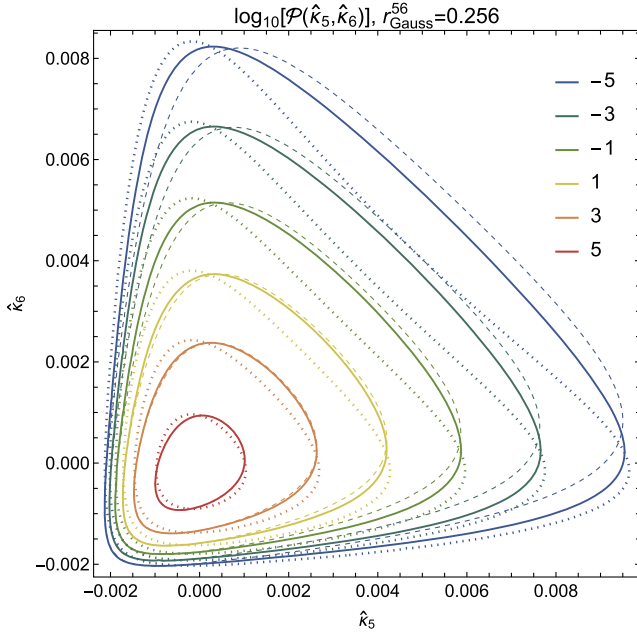


FIG. 7. Contour plot of the joint PDF  $P(\hat{\kappa}_5, \hat{\kappa}_6)$  between two neighboring nulled convergence bins as determined by the LDT prediction for  $P(\hat{\kappa}_5, \hat{\kappa}_6)$  (solid lines), the reconstructed PDF using the LDT prediction for the marginals  $P(\hat{\kappa}_5)$  and  $P(\hat{\kappa}_6)$  along with a Gaussian copula with correlation coefficient  $r_{\text{Gauss}}$  obtained from the LDT prediction (dashed lines) and the result one would obtain if the neighboring  $\hat{\kappa}$  bins were independent (dotted lines).

## 2. Measures of correlation

The linear correlation coefficient is defined as usual as

$$r_{ij} = \text{Corr}(\hat{\kappa}_i, \hat{\kappa}_j) = \frac{\langle \hat{\kappa}_i \hat{\kappa}_j \rangle}{\sqrt{\langle \hat{\kappa}_i^2 \rangle \langle \hat{\kappa}_j^2 \rangle}}. \quad (\text{B7})$$

In our particular case, we note that the value of the correlation coefficients  $r_{ij}$  will be closely related to the overlap between the nulled lensing kernels. For the case of discrete source planes, we found that the linear correlation coefficients measured from the joint bivariate PDFs are  $r_{56} \simeq r_{67} \simeq 0.25$ , which reflects the overlap between the neighboring lensing kernels that are of the order of 25% in physical space quantified by the comoving distance rather than redshift as shown in Fig. 1.

For the modeling of copulas, one typically uses *ranked correlation coefficients* such as Spearman's  $\rho$ . Spearman's  $\rho$  measures the correlation between the filling factors (obtained from the marginal CDFs),

$$\begin{aligned} \rho_{S,ij} &:= \text{Corr}(\mathcal{C}(\hat{\kappa}_i), \mathcal{C}(\hat{\kappa}_j)) \\ &= 12 \int_0^1 du_i \int_0^1 du_j C(u_i, u_j) - 3. \end{aligned} \quad (\text{B8})$$

It can be thought of as the linear correlation coefficient of abundance-matched  $\hat{\kappa}$ -values. For a Gaussian bivariate copula, Spearman- $\rho$  is obtained as  $\rho_S = \frac{6}{\pi} \arcsin(r/2)$

which is close to the linear correlation  $r$ . We invert this relation to obtain the Gaussian pairwise correlation coefficients  $r_{ij}^{\text{Gauss}} = 2 \sin(\pi \rho_{S,ij}/6)$  given the rank correlations  $\rho_{S,ij}$ .

If one is more interested in the tails of the distribution, one would like to know whether with large/small values of  $\hat{\kappa}_i$  also large/small values of  $\hat{\kappa}_j$  are expected. This can be quantified using the tail concentration function defined for every value of cumulative probability  $u$  as

$$\lambda_C(u) = \begin{cases} \frac{C(u,u)}{u} & u \in [0, 0.5[ \\ 2 + \frac{C(u,u)-1}{1-u} & u \in [0.5, 1]. \end{cases} \quad (\text{B9})$$

The tail concentration measures the slope of the copula and upper or lower tail dependence is signaled if  $\lambda_C(u \rightarrow 0) > 0$  or  $\lambda_C(u \rightarrow 1) > 0$ , respectively. We show the tail correlation for the joint PDF of neighboring  $\hat{\kappa}$  bins in Fig. 8. While they are positively correlated due to the overlap of their nulled lensing kernels, they do not exhibit upper or lower tail dependence, as  $\lambda_C \rightarrow 0$  as one approaches the rare event tails  $\mathcal{C}(\hat{\kappa}) \rightarrow 0$  and  $\mathcal{C}(\hat{\kappa}) \rightarrow 1$ , respectively. The correlation between regions of similar rarity is well described by a Gaussian copula (dashed lines), which does not exhibit any tail dependence. Since the  $\hat{\kappa}_i$  have similar CDFs, the tail concentration is related to a diagonal slice of the joint PDF  $\mathcal{P}(\kappa_i, \hat{\kappa}_j)$  such as the one shown in Fig. 7.

Note that while a Gaussian bivariate copula does not exhibit any tail dependence, conditionals obtained from a Gaussian multivariate copula or a Gaussian conditional copula can exhibit some tail dependence. We illustrate this

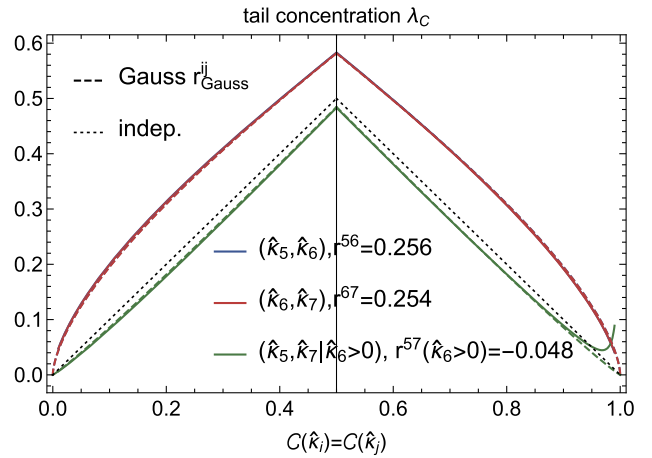


FIG. 8. Tail correlation between nulled  $\hat{\kappa}$  bins as predicted from LDT (solid lines) and reconstructed from Gaussian copulas with measured correlation (dashed lines) in comparison to the independent result (black dotted line). We show the almost identical results for two neighboring bin pairs  $\hat{\kappa}_{5/6}$  (blue) and  $\hat{\kappa}_{6/7}$  (red) with virtually identical positive correlation and separated bins  $\hat{\kappa}_{5/7}$  that are coupled through a constraint on the intermediate bin  $\hat{\kappa}_6 \geq 0$  leading to a slight negative correlation (green).

point with the tail correlation between physically separated nulled bins  $\hat{\kappa}_5$  and  $\hat{\kappa}_7$  that is caused by a positivity constraint on the intermediate ‘‘coupling’’ bin  $\hat{\kappa}_6 > 0$  as the green line in Fig. 8.

### 3. Modeling the trivariate PDF with a copula

#### a. Gaussian multivariate copula

A Gaussian multivariate copula is controlled by a *correlation matrix* containing the pairwise correlations  $r_{ij}$ . For the case of discrete source planes, only neighboring nulled  $\hat{\kappa}$ -bins are correlated. The simplest starting point to extract information from an  $n$ -dimensional multivariate PDF would then be to rely on the set of  $n$  univariate marginals  $\{\mathcal{P}(\hat{\kappa}_i)\}$  along with  $n - 1$  correlation coefficients.

When comparing the full joint PDF to the ‘‘reconstruction’’ based on marginals and a multivariate Gaussian copula with the correlation coefficients as input, one can get an idea of how much information is lost. For example, the trivariate PDF  $\mathcal{P}(\hat{\kappa}_5, \hat{\kappa}_6, \hat{\kappa}_7)$  considered in the main text would be reconstructed using three univariate marginals and a trivariate Gaussian copula with correlation coefficients  $r_{56}$ ,  $r_{67}$ , and  $r_{57} = 0$

$$\begin{aligned} \mathcal{P}(\hat{\kappa}_5, \hat{\kappa}_6, \hat{\kappa}_7) &= \mathcal{P}(\hat{\kappa}_5)\mathcal{P}(\hat{\kappa}_6)\mathcal{P}(\hat{\kappa}_7) \\ &\times c_{567}^{\text{Gauss}}(r_{56}, r_{67})(\mathcal{C}(\hat{\kappa}_5), \mathcal{C}(\hat{\kappa}_6), \mathcal{C}(\hat{\kappa}_7)), \end{aligned} \quad (\text{B10})$$

where the correlation coefficients  $r_{56}$  and  $r_{67}$  are determined from the joint bivariate PDFs as discussed in the context of Eq. (B6).

#### b. Conditional Gaussian bivariate copula

The trivariate PDF  $\mathcal{P}(\hat{\kappa}_5, \hat{\kappa}_6, \hat{\kappa}_7)$  is constructed from the bivariate CGFs  $\hat{\phi}(\lambda_5, \lambda_6)$  and  $\hat{\phi}(\lambda_6, \lambda_7)$  and the univariate CGF  $\hat{\phi}(\lambda_6)$ . Since those CGFs can be constructed from the bivariate PDFs  $\mathcal{P}(\hat{\kappa}_5, \hat{\kappa}_6)$  and  $\mathcal{P}(\hat{\kappa}_6, \hat{\kappa}_7)$  along with the univariate PDF  $\mathcal{P}(\hat{\kappa}_6)$ , it is natural to look for a reconstruction of the trivariate PDF in terms of bivariate PDFs and a conditional bivariate copula.

$$\begin{aligned} \mathcal{P}(\hat{\kappa}_5, \hat{\kappa}_6, \hat{\kappa}_7) &= \frac{\mathcal{P}(\hat{\kappa}_5, \hat{\kappa}_6)\mathcal{P}(\hat{\kappa}_6, \hat{\kappa}_7)}{\mathcal{P}(\hat{\kappa}_6)} \\ &\times c_{57|6}(\hat{\kappa}_6)(\mathcal{C}(\hat{\kappa}_5|\hat{\kappa}_6), \mathcal{C}(\hat{\kappa}_7|\hat{\kappa}_6)). \end{aligned} \quad (\text{B11})$$

For an illustration of this reconstruction, we will assume the conditional bivariate copula  $c_{57|6}(\hat{\kappa}_6) = c_{57|6}^{\text{Gauss}}(r_{57|6}(\hat{\kappa}_6))$  to

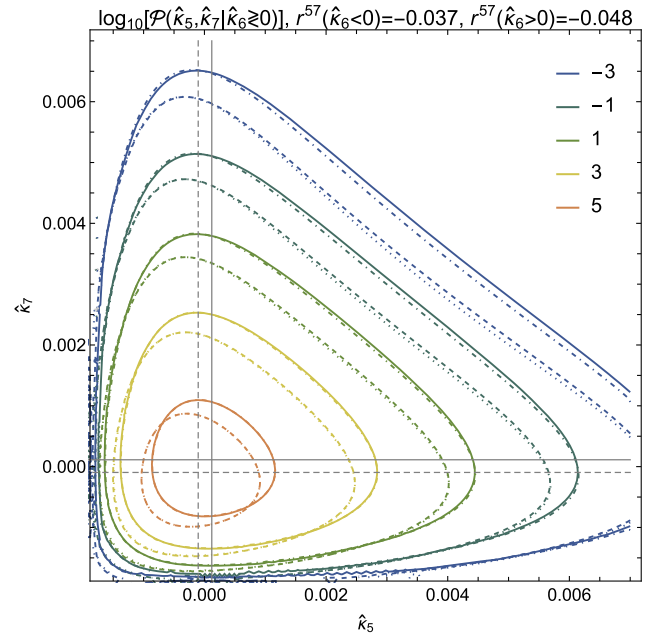


FIG. 9. Contour plot of the joint conditional PDF  $\mathcal{P}(\hat{\kappa}_5, \hat{\kappa}_7|\hat{\kappa}_6 \geq 0)$  between two nulled convergence bins coupled by a constraint on the intermediate bin  $\hat{\kappa}_6 \geq 0$  for the LDT prediction (solid/dashed lines) and the reconstructed PDF using the LDT prediction for the bivariate marginals  $\mathcal{P}(\hat{\kappa}_5, \hat{\kappa}_6)$  and  $\mathcal{P}(\hat{\kappa}_6, \hat{\kappa}_7)$  along with a Gaussian conditional copula with correlation coefficients  $r^{57}(\hat{\kappa}_6 \geq 0)$  from LDT (dot-dashed/dotted lines). We also show the means of  $\hat{\kappa}_{5/7}$  induced by  $\hat{\kappa} > 0$  (solid grey lines) and  $\hat{\kappa} < 0$  (dashed grey lines).

be Gaussian and hence given by a correlation coefficient function  $r_{57|6}(\hat{\kappa}_6)$  encoding the correlation coefficient between  $\hat{\kappa}_5$  and  $\hat{\kappa}_7$  given a value (or range) for  $\hat{\kappa}_6$ . While multivariate parametric copula models are only available for a few special classes such as Gaussian copulas, there is a plethora of bivariate parametric copula models that allows for more flexibility and could be used to refine the present model. We use the conditional Gaussian copula to reconstruct the bivariate conditional PDF  $\mathcal{P}(\hat{\kappa}_5, \hat{\kappa}_7|\hat{\kappa}_6 \geq 0)$  where a coupling between  $\hat{\kappa}_5$  and  $\hat{\kappa}_7$  is induced by a constraint on  $\hat{\kappa}_6 \geq 0$ , see Fig. 9. The good qualitative match of the Gaussian bivariate copula along a diagonal slice where  $\mathcal{C}(\hat{\kappa}_5) = \mathcal{C}(\hat{\kappa}_7)$  can also be observed in the tail correlation shown in Fig. 8. Interestingly, we see that putting a positivity constraint on  $\hat{\kappa}_6$  induces some positive tail dependence in the LDT prediction that is not captured by a simple Gaussian bivariate copula.

- [1] M. Kilbinger, Cosmology with cosmic shear observations: A review, *Rep. Prog. Phys.* **78**, 086901 (2015).
- [2] Ž. Ivezić, S. M. Kahn, J. A. Tyson, B. Abel, E. Acosta, R. Allsman, D. Alonso, Y. AlSayyad, S. F. Anderson, J. Andrew *et al.*, LSST: From science drivers to reference design and anticipated data products, *Astrophys. J.* **873**, 111 (2019).
- [3] R. Laureijs, J. Amiaux, S. Arduini, J. L. Auguères, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, and A. Ealet, Euclid definition study report, [arXiv:1110.3193](https://arxiv.org/abs/1110.3193).
- [4] K. Paton, J. Blazek, K. Honscheid, E. Huff, P. Melchior, A. J. Ross, and E. Suchyta, Cosmological constraints from the convergence 1-point probability distribution, *Mon. Not. R. Astron. Soc.* **472**, 439 (2017).
- [5] A. Boyle, C. Uhlemann, A. Barthelemy, O. Friedrich, S. Codis, and F. Bernardeau, Extracting cosmology from the weak lensing convergence PDF, *Mon. Not. R. Astron. Soc.* **505**, 2886 (2021).
- [6] L. Thiele, J. C. Hill, and K. M. Smith, Accurate analytic model for the weak lensing convergence one-point probability distribution function and its autocovariance, *Phys. Rev. D* **102**, 123545 (2020).
- [7] F. Bernardeau and P. Valageas, Construction of the one-point PDF of the local aperture mass in weak lensing maps, *Astron. Astrophys.* **364**, 1 (2000), <https://ui.adsabs.harvard.edu/abs/2000A%26A...364....1B/abstract>.
- [8] P. Reimberg and F. Bernardeau, Large deviation principle at work: Computation of the statistical properties of the exact one-point aperture mass, *Phys. Rev. D* **97**, 023524 (2018).
- [9] F. Bernardeau and P. Reimberg, Large deviation principle at play in large scale structure cosmology, *Phys. Rev. D* **94**, 063520 (2016).
- [10] C. Uhlemann, S. Codis, C. Pichon, F. Bernardeau, and P. Reimberg, Back in the saddle: Large-deviation statistics of the cosmic log-density field, *Mon. Not. R. Astron. Soc.* **460**, 1529 (2016).
- [11] O. Friedrich *et al.* (DES Collaboration), Density split statistics: Joint model of counts and lensing in cells, *Phys. Rev. D* **98**, 023508 (2018).
- [12] A. Barthelemy, S. Codis, C. Uhlemann, F. Bernardeau, and R. Gavazzi, A nulling strategy for modelling lensing convergence in cones with large deviation theory, *Mon. Not. R. Astron. Soc.* **492**, 3420 (2020).
- [13] A. Barthelemy, S. Codis, and F. Bernardeau, Probability distribution function of the aperture mass field with large deviation theory, *Mon. Not. R. Astron. Soc.* **503**, 5204 (2021).
- [14] F. Bernardeau, T. Nishimichi, and A. Taruya, Cosmic shear full nulling: Sorting out dynamics, geometry and systematics, *Mon. Not. R. Astron. Soc.* **445**, 1526 (2014).
- [15] P. L. Taylor, F. Bernardeau, and T. D. Kitching, k-cut cosmic shear: Tunable power spectrum sensitivity to test gravity, *Phys. Rev. D* **98**, 083514 (2018).
- [16] P. L. Taylor, F. Bernardeau, and E. Huff, x-cut Cosmic shear: Optimally removing sensitivity to baryonic and nonlinear physics with an application to the Dark Energy Survey year 1 shear data, *Phys. Rev. D* **103**, 043531 (2021).
- [17] A. S. Maniyar, E. Schaaf, and A. R. Pullen, A new probe of the high-redshift Universe: Nulling CMB lensing with interloper-free “LIM-pair” lensing, [arXiv:2106.09005](https://arxiv.org/abs/2106.09005).
- [18] Y. Mellier, Probing the universe with weak lensing, *Annu. Rev. Astron. Astrophys.* **37**, 127 (1999).
- [19] N. Kaiser, Nonlinear cluster lens reconstruction, *Astrophys. J. Lett.* **439**, L1 (1995).
- [20] P. Schneider, Detection of (dark) matter concentrations via weak gravitational lensing, *Mon. Not. R. Astron. Soc.* **283**, 837 (1996).
- [21] F. Bernardeau, The evolution of the large-scale structure of the universe: Beyond the linear regime, [arXiv:1311.2724](https://arxiv.org/abs/1311.2724).
- [22] To be more precise the only nonzero elements of  $p_{ij}$  satisfy  $i - 2 \leq j \leq i$ .
- [23] A. C. Deshpande *et al.*, Euclid: The reduced shear approximation and magnification bias for Stage IV cosmic shear experiments, *Astron. Astrophys.* **636**, A95 (2020).
- [24] M. A. Troxel and M. Ishak, The intrinsic alignment of galaxies and its impact on weak gravitational lensing in an era of precision cosmology, *Phys. Rep.* **558**, 1 (2015).
- [25] S. Codis, R. Gavazzi, Y. Dubois, C. Pichon, K. Benabed, V. Desjacques, D. Pogosyan, J. Devriendt, and A. Slyz, Intrinsic alignment of simulated galaxies in the cosmic web: implications for weak lensing surveys, *Mon. Not. R. Astron. Soc.* **448**, 3391 (2015).
- [26] M. Asgari *et al.*, KiDS-1000 cosmology: Cosmic shear constraints and comparison between two point statistics, *Astron. Astrophys.* **645**, A104 (2021).
- [27] J. Harnois-Déraps, N. Martinet, T. Castro, K. Dolag, B. Giblin, C. Heymans, H. Hildebrandt, and Q. Xia, Cosmic shear cosmology beyond two-point statistics: A combined peak count and correlation function analysis of DES-Y1, *Mon. Not. R. Astron. Soc.* **506**, 1623 (2021).
- [28] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, *Publ. l’Inst. Stat. l’Univ. Paris* **8**, 229 (1959).