

Theoretical and numerical perspectives on cosmic distance averages

Michel-Andrès Breton¹ and Pierre Fleury²

¹ Aix Marseille Univ., CNRS, CNES, LAM, Marseille, France
e-mail: michel-andres.breton@lam.fr

² Instituto de Física Teórica UAM-CSIC, Universidad Autónoma de Madrid, Cantoblanco 28049, Madrid, Spain
e-mail: pierre.fleury@uam.es

Received 16 December 2020 / Accepted 15 August 2021

ABSTRACT

The interpretation of cosmological observations relies on a notion of an average Universe, which is usually considered as the homogeneous and isotropic Friedmann-Lemaître-Robertson-Walker (FLRW) model. However, inhomogeneities may statistically bias the observational averages with respect to FLRW, notably for distance measurements, due to a number of effects such as gravitational lensing and redshift perturbations. In this article, we review the main known theoretical results on average distance measures in cosmology, based on second-order perturbation theory, and we fill in some of their gaps. We then comprehensively test these theoretical predictions against ray tracing in a high-resolution dark-matter N -body simulation. This method allows us to describe the effect of small-scale inhomogeneities deep into the non-linear regime of structure formation on light propagation up to $z = 10$. We find that numerical results are in remarkably good agreement with theoretical predictions in the limit of super-sample variance. No unexpectedly large bias originates from very small scales, whose effect is fully encoded in the non-linear power spectrum. Specifically, the directional average of the inverse amplification and the source-averaged amplification are compatible with unity; the change in area of surfaces of constant cosmic time is compatible with zero; the biases on other distance measures, which can reach slightly less than 1% at high redshift, are well understood. As a side product, we also confront the predictions of the recent finite-beam formalism with numerical data and find excellent agreement.

Key words. large-scale structure of Universe – distance scale – cosmology: theory – methods: numerical

1. Introduction

On very large scales, our Universe seems to be well described by a spatially homogeneous and isotropic Friedmann-Lemaître-Robertson-Walker (FLRW) model (Green & Wald 2014). This model allows us to predict the dynamics of cosmic expansion as a function of the Universe's content, and of the laws of gravitation. Furthermore, the FLRW model constitutes a rather efficient framework to interpret the observation of remote light sources; in particular, it provides the relation between their redshift z and their angular or luminosity distance D .

The distance-redshift relation $D(z)$ is prominent in cosmology, as it is involved in the interpretation of various observables. Its first derivative today defines the Hubble-Lemaître constant, $dD/dz|_0 = c/H_0$, whose exact value is still subject to a lively debate (Planck Collaboration VI 2020; Riess et al. 2019; Wong et al. 2019; Freedman et al. 2019). More generally, $D(z)$ constitutes the essence of the Hubble diagram of type-Ia supernovae (SNe, Scolnic et al. 2018; Abbott et al. 2019), which historically revealed the acceleration of cosmic expansion (Perlmutter & Aldering 1998; Riess et al. 1998), as well as the Hubble diagram of gravitational-wave standard sirens in the near future (Holz & Hughes 2005; Caprini & Tamanini 2016). The $D(z)$ relation is also essential in the analysis of the anisotropies of the cosmic microwave background (CMB, Planck Collaboration VI 2020), or in the baryon-acoustic oscillation signal observed in galaxy, Lyman- α or quasar surveys (Alam et al. 2021), because it converts the observed angular size of the

sound horizon θ_* into a physical distance $r_s = D(z_*)\theta_*$ that may be predicted by theory.

In the actual inhomogeneous Universe, however, the $D(z)$ relation is affected by various effects, such as gravitational lensing (Schneider et al. 1992) which tends to focus and distort light beams, thereby changing the apparent size and brightness of light sources; it is also affected by the peculiar velocities of the sources and the observer, which correct the observed redshift via the Doppler effect (Hui & Greene 2006; Davis et al. 2011). Such effects make $D(z)$ line-of-sight dependent, but it is generally assumed that the FLRW prediction is recovered on average.

The fundamental question of whether the average $D(z)$ is the same as the $D(z)$ of the average Universe goes back more than 50 years, when Zel'dovich (1964) and Feynman (in a colloquium given at Caltech the same year)¹ suggested the following: if the Universe is lumpy, then a typical light beam should mostly propagate through under-dense regions, and thereby be de-focussed with respect to FLRW; this should imply that $D(z)$ is actually biased up. Many developments and counter-arguments followed from that seminal idea; we refer the interested reader to the introduction of Kaiser & Peacock (2016, hereafter KP16) and the comprehensive review by Helbig (2020) for details.

In that debate, a significant step was made by Weinberg (1976), who showed that in a Universe sparsely filled with point masses, the average flux $\langle 1/D^2(z) \rangle$ is the same as if the matter in those lenses were homogeneously distributed in space.

¹ This talk was mentioned in the introduction of Gunn (1967).

Importantly, Weinberg's calculation was made at first order in the small projected density of the lenses². As such, it also implies that $\langle D(z) \rangle$ is unaffected by inhomogeneities at that order, because the difference between $\langle 1/D^2 \rangle$ and $1/\langle D \rangle^2$ only appears at second order. Weinberg nevertheless conjectured, on the basis of flux conservation, that the invariance of $\langle 1/D^2(z) \rangle$ may be exact and hold for any matter distribution. As noted by Ellis et al. (1998), this general flux-conservation argument is, in fact, incomplete because it implicitly assumes that the area of surfaces of constant redshift are unaffected by inhomogeneities, which is a mere reformulation of the whole problem.

For a long period of time, all this discussion remained mostly centred on the observation of individual sources, with the aim of predicting possible biases on the Hubble diagram; it became a somewhat marginal topic from the end of the 1980s, presumably because the precision of cosmological measurements was not sufficient to be sensitive to the expected biases on $\langle D(z) \rangle$. Interest in that matter was nevertheless revived by Clarkson et al. (2014), who made the rather surprising claim that lensing affects the distance to the last-scattering surface (LSS) at percent level, which would be dramatic for the standard interpretation of the CMB. This claim was then retracted by (almost) the same team in Bonvin et al. (2015a), who with KP16 clarified that: (i) the average distance to the LSS is not relevant to the standard CMB analysis; and (ii) one must distinguish between the concepts of directional averaging, source-averaging, or ensemble-averaging, which may yield different results (Bonvin et al. 2015b). Such considerations on cosmological averages were actually elaborating on an earlier work by Kibble & Lieu (2005).

In the end, for the CMB just as for the Hubble diagram, the whole problem boils down to the validity of Weinberg's conjecture which states that the area of LSS, A_* , or the area of constant-redshift surfaces, $A(z)$, are not significantly affected by inhomogeneities. KP16 undertook the difficult task to explicitly check this conjecture in the framework of cosmological perturbations at second order. With a rather intuitive approach, KP16 identified several key effects such as the shortening of the radius reached by rays due to their deflection, or the increase in A_* due to its wrinkles, and eventually reached the conclusion that A_* cannot be biased by more than a part in a million. They identified that the relevant structures responsible for such a bias are rather large in size, of the order of $50 h^{-1}$ Mpc.

Most of the theoretical work depicted above was done using cosmological perturbation theory on an FLRW background (see also Sasaki 1987; Bonvin et al. 2006; Ben-Dayan et al. 2012; Umeh et al. 2014; Yoo & Scaccabarozzi 2016). However, this theoretical framework is not guaranteed to provide a good representation of the Universe, as it does not access the highly non-linear regime of structure formation. That is why one may prefer to rely on numerical simulations and ray-tracing methods, in order to accurately describe the propagation of light in a realistic picture of the cosmos.

As a first step, a significant research endeavour was dedicated to ray tracing and distance measurements in cosmological toy-models, such as Swiss-cheese models (Brouzakis et al. 2007, 2008; Marra et al. 2008; Biswas & Notari 2008; Vanderveld et al. 2008; Valkenburg 2009; Clifton & Zuntz 2009; Bolejko 2009, 2011; Bolejko & C el erier 2010; Szybka 2011; Flanagan et al. 2013; Fleury et al. 2013; Fleury 2014; Troxel et al. 2014; Peel

et al. 2014; Lavinto & Rasanen 2015; Koksang 2017, 2019a,b, 2020a), plane-parallel Universes (Di Dio et al. 2012), or lattice models (Clifton & Ferreira 2009a,b, 2011; Clifton et al. 2012; Liu 2015; Bruneton & Larena 2013; Bentivegna et al. 2017; Sanghai et al. 2017; Koksang 2020b). These works generally agreed with the relevant theoretical predictions. Using N -body simulations, Odderskov et al. (2016) showed that at low redshift ($z < 0.1$), the averaged luminosity distance is very close from its value in an FLRW background. Within the field of numerical relativity, Giblin et al. (2016) showed that $\langle \log D \rangle$ (or averaged magnitude) was not affected by inhomogeneities, at least until $z = 1.5$. However, both studies used simulations with rather low resolution, which might be subject to large variance and therefore could not highlight second-order effects. More recently, Adamek et al. (2019) used the general-relativistic simulation *gevolution* (Adamek et al. 2016), and accurate ray tracing to find null geodesics between sources and observer and produce realistic halo catalogues. They found that when averaging over sources, $\langle 1/D^2(z) \rangle$ is very close to its value from a homogeneous Universe, until $z = 3$, thereby confirming Weinberg's conjecture, while $\langle D(z) \rangle$ is slightly biased as expected. Albeit a high-resolution, *gevolution* remains a particle-mesh code without adaptive-mesh refinement, which thus cannot access very small scales.

In the present article, we propose a short theoretical review, and a detailed numerical study of the bias in the distance-redshift relation with respect to the standard FLRW prediction. The theory part builds upon KP16 and fills minor conceptual gaps therein. In the main, numerical, part we use a high-resolution N -body simulation part of the 'Raygal' suite and propagate photons on null geodesics to infer distance measures, accounting for gravitational lensing and redshift perturbations. Taking advantage of very large statistics and wide redshift range (up to $z = 10$), we investigate the different averaging procedures and study the statistics to the related observables. Furthermore, we numerically estimate the area bias depending on the choice of light-cone slicing.

The article is organised as follows. Section 2 presents the formalism for light propagation and the bias on statistical quantities with respect to the homogeneous case, depending on the averaging procedure; we connect these notions to the area of slices of the light cone. The numerical simulation, ray-tracing methods, and analysis techniques, are presented in Sect. 3, while the results are exposed in Sect. 4. We conclude in Sect. 5.

Notation and conventions. Greek indices (μ, ν, \dots) run from 0 to 3 and Latin indices (i, j, \dots) from 1 to 3. Bold symbols denote Euclidean two-dimensional or three-dimensional vectors, and matrices. Over-barred symbols denote quantities computed in a homogeneous-isotropic FLRW model. We adopt units in which the speed of light is unity, $c = 1$.

2. Theory

This section gathers a number of already-established theoretical results about light propagation in the inhomogeneous Universe, as well as a few novel elements, such as the distinction between lensing magnification and amplification and its interpretation. We shall focus on the statistical averages of distance measures, and how they relate to the area of light-cone slices.

2.1. Light propagation in a perturbed FLRW Universe

We consider a cosmological space-time described by a spatially flat FLRW model with scalar perturbations. The associated line

² In a more modern language, we may say that the calculation was made at first order in the micro-lensing optical depth τ , which coincides with the convergence κ if the density of the lenses were smoothed out; see, for instance, Sect. II.C of Fleury & Garc a-Bellido (2020).

element reads, in the Newtonian gauge

$$ds^2 = a^2(\eta) \left[-(1 + 2\phi)d\eta^2 + (1 - 2\phi)d\mathbf{x}^2 \right], \quad (1)$$

where η denotes the conformal time (hereafter simply referred to as time), x^i are comoving coordinates, $a(\eta)$ the scale factor describing cosmic expansion, and ϕ the Bardeen potential (Bardeen 1980) caused by inhomogeneities in the matter density field. We assume that anisotropic stress is negligible so that this potential is unique. Except in the vicinity of compact objects, $\phi \ll 1$ can be treated as a perturbation. The time at the observation event (here and now) is denoted η_0 , where the scale factor is conventionally set to unity, $a_0 = a(\eta_0) = 1$.

Light propagates along null geodesics of the space-time geometry. In the absence of perturbations (that is, for $\phi = 0$), such geodesics are straight lines in comoving coordinates, travelled with unit coordinate speed. In the presence of perturbations, light rays are bent and the coordinate speed of light effectively varies (Schneider et al. 1992). These effects are encoded in the null geodesic equation $k^\nu \nabla_\nu k^\mu = 0$, with $k^\mu \equiv dx^\mu/d\lambda$ and λ denotes a past-oriented affine parameter for the light ray. The temporal and spatial components of the geodesic equation read

$$\frac{dk^0}{d\lambda} = -2\mathcal{H}(k^0)^2 - 2\frac{d\phi}{d\lambda}k^0 + 2\frac{\partial\phi}{\partial\eta}(k^0)^2, \quad (2)$$

$$\frac{dk^i}{d\lambda} = -2\mathcal{H}k^0k^i + 2\frac{d\phi}{d\lambda}k^i - 2\frac{\partial\phi}{\partial x^i}(k^0)^2, \quad (3)$$

where $\mathcal{H} \equiv a^{-1}da/d\eta$ is the conformal expansion rate. Equation (2) rules the evolution of light's frequency in the cosmic frame; combined with the latter, Eq. (3) describes light bending.

2.2. Gravitational lensing

Light bending implies that the images of light sources are displaced and distorted when seen through the inhomogeneous Universe. Let θ denote the position of such an image of a point source, and β its FLRW counterpart, that is, the position where the image would be seen in the absence of cosmological perturbations. It is customary to refer to β as the source position.

2.2.1. Geometric distortions of infinitesimal images

The distortions of an infinitesimal image are then fully encoded in the Jacobi matrix of the mapping $\theta \mapsto \beta$, also called distortion matrix. This matrix may be parameterised as

$$\mathcal{A} \equiv \frac{\partial\beta}{\partial\theta} = \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix} \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}, \quad (4)$$

with κ , $\gamma = \gamma_1 + i\gamma_2$, and ω are respectively called the convergence, complex shear, and rotation. As a rule of thumb, κ , γ are typically first order in cosmological perturbations, while ω is second order (see for example Fleury 2015, Sect. 2.3.2).

We define the signed geometric magnification of an image as

$$\mu = \frac{1}{\det\mathcal{A}} = \frac{1}{(1 - \kappa)^2 - |\gamma|^2}. \quad (5)$$

By definition of the determinant of a matrix, its absolute value $|\mu| = d^2\theta/d^2\beta$ is the ratio of the angular size of an infinitesimal image, $d^2\theta$, and the angular size of the underlying source, $d^2\beta$.

As indicated by its name and definition, the signed magnification of an image can be either positive or negative, which indicates its orientation relative to the source. An image at θ is

said to have positive parity if $\mu(\theta) > 0$, and negative parity otherwise. In a Universe made of transparent lenses, any source has an odd total number $2n + 1$ of images, with $n \geq 0$ images of negative parity and $n + 1$ images of positive parity (Burke 1981; Schneider et al. 1992).

The total geometric magnification of a source β is the sum of the absolute magnifications of its $2n + 1$ images $\theta_i(\beta)$,

$$\mu_{\text{tot}}(\beta) \equiv \sum_{i=1}^{2n+1} |\mu[\theta_i(\beta)]|. \quad (6)$$

It represents the total increase in apparent size of a source relative to its unlensed counterpart.

2.2.2. Geometric-magnification integrals

In a transparent Universe, the map $\theta \mapsto \beta(\theta)$, which to an image associates its source, is a well-defined surjective function of \mathbb{S}^2 onto \mathbb{S}^2 . In other words, any image has one and only one source, and every source has at least one image. These properties imply

$$\int_{\mathbb{S}^2} d^2\theta \mu^{-1}(\theta) = 4\pi, \quad (7)$$

$$\int_{\mathbb{S}^2} d^2\beta \mu_{\text{tot}}(\beta) = 4\pi, \quad (8)$$

which we refer to as the geometric-magnification integrals.

We note that in the absence of multiple imaging, $\theta \mapsto \beta(\theta)$ is a diffeomorphism of \mathbb{S}^2 , so that Eqs. (7) and (8) are merely changes of variables in an integral. The true interest of the magnification integrals is that they hold even in the presence of strong lensing and multiple images.

The total magnification integral (8) is the full generalisation of the result found by Weinberg (1976) at linear order and with point lenses. To the best of our knowledge, it was first formulated by Wucknitz (2008). The proof goes as follows. For each source element $d^2\beta$, $d^2\theta_{\text{tot}} = \mu_{\text{tot}}(\beta) d^2\beta$ is the total solid angle occupied by the associated images. As one sums over $d^2\beta$, the image sphere gets progressively covered. On the one hand, the whole sphere is eventually covered, because any image has a source – for any θ , there is always a corresponding β . On the other hand, every image point θ is covered only once, because an image cannot have more than one source.

The inverse-magnification integral (7) can be found in Kibble & Lieu (2005). Its proof relies on the relative number of positive- and negative-parity images, mentioned in Sect. 2.2.1. For each element $d^2\theta$ of the image sphere, $d^2\beta = |\mu^{-1}(\theta)| d^2\theta$ is the corresponding solid angle in the source sphere. As one sums over $d^2\theta$, the entire source sphere is covered, again because every source has at least one image. Multiple imaging implies, however, that some regions of the source sphere may be covered several times. When this occurs, since a source β always has $2n + 1$ images $\theta_i(\beta)$, n of which having negative parity, their contributions cancel two by two but one,

$$\sum_{i=1}^{2n+1} \mu^{-1}[\theta_i(\beta)] d^2\theta_i = \sum_{i=1}^{2n+1} (-1)^i d^2\beta = d^2\beta. \quad (9)$$

Therefore, each source element is eventually covered once and only once, which leads to Eq. (7). As pointed out by KP16, albeit correct the inverse-magnification integral has little practical interest, because it is difficult to observe the parity of an image.

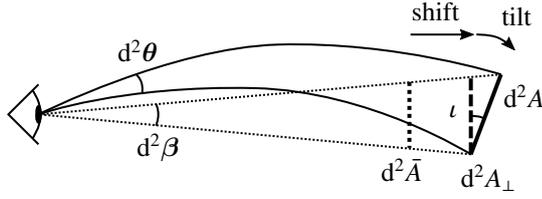


Fig. 1. Illustrating the difference between the geometrical magnification $\mu = d^2\theta/d^2\beta$ and the observable magnification $\tilde{\mu}(z) = [\bar{D}_A(z)/D_A(z)]^2 = (d^2\bar{A}/d^2A) \times (d^2\theta/d^2\beta)$. The causes of discrepancy are: (i) the position of a source at z may be shifted from the background to the perturbed case, $d^2\bar{A} \neq d^2A_\perp$; (ii) the beam's cross section may be tilted due to lensing, $d^2A_\perp = d^2A_z \cos \iota$. This illustration equally applies if z is replaced with λ, η, \dots .

2.2.3. Observable magnification: shift and tilt corrections

The geometric magnification $\mu = \pm d^2\theta/d^2\beta$ is a well-defined theoretical notion, but it is not the most observationally relevant one. This is because $d^2\beta$ represents the coordinate solid angle associated with an image, rather than the unlensed apparent size $d^2\theta$ of its source. There are two reasons why these quantities differ, namely the ‘shift’ and ‘tilt’, which we elaborate on below.

Observable magnification. We consider an infinitesimal source at redshift z with physical area d^2A . Let $d^2\theta$ be the apparent size of an image of that source, and $d^2\theta$ its unlensed counterpart, that is the solid angle under which d^2A would be seen at the same redshift in FLRW. The observable magnification is defined as

$$\tilde{\mu}(z, \theta) \equiv \pm \frac{d^2\theta}{d^2\theta}, \quad (10)$$

where the \pm sign indicates the image parity. By definition, the absolute observable magnification thus quantifies the change of the area distance D_A to an image due to cosmological perturbations,

$$|\tilde{\mu}(z, \theta)| = \left[\frac{\bar{D}_A(z)}{D_A(z, \theta)} \right]^2. \quad (11)$$

We note that the above relies on a notion of area distance associated with individual images θ .

Shift and tilt. We now relate the observational magnification $\tilde{\mu}$ to the geometric magnification μ . For simplicity, we identify the source with an infinitesimal patch of the surface of constant redshift. However, the results obtained in this paragraph are much more general; in particular, we refer the reader to Appendix A for an alternative approach based on a spherical source.

Let $d^2\beta$ be the coordinate solid angle covered by the source. We may multiply and divide the expression (11) of $|\tilde{\mu}|$ by $d^2\theta/d^2\beta$ to get

$$|\tilde{\mu}(z, \theta)| = \frac{d^2\theta}{d^2\beta} \times \frac{\bar{D}_A^2(z) d^2\beta}{D_A^2(z, \theta) d^2\theta} = |\mu(z, \theta)| \times \frac{d^2\bar{A}}{d^2A}, \quad (12)$$

where $d^2\bar{A} = \bar{D}_A^2(z) d^2\beta$ is the physical area sub-tended by the coordinate solid angle $d^2\beta$ in the absence of perturbations.

As illustrated in Fig. 1, $d^2\bar{A}$ differs from d^2A for two reasons. First, for a given redshift z , the time and radial position of the source event are not necessarily the same in the background ($\bar{\eta}, \bar{r}$) as in the perturbed Universe (η, r); the coordinates of that event are shifted. We call d^2A_\perp the area sub-tended by $d^2\beta$ at the

shifted event; we have

$$d^2A_\perp = a^2[\eta(z)] r^2(z) d^2\beta \neq a^2[\bar{\eta}(z)] \bar{r}^2(z) d^2\beta = d^2\bar{A}. \quad (13)$$

Second, because of light deflection, the orientation of the source is tilted by an angle ι with respect to how it would be seen in FLRW. Because they are sub-tended by the same solid angle $d^2\beta$, the tilted area d^2A is larger than its untilted counterpart $d^2A_\perp = d^2A \times \cos \iota$.

Summarising, the observable and geometrical magnifications are related as

$$\frac{\tilde{\mu}(z, \theta)}{\mu(z, \theta)} = \underbrace{\frac{d^2\bar{A}}{d^2A_\perp}}_{\text{shift}} \underbrace{\frac{d^2A_\perp}{d^2A}}_{\text{tilt}} = \frac{a^2[\bar{\eta}(z)] \bar{r}^2(z)}{a^2[\eta(z)] r^2(z)} \cos \iota. \quad (14)$$

We generally expect the shift to be the main driver of the difference between μ and $\tilde{\mu}$, because the tilt $\cos \iota \approx 1 - \iota^2/2$ is a second-order quantity. Specifically, in the numerical results discussed in Sect. 4.1, the effect of tilt will always be sub-dominant; it will be precisely quantified in Sect. 4.3.

Physical origin of the shift. While μ is a pure-lensing quantity, $\tilde{\mu}$ depends on other phenomena, such as time delays, Sachs-Wolfe (SW) and integrated Sachs-Wolfe (ISW) effects, or peculiar velocities. The latter in particular may lead to significant differences between $\mu(z)$ and $\tilde{\mu}(z)$ at low redshift. If a source has, for instance, a centripetal peculiar velocity with respect to the observer, then its redshift is smaller compared to a comoving source at the same position. Thus, for a given redshift z its comoving distance must be slightly larger than the one that it would have if it were comoving, $r(z) > \bar{r}(z)$. Because the source event belongs to the observer's past light cone, this also means that it happens slightly earlier, $\eta(z) < \bar{\eta}(z)$. At low z , this typically results in $d^2A_\perp > d^2\bar{A}$, implying that $\tilde{\mu}(z) < \mu(z)$. The conclusion would be opposite if the peculiar velocity were centrifugal.

To be more specific, at first order in the peculiar velocities of the source, \mathbf{v}_s and of the observer, \mathbf{v}_o , the shift³ reads (Kaiser 1987; Sasaki 1987)

$$\left(\frac{d^2A_\perp}{d^2\bar{A}} \right)_v \equiv 1 + 2\tilde{\kappa}_v, \quad \tilde{\kappa}_v \equiv \left(\frac{1}{\mathcal{H}r} - 1 \right) (\mathbf{v}_o - \mathbf{v}_s) \cdot \boldsymbol{\beta} + \mathbf{v}_o \cdot \boldsymbol{\beta}, \quad (15)$$

where $\boldsymbol{\beta}$ is the unit vector in the background direction of the source. The $1/(\mathcal{H}r)$ term in Eq. (15) shows that for sources at small distances, $\tilde{\kappa}_v$ may reach large values.

The quantity $\tilde{\kappa}_v$ is sometimes called ‘Doppler convergence’ (Bonvin 2008; Bolejko et al. 2013; Bacon et al. 2014), although it is unrelated to lensing. This expression and notation originate from the fact that we may define an observable distortion matrix $\tilde{\mathcal{A}}$, which is to the distortion matrix \mathcal{A} what $\tilde{\mu}$ is to μ . Namely, if

$$\tilde{\mathcal{A}}(z, \theta) \equiv \bar{D}_A^{-1}(z) \mathcal{D}(z, \theta), \quad (16)$$

where \mathcal{D} is the Jacobi matrix of the Sachs formalism (for example Fleury 2015, Sect. 2.2), then $\tilde{\mu} = 1/\det \tilde{\mathcal{A}}$. We may then introduce a convergence-shear decomposition of $\tilde{\mathcal{A}}$ similarly to Eq. (4), thereby defining $\tilde{\kappa}$, to which $\tilde{\kappa}_v$ is an important contribution.

³ In fact, Eq. (15) not only allows for the shift of the iso- z surface, but also for the optical aberration effect due to the observer's velocity. If the observer moves towards the source ($\mathbf{v}_o \cdot \hat{\mathbf{r}} > 0$), then the source appears smaller to them, $d^2\theta_\perp < d^2\theta_z$.

Fixing other parameters. In the above, we have defined the observable magnification $\tilde{\mu}$ at fixed redshift. This choice was made for concreteness, but the definition of $\tilde{\mu}$ could be adapted if we were to fix another parameter, such as the comoving radius, the emission time, or the affine parameter. A little intellectual challenge would consist in determining which light-cone slicing may ensure $\tilde{\mu} = \mu$. To our knowledge, there is currently no answer to that particular question.

2.2.4. Amplification and luminosity distance

Small or remote sources, such as SNe or quasars, are generally unresolved by telescopes. In that context, the key observable is the observed flux, that is the total power received from the source per unit of telescope area, rather than the apparent size of images. We may define the amplification of a source β at z as the ratio of the observed flux $S(z, \beta)$ with its unlensed counterpart $\bar{S}(z)$. By virtue of Etherington's reciprocity law (Etherington 1933), and assuming a transparent Universe, the amplification is nothing but the total observable magnification,

$$\frac{S(z, \beta)}{\bar{S}(z)} = \tilde{\mu}_{\text{tot}}(z, \beta) \equiv \sum_{i=1}^{2n+1} |\tilde{\mu}[z, \theta_i(\beta)]|. \quad (17)$$

By definition of the luminosity distance D_L , we also have

$$\tilde{\mu}_{\text{tot}}(z, \beta) = \left[\frac{\bar{D}_L(z)}{D_L(z, \beta)} \right]^2. \quad (18)$$

This could seem to be at odds with Eq. (11) and the well-known distance-duality relation $D_L = (1+z)^2 D_A$. This apparent paradox is due to the fact that we have defined D_A for a single image, while D_L accounts for all the images of a given source. The two approaches are reconciled if we consistently distinguish between image-based definitions and source-based definitions. For example, we could define the area distance of a multiply imaged source $D_A(z, \beta)$ from the total apparent area occupied by all its images. In that case $[\bar{D}_A(z)/D_A(z, \beta)]^2 = \tilde{\mu}_{\text{tot}}(z, \beta)$ consistently with distance duality.

Finally, we note that Eq. (17) is only valid if one compares the background and perturbed fluxes at the same redshift z . Had we compared the two situations, for instance, at fixed affine parameter, the background and perturbed redshift would have differed, which would have affected fluxes through the energy and the reception rate of individual photons.

2.3. Averaging in cosmology

The interpretation of cosmological observations, and their confrontation with theoretical predictions, involve various notions of averaging, which are non-trivially related in the presence of gravitational lensing. We review here the relevant definitions and properties of cosmological averages, elaborating on Bonvin et al. (2015b), Kaiser & Peacock (2016), Fleury et al. (2017a).

Importantly, from now on we shall neglect multiple imaging, except explicitly stated otherwise. Thus, the lens mapping $\theta \mapsto \beta$ is assumed to be a diffeomorphism of \mathbb{S}^2 , and the resulting magnifications are positive. In that context, there is no difference between signed, absolute, and total magnifications any more. We may also treat observable magnification and amplification as synonyms, both denoted $\tilde{\mu}$. This assumption is justified by the relatively rare occurrence of strong lensing from a cosmological perspective, and by the huge gain of simplicity that it brings to the discussions of this section.

2.3.1. Directional averaging

Let $X(\theta)$ be an observable in the direction θ on the observer's celestial sphere, such as the temperature anisotropies of the cosmic microwave background, or the apparent surface density of galaxies. Directional averaging $\langle \dots \rangle_d$ corresponds to a statistical average of $X(\theta)$ where all the observation directions θ have the same statistical weight; the average is thus weighted by the image solid angle $d^2\theta$,

$$\langle X \rangle_d \equiv \frac{1}{4\pi} \int_{\mathbb{S}^2} d^2\theta X(\theta). \quad (19)$$

One may ask how lensing affects directional averages, in particular for distance measurements. We first note that, by virtue of Eq. (7), the directional average of the inverse geometric magnification is unity,

$$\langle \mu^{-1} \rangle_d \equiv \frac{1}{4\pi} \int_{\mathbb{S}^2} d^2\theta \mu^{-1}(\theta) = 1. \quad (20)$$

This property is exact and applies to any slicing of the light-cone. However, as pointed out in Sect. 2.2.3, Eq. (20) has only little observational relevance, because the actually observable quantity is $\tilde{\mu}$, which differ from μ by the shift and tilt described in Fig. 1. Despite that concern, we may still conclude that

$$\langle \tilde{\mu}^{-1} \rangle_d \approx 1, \quad (21)$$

in the limit where the tilt/shift corrections are sub-dominant compared to the most relevant gravitational-lensing effects.

Unlike Eq. (20), the accuracy of Eq. (21) depends on which parameter is fixed in the definition of $\tilde{\mu}$. For instance, Kibble & Lieu (2005) argued that $\langle \tilde{\mu}^{-1}(\lambda) \rangle_d = 1$ was accurate for sources at fixed affine parameter λ ; this was checked numerically with ray tracing in post-Newtonian cosmological modelling (Sanghai et al. 2017). However, we shall see in Sect. 4.2.5 that the use of the affine parameter is quite risky at very high redshift. If instead the redshift is kept fixed, then significant departures from $\langle \tilde{\mu}^{-1}(z) \rangle_d = 1$ are expected at low z due to peculiar velocities.

2.3.2. Source-averaging and areal averaging

We now consider an observable Y which is associated with a specific population of sources, such as the distance to type-Ia supernovae or the Lyman- α absorption in quasar spectra. The natural averaging procedure associated with such an observable is called source averaging $\langle \dots \rangle_s$, and is defined as

$$\langle Y \rangle_s = \frac{1}{N} \sum_{s=1}^N Y_s = \int_{\mathbb{S}^2} d^2\theta \frac{1}{N} \frac{d^2N}{d^2\theta} Y(\theta), \quad (22)$$

where N denotes the number of observed sources, and in the second equality we took the continuous limit. The difference with directional averaging is that the sky is not necessarily homogeneously sampled. Clearly, if the sources are not homogeneously distributed in the Universe, then their projected density $N^{-1} d^2N/d^2\theta$ tends to favour some regions of the sky more than others, thereby breaking the apparent statistical isotropy.

But even if sources are homogeneously distributed in space, gravitational lensing implies that they do not evenly sample the observer's sky. Indeed, lensing tends to make light beams 'avoid' over-dense regions of the Universe, thereby favouring under-dense regions in source-averages. This specific effect may be

captured in the notion of areal averaging. For example, if all the sources are observed at the same redshift z , we may define the areal average of Y as

$$\langle Y(z) \rangle_a \equiv \frac{1}{A(z)} \int_{\Sigma(z)} d^2 A_z Y(\mathbf{x}), \quad (23)$$

with $\Sigma(z)$ the surface of constant redshift z and $A(z)$ its total proper area. The definition must be adapted if the sources are observed on other slices of the light cone, for instance all at the same emission time η or affine parameter λ .

Using the area distance, $d^2 A_z = D_A^2(z) d^2 \theta$, we may convert areal averages in terms of directional averages as follows,

$$\langle Y(z) \rangle_a = \frac{\int_{\mathbb{S}^2} d^2 \theta D_A^2(z, \theta) Y(z, \theta)}{\int_{\mathbb{S}^2} d^2 \theta D_A^2(z, \theta)} = \frac{\langle \tilde{\mu}^{-1}(z) Y(z) \rangle_d}{\langle \tilde{\mu}^{-1}(z) \rangle_d}, \quad (24)$$

from which we immediately conclude, substituting $Y = \tilde{\mu}$, that

$$\langle \tilde{\mu}(z) \rangle_a = \frac{1}{\langle \tilde{\mu}^{-1}(z) \rangle_d} \approx 1, \quad (25)$$

by virtue of Eq. (21). Areal averaging exactly coincides with source-averaging if the sources are homogeneously distributed on $\Sigma(z)$, because then the number of observed sources scales as the area that they occupy, so that $N^{-1} d^2 N / d^2 \theta = d^2 A_z / d^2 \theta = D_A^2(z)$. If not, corrections arise from the correlation between the fluctuations of the density of sources and the amplification; further corrections such as redshift-space distortions, must also be accounted for if the sources are observed in redshift bins (Fleury et al. 2017a; Fanizza et al. 2020). Such discrepancies between source-averaging and areal averaging typically remain below 10^{-5} , and hence they may be neglected. Combining this approximation with Eq. (25) then yields

$$\langle \tilde{\mu}(z) \rangle_s \approx 1. \quad (26)$$

Equation (26) was shown to be accurate at the 10^{-3} level up to $z = 3$ by Adamek et al. (2019).

2.3.3. Ensemble averaging and cosmic variance

We shall close this discussion with the notion of ensemble averaging. Within the standard lore, we envisage all cosmological structures as originating from quantum fluctuations in the primordial Universe (Peter & Uzan 2013). From that point of view, $\phi(\eta, \mathbf{x})$ is a particular realisation of an intrinsically stochastic field, which is believed to be initially Gaussian. In that framework, the ensemble average of any field $Z(\eta, \mathbf{x})$ that depends on ϕ , which we may simply denote as $\langle Z(\eta, \mathbf{x}) \rangle$, would be its expectation value over an infinite number of realisations the Universe.

Contrary to directional, areal, or source-averages, ensemble-averaging is thus a strictly theoretical procedure, which is nevertheless used in any cosmological prediction. Ensemble averaging may be connected to other averaging procedures via the ergodicity principle. Which observable averaging is mimicked by ensemble averaging then depends on which quantities are kept fixed when making multiple realisations of the Universe as illustrated in Fig. 2. For example, if the observed direction of light θ and redshift z are kept fixed, then we get directional averaging,

$$\langle X(z, \theta) \rangle = \langle X(z) \rangle_d, \quad (27)$$

because any θ is virtually affected the same statistical weight. In this scenario, the source position β may change from one cosmic

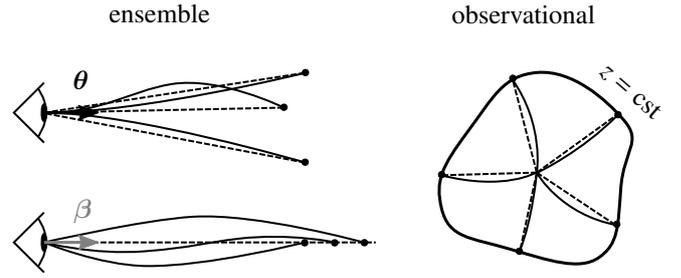


Fig. 2. Correspondence between ensemble averaging and observational averaging procedure depends on which quantity is kept fixed.

realisation to another. An alternative scenario would consist, on the contrary, in fixing β while allowing θ to vary from one realisation to another; this yields

$$\langle X(z, \beta) \rangle = \langle \mu^{-1}(z) X(z) \rangle_d. \quad (28)$$

We may divide the above with $\langle \mu^{-1} \rangle_d$ if directional average is taken on a fraction of the sky only. Other possibilities would consist in fixing another parameter than the redshift, such as time or affine parameter, which would correspond to averaging across other slices of the light cone.

Importantly, ergodicity is sensible only if the region of the Universe over which an observational averaging is performed is statistically homogeneous. In other words, there should not be super-sample inhomogeneity modes. Such an assumption is not satisfied in the standard lore, which predicts inhomogeneity modes at all scales. Thus, any observational average is subject to an irreducible source of uncertainty, called cosmic variance. Equations (27) and (28) only hold up to cosmic variance.

2.4. Biased distance measurements

Equations (21) and (26) show that only very specific quantities are (almost) unbiased by cosmic inhomogeneities; in particular, most distance measurements happen to be biased. We describe here the nature and amplitude of these biases.

We introduce for convenience the dimension-less distance

$$d(z, \theta) \equiv \frac{D_A(z, \theta)}{\bar{D}_A(z)} = |\tilde{\mu}(z, \theta)|^{-1/2}. \quad (29)$$

Because d is a non-linear function of $\tilde{\mu}$, it exhibits a statistical bias for both directional and source-averaging. For directional averaging we may expand d at second order in $\tilde{\mu}^{-1} - 1$ and use Eq. (21) to get

$$\langle d(z) \rangle_d - 1 \approx -\frac{1}{8} \langle [\tilde{\mu}^{-1}(z) - 1]^2 \rangle_d < 0. \quad (30)$$

Similarly, for areal or source-averaging we may expand d in terms of $\tilde{\mu} - 1$ which, together with Eq. (26) yields

$$\langle d(z) \rangle_s - 1 \approx \frac{3}{8} \langle [\tilde{\mu}(z) - 1]^2 \rangle_s > 0. \quad (31)$$

The biases appearing in Eqs. (30) and (31) are not independent, and are usually expressed in terms of the convergence. Indeed, if the amplification is expressed in terms of some convergence and shear similarly to Eq. (5), that is, $\tilde{\mu} = [(1 - \tilde{\kappa})^2 - |\tilde{\gamma}|^2]^{-1}$, then at second order in $\tilde{\kappa}, \tilde{\gamma}$,

$$(\tilde{\mu} - 1)^2 = (\tilde{\mu}^{-1} - 1)^2 = 4\tilde{\kappa}^2. \quad (32)$$

Since $\tilde{\kappa}^2$ is a second-order quantity, the difference between its directional, source-, or ensemble-average would be of higher order, and hence it should not matter much which averaging procedure is considered when substituting $\tilde{\kappa}^2$ in Eqs. (30) and (31). Furthermore, if we neglect again the shift and tilt corrections⁴ and write $\tilde{\kappa} = \kappa$, then we simply have

$$\langle d(z) \rangle_d \approx 1 - \frac{1}{2} \langle \kappa^2(z) \rangle, \quad (33)$$

$$\langle d(z) \rangle_s \approx 1 + \frac{3}{2} \langle \kappa^2(z) \rangle. \quad (34)$$

If Eq. (34) is applied at the redshift of the CMB, $z_* \approx 1100$, and $\kappa(z)$ is computed from linear perturbation theory, then the corresponding bias reaches the percent level. This is how Clarkson et al. (2014) concluded that the standard analysis of the CMB, which does not account for such a bias, might be flawed. That conclusion was shown to be incorrect by Bonvin et al. (2015a), Kaiser & Peacock (2016), because the analysis of the CMB is in fact not sensitive to $\langle d(z_*) \rangle_s$. However, supernova cosmology is. In supernova surveys, it is customary to use the distance modulus rather than the luminosity distance as a distance measure; its perturbation due to inhomogeneities reads $\Delta m = 5 \log_{10} d$, and hence its source-averaged bias is

$$\langle \Delta m(z) \rangle_s \approx \frac{5}{4 \ln 10} \langle [\tilde{\mu}(z) - 1]^2 \rangle_s \approx \frac{5}{\ln 10} \langle \kappa^2(z) \rangle. \quad (35)$$

For $z < 2$, this bias remains below 10^{-3} and hence is negligible in current SN surveys, except for reconstructions of the evolution of the dark-energy equation of state (Fleury et al. 2017a). It would be easily removed if next-generation surveys were using $1/D_L^2(z)$ instead of magnitude as a distance indicator.

We finally note that all the above only holds in a transparent Universe. If this assumption is relaxed, then distance measurements may be further biased by selection effects. For instance, in a Universe made of opaque matter lumps, observed light beams do not evenly sample the density field – they experience an effectively under-dense Universe. This result in an effective de-focussing of light as originally described by Zel'dovich (1964), later generalised by Dashevskii & Slysh (1966) and Dyer & Roeder (1974) on the basis of Einstein-Straus Swiss-cheese models (Kantowski 1969; Fleury 2014). The resulting bias on luminosity distance measurements typically reaches 10% at $z = 1$ for very lumpy models (Fleury et al. 2013). In Okamura & Futamase (2009), the authors made an attempt to determine the fraction of such opaque lumps based on the halo model of Sheth & Tormen (1999). To date, however, there is no compelling evidence of any large effect of opaque lumps on distance measurements in our Universe (Helbig 2020).

2.5. Reformulation: the area of light-cone slices

We may now rephrase the average-amplification rules (21) and (26) in terms of the area of light-cone slices, such as surfaces of constant redshift. We consider for instance the directional average of the inverse amplification:

$$\langle \tilde{\mu}^{-1}(z) \rangle_d = \frac{1}{4\pi \bar{D}_A^2(z)} \int_{\mathbb{S}^2} d^2\theta D_A^2(z, \theta) = \frac{A(z)}{\bar{A}(z)}, \quad (36)$$

where in the last equality we introduced the area $A(z)$ of the surface of constant redshift, $\Sigma(z)$, as well as its background counterpart $\bar{A}(z) = 4\pi \bar{D}_A^2(z)$. Equation (36) thus tells us that $\langle \tilde{\mu} \rangle_d \approx 1$

⁴ This approximation fails at low- z , where peculiar velocities generate Malmquist bias (Ben-Dayan et al. 2014; Kaiser & Hudson 2015).

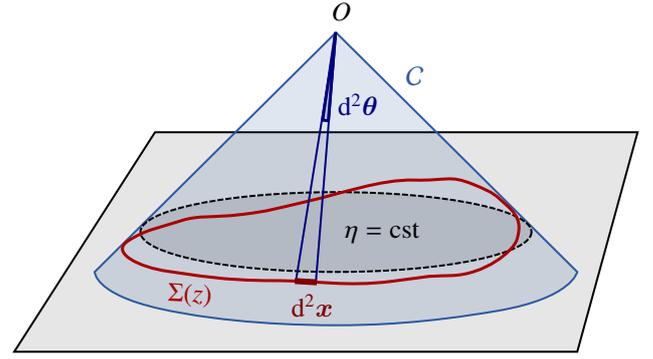


Fig. 3. Surface of constant redshift, $\Sigma(z)$ (red), is a particular slicing of the light cone C (blue) that is not included in constant-time hypersurfaces (grey). We represented light rays as straight lines for simplicity.

would be equivalent to $A(z) \approx \bar{A}(z)$; meaning that the area of iso- z surfaces is mostly unaffected by inhomogeneities.

Although it may seem quite natural, the last equality of Eq. (36) is, in fact, not obvious. In the background FLRW space-time, $\bar{\Sigma}(z)$ is a sphere (in comoving coordinates) at constant cosmic time; hence its proper area is clearly $\bar{A}(z) = 4\pi \bar{r}^2(z)/(1+z) = 4\pi \bar{D}_A^2(z)$. But things are less clear in the inhomogeneous Universe, where $\Sigma(z)$ is wrinkly and is not limited to a constant-time hypersurface. The definition of its proper area $A(z)$ is then subject to several questions about its uniqueness, if it is frame-dependent and how it relates to the angular distance. We propose to clarify these subtleties below.

2.5.1. Surfaces of constant redshift and their area

Surfaces of constant redshift, $\Sigma(z)$, correspond to a particular slicing of the light cone C of the observation event O , as illustrated in Fig. 3. We note that this slicing is generally not performed at constant time, $\eta(z) \neq \text{const}$. This last property raises the question of how to actually define the proper area of $\Sigma(z)$.

Let $d^2\mathbf{x}$ be the element of $\Sigma(z)$ subtended by the solid angle $d^2\theta$ at O . For causality reasons, $d^2\mathbf{x}$ must be space-like; thus, there exists a frame such that $d^2\mathbf{x}$ is strictly spatial. We shall call it the ‘natural frame’⁵ of $d^2\mathbf{x}$, and define the area d^2A_z of $d^2\mathbf{x}$ in that frame. Applying that construction to all elements $d^2\mathbf{x}$ of $\Sigma(z)$ and integrating over them then defines its total area $A(z)$.

Now that we have defined the area of an iso- z surface, we shall see how it relates to the angular distance $D_A(z)$. For that purpose, we note that $d^2\mathbf{x}$ is orthogonal to direction of light propagation in its natural frame. We shall now prove this point. We may see C as the hyper-surface defined by all the events that are in phase with O , for a spherical wave converging at the observer. If w denotes the phase of that wave and $k_\mu = \partial_\mu w$ is the associated wave four-vector, then any displacement $d\mathbf{x}^\mu$ across C satisfies $k_\mu d\mathbf{x}^\mu = dw = 0$. This applies, in particular, to any $d\mathbf{x}^\mu \in d^2\mathbf{x} \subset \Sigma(z) \subset C$. In the natural frame of $d^2\mathbf{x}$, this four-dimensional orthogonality becomes three-dimensional because $dx^0 = 0$; in other words, $\mathbf{k} \cdot d\mathbf{x} = 0$, where \mathbf{k} is the wave-vector in the natural frame.

The spatial orthogonality between \mathbf{k} and $d^2\mathbf{x}$ implies that $d^2\mathbf{x}$ forms a Sachs screen space in its natural frame. Thus, d^2A_z

⁵ The natural frame is not unique; there is in fact a class of natural frames which are all related by Lorentz boosts in the local direction of light propagation. The area of $d^2\mathbf{x}$ is invariant under such boosts, as long as they go from one natural frame to another one.

is not only the proper area of $d^2\mathbf{x}$, but also the cross-sectional area of the light beam subtended by $d^2\theta$ in that frame. By virtue of Sachs' shadow theorem (Sachs 1961, see also Sect. 2.1.2 of Fleury 2015), the area of a beam is independent of the frame in which it is evaluated, as long as it is projected on a Sachs screen. This unique notion of a beam's cross-sectional area then defines the angular distance according to

$$d^2A_z = D_A^2(z) d^2\theta. \quad (37)$$

This confirms that the area of $\Sigma(z)$ is indeed related to the angular distance as $A(z) = 4\pi\langle D_A^2(z) \rangle_d$, thereby validating the last equality of Eq. (36).

We finally note that the above reasoning actually applies to any slice of the light cone. In other words, for any parameter p such that the iso- p surface $\Sigma(p)$ is space-like (p may stand, for instance, for the affine parameter λ , time η , the comoving radius r , etc.) the area of the element $d^2\mathbf{x}$ subtended by the solid angle $d^2\theta$ reads $d^2A_p = D_A^2(p)d^2\theta$ and the total area of the iso- p surface is

$$A(p) = 4\pi\langle D_A^2(p) \rangle_d = \langle \tilde{\mu}^{-1}(p) \rangle_d \bar{A}(p). \quad (38)$$

2.5.2. The photon-flux conservation argument

In the spirit of the second part of Weinberg (1976), we may also connect the area-averaged amplification to $A(z)$ on the basis of photon conservation. Let F_0 be the total number of photons received per unit time by an observer. If the photon number is conserved, then the same photons crossed $\Sigma(z)$ at a rate $F_z = (1+z)F_0$, where the $(1+z)$ factor accounts for time dilation. Importantly, the latter relation holds regardless of the geometry of $\Sigma(z)$; in other words, $F_z = \bar{F}_z$.

Now, the photon flux may be written as

$$F_z = \int_{\Sigma(z)} d^2\mathbf{x} \mathbf{n} \cdot \mathbf{J}(z) = A(z)\langle |\mathbf{J}(z)| \rangle_a, \quad (39)$$

where \mathbf{n} is the outgoing normal to $\Sigma(z)$, \mathbf{J} is the photon flux density vector, and in the second equality we used that in its natural frame \mathbf{n} is aligned with \mathbf{k} and hence to \mathbf{J} . Since $\mathbf{J} = \mathbf{\Pi}/(\hbar\omega)$, where $\mathbf{\Pi} \propto 1/D_L^2(z)$ is the Poynting vector, we conclude that $|\mathbf{J}(z)| \propto \tilde{\mu}(z)$. Combining this with $F_z = \bar{F}_z$ then yields

$$\langle \tilde{\mu}(z) \rangle_a = \frac{\bar{A}(z)}{A(z)}. \quad (40)$$

Therefore, Eq. (25) is equivalent to stating that inhomogeneities do not affect the area of iso- z surfaces. We stress again here that Weinberg's photon-flux-conservation argument does not imply that $\langle \tilde{\mu}(z) \rangle_a = 1$, but rather shows that such an equality is equivalent to $A(z) = \bar{A}(z)$.

2.5.3. CMB and the area of the last-scattering surface

Hitherto, our discussion has been focused on surfaces of constant redshift because of their connection with observational averages. The archetypal application would be the analysis of the Hubble diagram in a non-homogeneous Universe, which involves the source-averaged distance modulus. However, shall we be more interested in the analysis of the CMB than in SNe, more relevant would be the last-scattering surface (LSS) and its area A_* .

The area of the LSS is a relevant quantity indeed. As illustrated in Fig. 4, A_* essentially drives how many sound horizons r_s the observer may count in the CMB, and hence their average

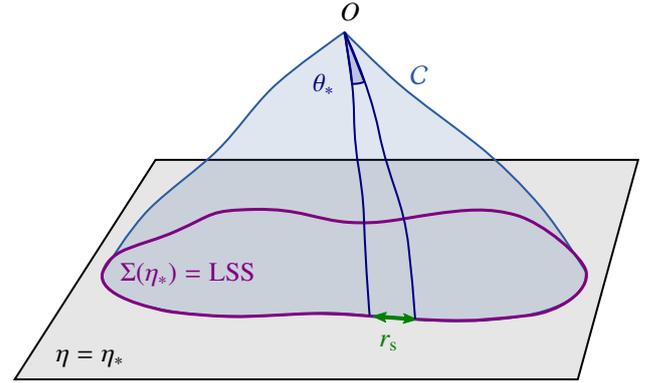


Fig. 4. Last scattering surface, approximated as a constant-time slice $\Sigma(\eta_*)$ of the light cone C . Its area is connected to the number of sound horizons r_s that appear on the observer's CMB, and hence to its average apparent size θ_* .

angular size θ_* which is one of the main direct CMB observables. The problem is then to estimate to which extent is A_* affected by the inhomogeneities of our Universe. Following KP16, we shall approximate the LSS as a surface of constant time⁶, $\Sigma(\eta_*)$. By virtue of Eq. (38) for $p = \eta_*$, we have $A_* \approx \bar{A}_*$ where the departure from strict equality stems from the difference between $\tilde{\mu}(\eta_*)$ and $\mu(\eta_*)$, that is from the small shift and tilt effects emphasised in Sect. 2.2.3.

From a pedagogical point of view, surfaces of constant time $\Sigma(\eta)$ are very attractive because their natural frame (in the sense of Sect. 2.5.1) is the comoving frame. Indeed, for any displacement $dx^\mu \in \Sigma(\eta)$ we have by definition $dx^0 = 0$ in comoving coordinates. Because of this, the shift and tilt corrections to the area A_* of the LSS can be made particularly explicit. Expressing A_* as the area of a polar surface, we have indeed

$$A(\eta_*) = a^2(\eta_*) \int_{\mathbb{S}^2} d^2\boldsymbol{\beta} \frac{r_*^2(\boldsymbol{\beta})}{\cos \iota_*(\boldsymbol{\beta})}, \quad (41)$$

$$\cos \iota_* = \left(1 + \left| \frac{1}{r_*} \frac{\partial r_*}{\partial \boldsymbol{\beta}} \right|^2 \right)^{-1/2}, \quad (42)$$

where $r_*(\boldsymbol{\beta}) \equiv r(\eta_*, \boldsymbol{\beta}) \neq \bar{r}(\eta_*)$ is the comoving radial coordinate of the point of LSS with angular coordinates $\boldsymbol{\beta}$, which is generally shifted with respect to its background counterpart $\bar{r}_*(\boldsymbol{\beta})$. The angle $\iota_*(\boldsymbol{\beta})$ is the tilt between the normal to the LSS and the radial direction; it encodes the wrinkles of the LSS which tend to increase its area. Recall that $\boldsymbol{\beta}$ denotes the 'true' angular position of a point of the LSS, not to be confused with the direction θ in which that point would be observed.

In order to get a theoretical prediction for $\delta A_*/\bar{A}_*$, we may expand Eq. (41) at second order in cosmological perturbations, and assume ergodicity to turn integrations over $\boldsymbol{\beta}$ into ensemble averages (see Sect. 2.3.3). This yields

$$\frac{\delta A_*}{\bar{A}_*} = \underbrace{\left\langle \frac{2\delta r_*(\boldsymbol{\beta})}{\bar{r}_*} + \frac{\delta r_*^2(\boldsymbol{\beta})}{\bar{r}_*^2} \right\rangle}_{\text{shift}} + \underbrace{\frac{1}{2} \langle \iota_*^2(\boldsymbol{\beta}) \rangle}_{\text{tilt}}, \quad (43)$$

with $\delta A_* \equiv A_* - \bar{A}_*$.

KP16 proposed a quite intuitive analysis of the shift term, δr_* ; we shall paraphrase their idea here, while further details and

⁶ This is of course a gauge-dependent statement, see for example Ellis & Durrer (2018) for a discussion.

minor corrections are given in Appendix C. The first effect of inhomogeneities is the presence of the gravitational potential ϕ , which changes the effective (coordinate) speed of light as

$$c_{\text{eff}} = \left| \frac{d\mathbf{x}}{d\eta} \right| = \sqrt{\frac{1+2\phi}{1-2\phi}} = 1 + 2\phi + \mathcal{O}(\phi^2). \quad (44)$$

As a consequence, during a fixed travel time $\eta_0 - \eta$, the comoving distance travelled⁷ s is slightly changed, $s(\eta) = \eta_0 - \eta + \delta s(\eta)$. At the LSS, this reads

$$\delta s_*(\boldsymbol{\beta}) = \int_{\eta_*}^{\eta_0} d\eta \, 2\phi[\eta, \mathbf{x}(\eta)], \quad (45)$$

where $\mathbf{x}(\eta)$ is the photon trajectory connecting the observer to the point $\boldsymbol{\beta}$ of the LSS. We note that this is nothing but the usual Shapiro time delay seen from a different point of view.

Second, because of gravitational lensing, light rays are wiggly, and hence the comoving radius r that they reach after travelling a comoving distance s is slightly smaller than s . At the LSS we may write $r_* = s_* + \delta r_{\text{geo}}$, with

$$\delta r_{\text{geo}}(\boldsymbol{\beta}) = \int_0^{s_*} ds [\cos \iota(s) - 1] \approx -\frac{1}{2} \int_0^{\bar{r}_*} dr \, \iota^2(r), \quad (46)$$

at second order in perturbations, and where ι is the angle made between the instantaneous photon propagation direction and the axis spanned by $\boldsymbol{\beta}$. It coincides with ι_* at the LSS.

When both effects (time delay and wiggles) are taken into account, the radial shift of the LSS with respect to its background counterpart reads

$$\delta r_*(\boldsymbol{\beta}) \equiv r_*(\boldsymbol{\beta}) - \bar{r}_* = \delta s_*(\boldsymbol{\beta}) + \delta r_{\text{geo}}(\boldsymbol{\beta}). \quad (47)$$

We note that δs_* is first-order, while δr_{geo} is second-order in cosmological perturbations. It is thus essential to go beyond the Born approximation when evaluating δs_* for consistency. Because of that hierarchy, δs_* may also be considered the main driver of the wrinkles ι_* of the LSS.

Once ensemble average is taken, Eq. (43) yields

$$\frac{\delta A_*}{\bar{A}_*} = \int_0^{\bar{r}_*} dr \frac{(2\bar{r}_* - r)r}{\bar{r}_*^2} J(r) \approx 5 \times 10^{-7}, \quad (48)$$

$$J(r) \equiv 2 \int_0^\infty \frac{dk}{2\pi} k^3 P_\phi(\eta_0 - r, k), \quad (49)$$

where P_ϕ denotes the power spectrum of the gravitational potential. More details can be found in Appendix C. Equation (48) agrees with Eq. (A.44) of KP16, albeit obtained via a slightly different path.

2.6. Summary and goal of the remainder of this article

Inhomogeneities may bias cosmological observations, notably via the effect of gravitational lensing on distance measurements. Biases depend on the notion of averaging that is involved. By virtue of the inverse-magnification integral $\langle \mu^{-1} \rangle_{\text{d}} = 1$, some specific observables are expected to be almost unbiased: $\langle d^2(z) \rangle_{\text{d}} \approx 1 / \langle d^{-2}(z) \rangle_{\text{s}} \approx 1$. Other combinations of d , such as the magnitude, generally exhibit a potentially much larger bias, on the order of $\langle \kappa^2 \rangle$. Departures from the exact $\langle d^2(z) \rangle_{\text{d}} = 1$ stem from $\tilde{\mu} \neq \mu$ and may be interpreted as being due to shifts and tilts

⁷ In KP16 that same quantity is denoted λ . We adopt s instead in order to avoid confusions with the affine parameter.

of iso- z surfaces with respect to their background counterpart. Apart from these shift and tilt effects, the area of iso- z surfaces is unaffected by inhomogeneities. An equivalent reasoning may be applied to other slices of the light cone, such as surfaces of constant time whose area is relevant for CMB observations.

In the remainder of this article, we propose to numerically evaluate: (i) the accuracy of $\langle d^2(z) \rangle_{\text{d}} \approx 1 / \langle d^{-2}(z) \rangle_{\text{s}} \approx 1$; (ii) the amplitude of the $\mathcal{O}(\langle \kappa^2 \rangle)$ bias on other observables; (iii) the performance of the prediction (48) for the area of iso- η surfaces. Our investigation will be based on accurate ray tracing in a high-resolution N -body simulation, so as to fully capture non-linear effects which are difficult to control in a pure-theory approach.

3. Numerical methods

In this section, we present the numerical set-up and the various tools that are used to obtain the results reported in Sect. 4.

3.1. Simulation

We use the N -body code RAMSES (Teyssier 2002; Guillet & Teyssier 2011) with dark matter (DM) only. RAMSES uses a Particle-Mesh with Adaptive-Mesh-Refinement (PM-AMR) method, which computes the evolution of the gravitational potential and density field from particles and gravity cells. AMR allows one to probe high-density regions and hence the highly non-linear regime of structure formation.

The simulation's box comoving length is $2625 h^{-1}$ Mpc with 4096^3 particles in a Λ CDM cosmology with WMAP-7 best-fit parameters (Komatsu et al. 2011), namely $h = 0.72$, total-matter density $\Omega_{\text{m}} = 0.25733$, baryon density $\Omega_{\text{b}} = 0.04356$, radiation density $\Omega_{\text{r}} = 8.076 \times 10^{-5}$, spectral index $n_{\text{s}} = 0.963$ and power-spectrum normalisation $\sigma_8 = 0.801$. The corresponding DM-particle mass is $1.88 \times 10^{10} h^{-1} M_{\odot}$. The initial power spectrum is computed with CAMB (Lewis et al. 2000). Initial conditions are generated using a 2LPT version of MPMGRAFIC (Prunet et al. 2008) to avoid transients (Scocimarro 1998), which allows us to start the simulation at $z = 46$.

Fidler et al. (2015, 2016) showed that Newtonian N -body simulations, such as the one used in this article, yield physical quantities computed in the so-called N -body gauge. In principle, a small relativistic correction must be applied to translate such results into the Newtonian gauge (Chisari & Zaldarriaga 2011). We choose to neglect these corrections, and hence we identify the coordinates and the gravitational potential computed from the simulation with the coordinates and metric perturbation ϕ in Eq. (1).

3.2. Light cones

To produce light cones from our simulation we use the onion-shell method (Fosalba et al. 2008; Teyssier et al. 2009). At each synchronisation (coarse) time step of the simulation, we output a thin spherical shell whose mean radius is the comoving distance to a central observer at the snapshot time. The shells contain all the required information about the particles (positions and velocities) and about the grid cells (gravitational potential and acceleration). Furthermore, the shells are produced with a non-zero thickness, in the sense that every spatial cell appears at different times, which allows us to compute time derivatives.

We produce three different light cones for a given observer at the centre of the simulation, which correspond to three different depths and sky coverage. The simulation box size allows us to

build a full-sky cone up to a radius equal to half the box length, corresponding to $z \lesssim 0.5$. Going further would imply that some parts of the cone would repeat due to the periodic boundary conditions of the simulation. Such replication effects are suppressed by reducing the angular width of the light cone beyond $z \approx 0.5$. An intermediate narrow cone is built up to $z = 2$ with a sky coverage of 2500 deg^2 , while our deep narrow cone goes up to $z = 10$ covering 400 deg^2 . The cones are oriented so that light rays do not cross the same structures at different times.

Halo on the light cones are identified using the parallel Friend-of-Friend (PFoF) code (Roy et al. 2014) with linking length $b = 0.2$ and at least 100 DM particles. A halo's position is defined from its centre of mass, while its velocity is defined as the mean velocity of the particles that it contains. The properties of the DM haloes of the present simulation have been studied in the appendix of Corasaniti et al. (2018).

We choose to model neither the haloes' intrinsic luminosity, nor the luminosity threshold for their detection by the observer. Thus, the results presented in this article exploit all the available data within the redshift ranges of interest.

3.3. 3D relativistic ray tracing

Observables are extracted from the light cones using a fully relativistic ray-tracing procedure based on the MAGRATHEA library (Reverdy 2014). Ray-tracing is performed backwards, that is, towards the past starting from the observation event where $\lambda = 0$. Initial conditions are fixed by the observation direction \mathbf{n} and by setting $k^0 = 1$. This means that the affine parameter coincides with conformal time at O. The observer is chosen to be comoving, meaning that its peculiar velocity is set to zero, $\mathbf{v}_o = \mathbf{0}$. This implies that $k^i \propto n^i$, where the proportionality factor is such that $k^\mu k_\mu = 0$.

From these initial conditions, the geodesic Eqs. (2) and (3) are integrated numerically with a fourth-order Runge-Kutta integrator. Specifically, photon trajectories are computed within the 3D AMR structure with four steps per AMR cell. Since the underlying N -body code uses a Triangular Shaped Cloud (TSC) interpolation scheme, we use an inverse TSC to estimate the gravitational potential and acceleration at the exact position of a photon. Using another interpolation method may lead to inconsistencies, such as self-accelerating particles.

The 3D TSC scheme requires 27 cells with the same refinement level to interpolate the value of a field at a position \mathbf{x} . In practice, we start with the finest level, that is, the level of the smallest cell that contains \mathbf{x} ; if there are less than 27 neighbouring cells with the same refinement level, then we try again with the next coarser level, and so on.

We stop the ray tracing if (and only if) that operation is impossible even at the coarse level, which means that the ray reaches the limits of our numerical background light-cones (described in Sect. 3.2) and there is no more data available to pursue its propagation. Importantly, we save all the information about every integration step of each ray's trajectory. Besides, rays are traced irrespective of the structures that they encounter; in other words, matter is assumed to be transparent.

3.4. Infinitesimal beams

The most common way to numerically evaluate the distortion matrix \mathcal{A} is based on the multi-plane lensing formalism (Blandford & Narayan 1986), where the matter distribution near the line of sight (LOS) is projected onto various planes which are then treated as thin lenses (Jain et al. 2000; Hilbert et al.

2009). Here we want to fully exploit the 3D information of the RAMSES AMR octree. For that purpose, a first option consists in integrating the projected Hessian matrix $\nabla_a \nabla_a \phi$ of the gravitational potential along the actual trajectories of light rays,

$$\mathcal{A}_{ab} = \delta_{ab} - \frac{2}{c^2} \int_0^{r_s} dr \frac{(r_s - r)r}{r_s} \nabla_a \nabla_b \phi[\eta(r), \mathbf{x}(r)], \quad (50)$$

where r_s is the comoving distance to the source, a, b take the values 1, 2, and the two-dimensional gradient ∇_a is transverse to \mathbf{n} (to the LOS). In practice, the 3D Hessian $\partial_i \partial_j \phi$ is computed on the mesh, and then converted in spherical coordinates to extract its angular (transverse) part $\nabla_a \nabla_b \phi$.

We shall refer to this approach as the 'infinitesimal-beam method', because it describes the distortions of an infinitesimal light source. We note that this way of computing \mathcal{A} is comparable to the method used in RAY-RAMES (Barreira et al. 2016), except that here $\nabla_a \nabla_b \phi$ is evaluated on the actual ray trajectory rather than on the background trajectory. In other words, we do not resort to the Born approximation.

3.5. Ray bundles and finite-beam effect

The second option to compute the distortion matrix \mathcal{A} is based on a bundle of rays (*a minima* three), which may be seen as a finite light beam subtended by an extended light source (Fluke et al. 1999; Fluke & Lasky 2011). In that 'ray-bundle method', each ray is accompanied with four auxiliary rays making an angle ε with the central one, as depicted in Fig. 5. The components of \mathcal{A} are then computed from finite coordinate differences between the rays, rather than from gradients.

More precisely, our method goes as follows: First, stop the central ray when the relevant parameter (such as redshift, time or comoving distance) has reached the desired value; this defines the fiducial source event S. Second, stop the auxiliary rays at the same affine parameter λ as the central ray's at S. This criterion is arbitrary and other possibilities are implemented in the code. Third, project the relative positions of the auxiliary rays on some source plane. For simplicity, we chose it to be orthogonal to the LOS⁸ θ . This defines the transverse separation ξ between the auxiliary rays and the central one.

Our estimator $\hat{\mathcal{A}}$ for the distortion matrix is then motivated by the fact that if two rays separated by a small $\Delta\theta$ at O should have angular coordinates differing by $\Delta\beta = \Delta\xi/r = \hat{\mathcal{A}} \Delta\theta$,

$$\hat{\mathcal{A}} \equiv \frac{1}{2r\varepsilon} \begin{bmatrix} (\xi_A - \xi_C) \cdot \mathbf{e}_1 & (\xi_B - \xi_D) \cdot \mathbf{e}_1 \\ (\xi_A - \xi_C) \cdot \mathbf{e}_2 & (\xi_B - \xi_D) \cdot \mathbf{e}_2 \end{bmatrix}, \quad (51)$$

where r is the radial position where the central ray was stopped, and $\mathbf{e}_1, \mathbf{e}_2$ are unit vectors defining the initial separation of the auxiliary rays with respect to the central ray.

We note that the choices made in step 2 and 3 induce a spurious tilt in our estimate of the distortion matrix. We have checked that the effect of this tilt is negligible in all the results involving $\hat{\mathcal{A}}$ in this article. The specific analysis of the tilt in Sect. 3.8, which require a better accuracy, will rely on a different method.

The finite separation of the rays in the bundle method may cause some discrepancies with the infinitesimal-beam approach. Those may be quantified using the finite-beam formalism developed by Fleury et al. (2017b, 2019a,b)⁹. In particular, the

⁸ We have also implemented another prescription where the screen is orthogonal to the photon direction.

⁹ This formalism was initially developed to tackle the Ricci-Weyl paradox in the gravitational-lensing theory. It was later applied to determining the effect of the finite size of light sources in weak-lensing surveys.

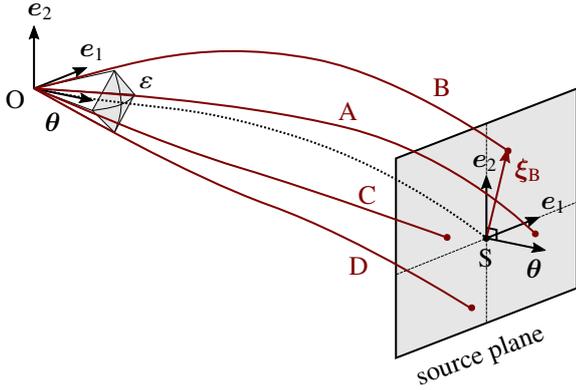


Fig. 5. Ray-bundle method. A central light ray (dotted line) is accompanied with four auxiliary rays A, B, C, D (red solid lines). Each auxiliary ray makes an angle ε at O with respect to the central ray. The distortion matrix is estimated by comparing the relative positions of the auxiliary rays in a plane orthogonal to the line of sight θ .

finite-beam corrections to the angular power spectrum of convergence, P_κ , and shear, P_γ , are found to read

$$\frac{P_\kappa(\ell; \varepsilon)}{P_\kappa(\ell; 0)} = \frac{1 + J_2(2\varepsilon\ell) - J_0(2\varepsilon\ell) + 2J_2(\sqrt{2}\varepsilon\ell)}{2(\varepsilon\ell)^2}, \quad (52)$$

$$\frac{P_\gamma(\ell; \varepsilon)}{P_\gamma(\ell; 0)} = \frac{1 - J_0(2\varepsilon\ell)}{(\varepsilon\ell)^2}, \quad (53)$$

where $P_{\kappa,\gamma}(\ell; 0)$ denote the power spectra computed with the infinitesimal-beam approach described in Sect. 3.4.

The complete derivation of Eqs. (52) and (53) is provided in Appendix B; it relies on the weak-lensing, flat-sky, and Limber approximations. We note that Eqs. (52) and (53) differ from the results highlighted in Fleury et al. (2019a, Eqs. (120) and (121)), because they correspond to different beam geometries. The latter were computed from the distortions of circular beams, while the former correspond to square-shaped beams as depicted in Fig. 5.

Figure 6 compares the predictions of Eqs. (52) and (53) with ray tracing. Three different beam semi-apertures are considered: $\varepsilon = 35$ arcmin, 3.5 arcmin and 0.35 arcmin, to which we may add $\varepsilon = 0$ corresponding to infinitesimal beams. For each value of ε but 0, we compute the convergence and shear using the ray-bundle method, at $z = 1.95$ on the intermediate narrow light cone, and for LOS dictated by HEALPIX (Górski et al. 2005). Power spectra are extracted using POLSPICE (Szapudi et al. 2001; Chon et al. 2004), so as to correctly allow for the angular selection function associated with the narrow cone's geometry.

Power-spectrum estimates from HEALPIX turn out to be robust until $\ell \approx \text{nside}^{10}$. Since finite-beam effects typically kick in from $\ell \sim \varepsilon^{-1}$, we set $\text{nside} = 4096$ for $\varepsilon = 35$ arcmin and $\varepsilon = 3.5$ arcmin, while we set $\text{nside} = 8192$ for the smallest beam size $\varepsilon = 0.35$ arcmin, so as to ensure that the power spectra are reliable at the scales of interest.

The excellent agreement between Eqs. (52) and (53) and ray tracing, as shown in Fig. 6, is the first numerical evidence of the accuracy of the finite-beam formalism. This confirms that the finite-beam corrections that may arise in the present work are well understood and under control. In particular, the damping of $P_\kappa(\ell; \varepsilon)$ $P_\gamma(\ell; \varepsilon)$ is expected to slightly reduce the variance of

¹⁰ The total number of pixels in a full-sky HEALPIX map is given by $N_{\text{pix}} = 12 \times \text{nside}^2$.

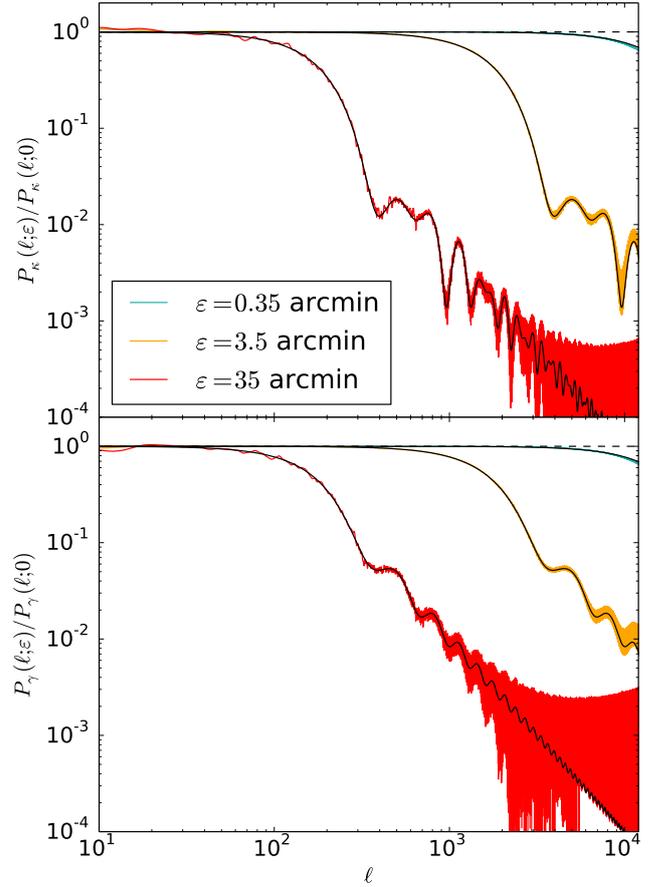


Fig. 6. Finite-beam corrections to the angular power spectrum of convergence (top panel) and shear (bottom panel) at $z = 1.95$, for different semi-aperture sizes. Black lines indicate the theoretical predictions of Eqs. (52) and (53), while coloured lines indicate ray-tracing results.

convergence and shear, which are involved in distance biases. Such effects will not change the conclusions of our analysis.

Except otherwise stated, in the remainder of this article we set $\varepsilon = 0.35$ arcmin. Smaller beam sizes are excluded because they would exceed the resolution of the simulation.

3.6. Producing observables for statistical averages

We now turn to the generation of observables, for the purpose of computing statistical averages. As seen in Sect. 2, directional averaging and source averaging are distinct operations for which different numerical techniques must be applied.

3.6.1. HEALPIX maps for directional averages

Directional averaging consists in affecting equal weights to all directions of the observer's sky. This condition is easily satisfied by dividing the sky into pixels of equal area, which is the purpose of HEALPIX. In order to estimate the directional average $\langle X \rangle_d$ of an observable X , we thus shoot a ray bundle in each direction θ dictated by HEALPIX, compute $X(\theta)$, and take their average.

3.6.2. Halo catalogues for source averaging

Source averaging gives the same statistical weight to each source on the observer's light cone. Thus, computing source averages

Table 1. Three light cones are used for the present work.

Cone	Area (deg ²)	z_{\max}	N_{haloes}	Nside
Full-sky	–	0.5	1.4×10^7	2048
Intermediate narrow	2500	2	1.2×10^7	4096
Deep narrow	400	10	3×10^6	8192

Notes. This table indicates the type of light cone, the area covered, the maximum redshift, the number of DM haloes in our catalogues and the nside parameter used for HEALPIX maps (except otherwise stated).

requires to produce a source catalogue, and to determine the null geodesic that connects each source to the observer.

In this work, sources are identified with the DM haloes, which are extracted from the simulation as described in Sect. 3.2. Geodesic identification, besides, follows Breton et al. (2019) (see also Adamek et al. 2019). In a nutshell, a photon is shot towards the comoving direction of a source; due to gravitational lensing the photon generally misses the source, so that LOS must be corrected and the operation iterated upon convergence at the source (for an illustration, see Fig. 1 in Breton et al. 2019). This procedure eventually yields the full trajectory of light for each source, as well as its observed position. We note that we do not account for multiple images of the same source, meaning that we stop the geodesic-finding algorithm as soon as one valid ray is found.

From the N -body code we also know the gravitational potential and velocity of each source in the catalogue. This data notably allow us to accurately compute the redshift, accounting for all the special- and general-relativistic effects at first order in the metric perturbation. The quantitative features of the mocks¹¹ used in this work are summarised in Table 1.

3.7. Surfaces on the light cone

In this article, we shall consider various ways to slice the observer’s past light cone, depending on which parameter is fixed; namely: Surfaces of constant redshift (iso- z), constant time (iso- η), constant comoving distance travelled (iso- s), and constant affine parameter (iso- λ). In the background FLRW model, all these surfaces are spherical and correspond to each other following specific one-to-one relations. These are denoted with an over-bar; for instance, $\bar{z}(\eta)$ is the background redshift on the background iso- η surface. In practice, we determine these background relations by shooting a single ray in the simulation with $\phi = 0$.

In the inhomogeneous case, the surfaces are determined by shooting rays in directions θ set by HEALPIX. Since all the properties of the ray and its location are saved at each integration step, it is straightforward to determine the perturbed surfaces, such as iso- η surfaces $\mathbf{x}(\eta)$, as well as the value of all the other parameters across the surfaces, so that $z(\eta, \theta) \neq \bar{z}(\eta)$. The comoving distance travelled s is computed at each integration step according to $s_{i+1} = s_i + |\mathbf{x}_{i+1} - \mathbf{x}_i|$.

Subtleties arise in the case of iso- z surfaces. The significant contribution of peculiar velocities to the observed redshift raises two issues: First, since velocities are only defined for particles, interpolation on the grid is necessary to estimate z at each time step. For that purpose, we use a TSC interpolation using all the

DM particles in the redshift range of interest with a buffer zone. Second, it may happen that a light ray meets the same redshift multiple times during its propagation. In other words, the function $\lambda \mapsto z(\lambda, \theta)$ is not one-to-one in the inhomogeneous Universe; iso- z surfaces are not uniquely defined. In this work we restrict the analysis to the two extremal iso- z surfaces, namely the closest to the observer $\{\mathbf{x}[r_{\min}(z, \theta), \theta]\}$, and the farthest from the observer $\{\mathbf{x}[r_{\max}(z, \theta), \theta]\}$. We denote these surfaces $\Sigma_-(z)$ and $\Sigma_+(z)$ respectively.

3.8. Computing the area of wrinkly iso- η surfaces

In order to check the theoretical predictions of Sect. 2.5.3 regarding the area of the LSS, and more generally of the iso- η surfaces, we need to numerically evaluate the expression

$$A(\eta) = a^2(\eta) \int_{\mathbb{S}^2} d^2\boldsymbol{\beta} r^2(\eta, \boldsymbol{\beta}) \sqrt{1 + \left| \frac{\partial \ln r}{\partial \boldsymbol{\beta}} \right|^2}, \quad (54)$$

where $\boldsymbol{\beta}$ denotes the ‘true’ position of a point of the iso- η surface, as opposed to the direction θ in which it would be observed. Such a computation thus requires the numerical determination of $r(\eta, \boldsymbol{\beta})$ and its gradient $\partial r / \partial \boldsymbol{\beta}$.

In practice, however, we have a more direct access to $r(\eta, \theta)$ because the iso- η surface is determined by ray shooting (see Sect. 3.7), which yields $r(\eta, \theta)$ and $\boldsymbol{\beta}(\eta, \theta)$ for each θ of a HEALPIX map. One could in principle compute $r(\eta, \boldsymbol{\beta})$ by finding the null geodesics between the observer and the direction $\boldsymbol{\beta}$ at each iso- η surface, but this procedure would be computationally expensive. Another option consists in directly building a lower-resolution HEALPIX $\boldsymbol{\beta}$ -map, such that in each pixel $r(\eta, \boldsymbol{\beta})$ is the average of the $r(\eta, \theta)$ for which $\boldsymbol{\beta}(\eta, \theta)$ falls into that pixel.

An even cheaper possibility consists in using the fact that the conversion between θ and $\boldsymbol{\beta}$ is dictated by lensing quantities, which we do compute for each ray. We shall adopt this method here. Specifically, in terms of θ , Eq. (54) reads

$$A(\eta) = a^2(\eta) \int_{\mathbb{S}^2} d^2\theta \frac{r^2(\eta, \theta)}{\mu(\eta, \theta)} \sqrt{1 + \left| \mathcal{A}(\eta, \theta) \frac{\partial \ln r}{\partial \theta} \right|^2}, \quad (55)$$

$$= 4\pi a^2(\eta) \left\langle \frac{r^2(\eta)}{\mu(\eta)} \sqrt{1 + \left| \mathcal{A}(\eta) \frac{\partial \ln r}{\partial \theta} \right|^2} \right\rangle_d. \quad (56)$$

The quantities r, \mathcal{A}, μ are indeed evaluated in each direction θ of our HEALPIX maps, thereby making the computation of $A(\eta)$ much easier. In fact, the corrections due to the presence of μ and \mathcal{A} , that is, the difference between integrating over θ or $\boldsymbol{\beta}$, turn out to be very small – about 1% of $\delta A(\eta) / \bar{A}(\eta)$; these corrections could thus be neglected in first approximation.

We use two different methods to compute the gradient $\partial r / \partial \theta$ from the map $r(\eta, \theta)$ so as to better control numerical artefacts: First, the ‘finite differences’ method, where we estimate derivatives from finite differences between pixels. Second, the ‘spherical harmonics’ method, where we first decompose the map into spherical harmonics, $r(\eta, \theta) = \sum_{\ell, m} r_{\ell m}(\eta) Y_{\ell m}(\theta)$ and then compute gradients from the gradients of spherical harmonics. The same procedure is applied to the mask (with zero padding), which so as to normalise the gradient of the original map. We use `healpy` routines (Zonca et al. 2019).

In practice, the spherical-harmonics method requires a smoothing beforehand to ensure that we recover the initial map through the operation map $\rightarrow r_{\ell m} \rightarrow$ map. The smoothing scale must be as small as possible but larger than pixel size; we thus

¹¹ The halo catalogues, as well as convergence and magnification HEALPIX maps are available at <http://cosmo.obspm.fr/raygalsim-relativistic-halo-catalogs>

adopt a Gaussian beam with $FWHM = 5$ arcmin. Although smoothing is not needed in the finite-difference method, we apply it as well to ensure that their results are comparable.

3.9. Uncertainties on numerical averages

When computing the (source or directional) average $\langle X \rangle$ of an observable X from mock data, the result generally differs from theoretical predictions; in other words, the ergodicity principle is not exactly satisfied. This may happen for two reasons: First, the number of mock observations in the sample is finite; this leads to a Poisson uncertainty on the estimation of any average quantity. Second, there may be super-sample inhomogeneity modes (Hui & Greene 2006), which may bias the estimator of $\langle X \rangle$. This is particularly relevant to the mock data extracted from the narrow cones, which may be, for example, slightly over-dense or under-dense with respect to the simulation box.

We shall account for this uncertainty by adding error bars on any numerical average presented in the next section. The size of such error bars, that is, the uncertainty σ on $\langle X \rangle$, is computed as

$$\sigma^2 = \underbrace{\frac{1}{N} \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} C_{\ell}^X}_{\text{Poisson}} + \underbrace{\sigma_{\text{ss}}^2}_{\text{super sample}}. \quad (57)$$

In Eq. (57), the first term represents the Poisson uncertainty due to the finite sample size; C_{ℓ}^X denotes the ℓ th multipole of X , and N is the number of mock observations – the number of pixels in the map for directional average, or the number of sources for source-averaging. When $N \gg 1$, this first term may be neglected.

The second term in Eq. (57), σ_{ss}^2 , is the super-sample variance. This contribution may be understood as a generalisation of the cosmic variance mentioned in Sect. 2.3.3. Suppose that we compute the directional average of X within a cone with half angle α at the observer,

$$\langle X \rangle_{\alpha} = \int_0^{\alpha} d\vartheta \sin \vartheta \int_0^{2\pi} d\varphi X(\vartheta, \varphi) = \int_{\mathbb{S}^2} d^2\theta W(\theta) X(\theta), \quad (58)$$

with $\theta = (\vartheta, \varphi)$, and where

$$W(\theta) \equiv \frac{[\vartheta < \alpha]}{2\pi(1 - \cos \alpha)}, \quad (59)$$

is the cone's window function. We note that $\langle X \rangle_{\alpha=\pi} = \langle X \rangle_{\text{d}}$ by definition. $\langle X \rangle_{\alpha}$ is a random variable, because it depends on the actual orientation of the cone. Its variance is the super-sample variance σ_{ss}^2 that we are looking for,

$$\sigma_{\text{ss}}^2 = \langle \langle X \rangle_{\alpha}^2 \rangle = \int d^2\theta d^2\theta' W(\theta) W(\theta') \langle X(\theta) X(\theta') \rangle. \quad (60)$$

Decomposing the window function X and observable W in spherical harmonics, for which

$$W_{\ell m} = \begin{cases} \frac{\delta_{m0}}{\sqrt{4\pi}} & \ell = 0, \\ \frac{\delta_{m0}}{\sqrt{4\pi(2\ell+1)}} \frac{P_{\ell-1}(\cos \alpha) - P_{\ell+1}(\cos \alpha)}{1 - \cos \alpha} & \ell \geq 1, \end{cases} \quad (61)$$

where P_{ℓ} are Legendre polynomials, we find

$$\sigma_{\text{ss}}^2 = \sum_{\ell, m} |W_{\ell m}|^2 C_{\ell}^X, \quad (62)$$

$$= \frac{1}{4\pi} \left[C_0^X + \sum_{\ell=1}^{\infty} \frac{C_{\ell}^X}{2\ell+1} \left| \frac{P_{\ell-1}(\cos \alpha) - P_{\ell+1}(\cos \alpha)}{1 - \cos \alpha} \right|^2 \right]. \quad (63)$$

In the full-sky limit ($\alpha = \pi$), the uncertainty on $\langle X \rangle_{\alpha}$ is expectedly dictated by the monopole only, $\sigma_{\text{ss}}^2 = C_0^X/4\pi$.

3.10. Variance within a finite simulation box

The variance derived in Sect. 3.9 depends on C_{ℓ}^X , which ultimately depends on the matter density power spectrum $P(k)$. As such, it would seem natural to use the information from all the wavelengths available in $P(k)$. However there is a subtlety when estimating the variance from N -body simulations: these are usually cubic boxes with periodic boundary conditions, with a mean density equal to zero inside the cubic volume by definition. This means that, unlike the real Universe, there can be no inhomogeneity modes with wavelengths larger than the box itself.

To mimic this effect, Gelb & Bertschinger (1994) imposed a cut-off in the matter power spectrum at $k_{\text{min}} = 2\pi/L$ with L the comoving size of the box. This approach has been widely studied either to estimate 3D statistics (Bagla & Ray 2005; Power & Knebe 2006) or 2D weak-lensing analysis (Harnois-Déraps & van Waerbeke 2015). To go further, one may convolve the power spectrum with the appropriate cubic window function in real space (Pen 1997; Sirko 2005); we found that this last correction was negligible because our box is large enough.

Finite-box corrections effectively change the low- k behaviour of the power spectrum of any quantity that depends on the gravitational potential or on the density contrast. As a consequence, the angular power spectrum C_{ℓ}^X of any related observable ℓ is modified at low ℓ compared to its theoretical predictions in an infinite Universe. Depending on the shape of C_{ℓ}^X , this effect may be more or less pronounced; in particular, we expect a strong impact when most of the power is carried by large scales. In that case, it is crucial to carefully account for finite-box effects as well as evaluating power spectra beyond Limber's approximation. See Kilbinger et al. (2017) and references therein for a review about low- ℓ corrections in weak-lensing studies.

We finally mention that, besides the aforementioned low- k corrections, there are high- k corrections due to mass assignment, shot noise and aliasing. These are already well known (Hockney & Eastwood 1981) and negligible in the present study.

3.11. Constrained Gaussian random field and ensemble averaging

The ergodic principle does not hold when averaging over a small volume. For example, the ensemble average of the gravitational potential vanishes by definition, $\langle \phi \rangle = 0$; yet, its average over a small spatial region around the observer should be almost equal to ϕ_0 , which in general is non-zero. This mis-match, due to small-scale correlations, may lead to spurious discrepancies between ensemble averages and numerical averages at low redshift.

It is possible to smoothly transition from the constraint at the observer to the expectation from ensemble average through the constrained random field formalism (Hoffman & Ribak 1991; van de Weygaert & Bertschinger 1996; Mitsou et al. 2020). Take again the example of the gravitational potential ϕ near the observation event, which is subject to the constraint $\phi(\mathbf{0}) = \phi_0$. Following Desjacques et al. (2021), the constrained ensemble average of ϕ then reads

$$\langle \phi(x) | \phi_0 \rangle = \phi_0 \frac{\xi_{\phi}(r)}{\xi_{\phi}(0)}, \quad (64)$$

where $\xi_{\phi}(r)$ is the unconstrained two-point correlation function of the gravitational potential and $\xi_{\phi}(0) = \sigma_{\phi}^2$ its variance at $z = 0$. We implicitly assume that all quantities are evaluated on the light cone to alleviate notation, $\xi_{\phi}(r) \equiv \xi_{\phi}(\eta_0 - r, r)$. For numerical

applications, $\xi_\phi(r)$ is estimated from the linear power spectrum with an infrared cutoff at $k_{\min} = 2\pi/L$, as per Sect. 3.10.

The constraint $\phi(\mathbf{0}) = \phi_0$ also impacts the two-point correlation function of ϕ . Indeed, since its value is fixed at $\mathbf{0}$, we expect the variance of ϕ to vanish as we approach that point. The constrained two-point correlation function reads

$$\zeta_\phi(\mathbf{x}, \mathbf{x}') \equiv \langle \phi(\mathbf{x})\phi(\mathbf{x}') | \phi_0 \rangle - \langle \phi(\mathbf{x}) | \phi_0 \rangle \langle \phi(\mathbf{x}') | \phi_0 \rangle, \quad (65)$$

$$= \xi_\phi(|\mathbf{x} - \mathbf{x}'|) - \frac{\xi_\phi(r)\xi_\phi(r')}{\xi_\phi(0)}. \quad (66)$$

Equations (64) and (66) are particularly useful for fields that are mostly correlated on large scales, such as the gravitational potential whose power spectrum scales as $P(k)/k^4$.

4. Results

In this section, we confront the theoretical predictions of Sect. 2 with numerical results obtained with the methods described in Sect. 3. Directional averages and source-averages of cosmological distance indicators are considered in Sect. 4.1. In the next sub-sections, we then focus on the rather subtle shift (Sect. 4.2) and tilt (Sect. 4.3) corrections to the amplification $\tilde{\mu}$, that is, to the area of various light-cone slices.

4.1. Directional and source averages

4.1.1. Directional averaging

We analyse here the statistical properties of the HEALPIX maps generated as described in Sect. 3.6.1 for directional averages. To avoid numerical uncertainties due to short light propagation, we focus on data with $z \geq 0.2$.

Figure 7 shows how much the directional average of the inverse geometric and observable magnifications depart from unity. Dots indicate numerical averages, while error bars account for both Poisson and super-sample variance as described in Sect. 3.9; an exception is the full-sky estimate of $\langle \mu^{-1} \rangle_d$ for $z < 0.5$, which is only affected by Poisson variance. The exact expressions that we use are provided in Appendix D.1.

First of all, we note that we do not recover exactly $\langle \mu^{-1}(z) \rangle_d = 1$, which yet should be exactly satisfied. For the full-sky cone ($z < 0.5$) the discrepancy is extremely small ($< 10^{-6}$) and is attributed to the discretisation of the full-sky map: averages are performed over a large but finite number of points. This interpretation is supported by the fact that $\langle \mu^{-1}(z < 0.5) \rangle_d$ is consistent with unity within the Poisson uncertainty. For the two other cones ($z > 0.5$), the discrepancy is larger ($\sim 10^{-4}$). This should not come as a surprise, because $\langle \mu^{-1} \rangle_d = 1$ is exact on the full sky only. The intermediate and deep narrow cones are simply slightly under-dense or over-dense with respect to the average box. Again, this interpretation is supported by the fact that $\langle \mu^{-1}(z) \rangle_d - 1$ falls well into the error bars accounting for super-sample variance.

The direction-averaged inverse amplification $\langle \tilde{\mu}^{-1}(z) \rangle_d$ departs from 1 by almost 10^{-3} even for the full-sky data. We may note that, unlike μ , $\tilde{\mu}$ is subject to super-sample variance even on a full sky; however, the main reason for which $\langle \tilde{\mu}^{-1}(z) \rangle_d \neq 1$ at low z is the shift effect due to peculiar velocities (see Sect. 2.2.3). For the other two cones $\langle \tilde{\mu}^{-1}(z) \rangle_d - 1$ is mostly due to the super-sample variance as the one affecting $\langle \mu^{-1}(z) \rangle_d - 1$.

Our results show that the approximation $\langle \tilde{\mu}^{-1}(z) \rangle_d \approx 1$ is accurate up to 10^{-3} up to $z = 10$. They also indicate that incomplete sky coverage in actual observations may in the end be the main cause of any departure from $\langle \tilde{\mu}^{-1}(z) \rangle_d = 1$ at high redshift.

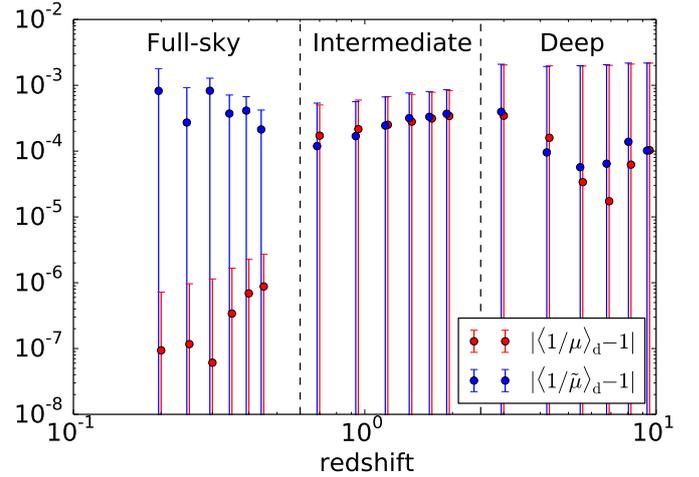


Fig. 7. Departures from 1 of the directional average of the inverse geometric magnification $\langle \mu^{-1}(z) \rangle_d$ and of the inverse observable magnification $\langle \tilde{\mu}^{-1}(z) \rangle_d$ at constant redshift z . Dots indicate numerical averages over HEALPIX maps, while error bars allow for super-sample and Poisson variance; see Appendix D.1 for their expressions. The full-sky variance on $\langle \mu^{-1}(z < 0.5) \rangle_d$ does not contain super-sample variance. The three vertical dashed lines indicate the limits of our three light cones: full-sky ($z < 0.5$), intermediate narrow ($0.5 < z < 2$), and deep narrow ($2 < z < 10$).

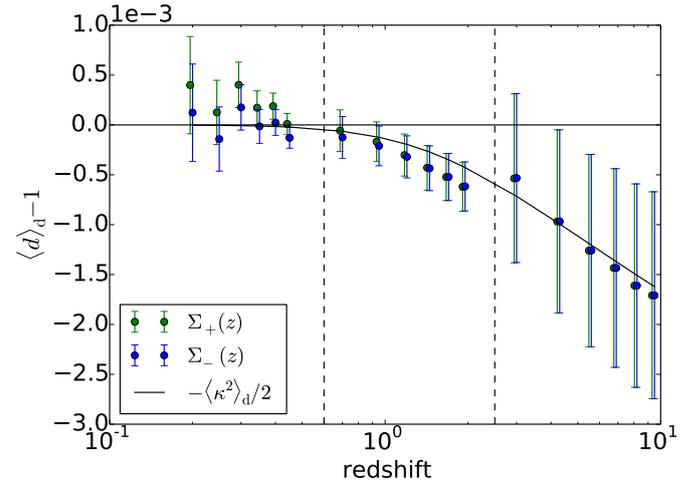


Fig. 8. Bias on the distance-redshift relation when averaged over directions, compared with the theoretical prediction $\langle d(z) \rangle_d = 1 - \langle \kappa^2(z) \rangle / 2$. Since the iso- z surface is not unique, we indicated results for the closest surface $\Sigma_-(z)$ and farthest surface $\Sigma_+(z)$ from the observer. Expressions for the error bars may be found in Appendix D.1.

We now turn to the bias of angular or luminosity distance. For directional averaging, we expect $d(z)$ to be negatively biased according to $\langle d(z) \rangle_d \approx 1 - \langle \kappa^2(z) \rangle / 2$. We evaluate $\langle \kappa^2 \rangle$ from the data directly as $\langle \kappa^2 \rangle_d$, and we have checked that a pure-theory estimate based on the matter power spectrum gives the same results. As shown in Fig. 8, that theoretical prediction is in good agreement with numerical results within the error bars dominated by super-sample variance. Again, low-redshift departures are due to peculiar velocities whose shift effect is not accounted for in κ^2 . While $\langle \mu^{-1}(z) \rangle_d$ and $\langle \tilde{\mu}^{-1}(z) \rangle_d$ remains unity within error bars, $\langle d(z) \rangle_d - 1$ clearly departs from zero for $z > 1$.

We finally evaluate the accuracy of the expansion that allowed us to express all distance biases in terms of $\langle \kappa^2(z) \rangle$ in Sect. 2.4. We shall take the direction-averaged magnification in

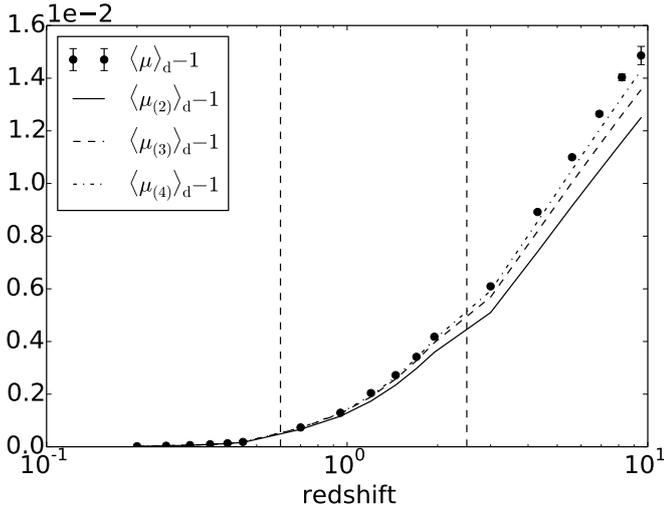


Fig. 9. Accuracy of the second-order expansion for the direction-averaged magnification. The error bars are only due to Poisson variance.

order to illustrate that point. Taylor-expanding μ at order n in $\mu^{-1} - 1$ yields

$$\mu = \left[1 - (1 - \mu^{-1})\right]^{-1} = \underbrace{\sum_{k=0}^n (1 - \mu^{-1})^k}_{\mu^{(n)}} + O(1 - \mu^{-1})^{n+1}. \quad (67)$$

Figure 9 confronts $\langle\mu\rangle_d$ with its Taylor-expansion $\langle\mu^{(n)}\rangle_d$ for $n = 2, 3, 4$, where all these quantities are evaluated numerically. The error bars only allow for Poisson errors, because super-sample variance affects μ and $\mu^{(n)}$ in the same way. We see that the quadratic expansion $\langle\mu^{(2)}\rangle_d$ provides a good approximation of the exact result for $z < 1$. Beyond that, we observe discrepancies reaching about 10% at $z = 10$. These are due to departures from the weak-lensing regime ($|\mu^{-1} - 1| \ll 1$), which are more likely to happen as the redshift increases. Figure 9 also illustrates the convergence of the series expansion of Eq. (67).

4.1.2. Source averaging

We now analyse the mock halo catalogues produced for the three light cones (full sky, intermediate narrow and deep narrow) as described in Sect. 3.6.2. We arrange the haloes in tomographic bins of width $\Delta z = 0.08, 0.2$ and 1.5 for the full sky, intermediate and deep light cones, respectively.

We have seen in Sect. 2.3.2 that the source-average of the geometric and observable magnifications are expected to be almost unity $\langle\mu(z)\rangle_s \approx \langle\tilde{\mu}(z)\rangle_s \approx 1$ ¹². Unlike directional-averaging, departures from equality may be caused by the non-trivial clustering of sources in addition to the shift and tilt corrections responsible for $\tilde{\mu} \neq \mu$. Numerical results for $\langle\mu(z)\rangle_s, \langle\tilde{\mu}(z)\rangle_s$ are

¹² In principle these (approximate) relations should apply to the total absolute magnification rather than to the signed magnification. As mentioned in Sect. 2.3, they coincide in the absence of multiple imaging, which is a good approximation here. In fact, no negative-parity image was found in our halo catalogue. This is mostly due to the source-averaging procedure, which is expected to give less strong-lensing events than the directional averaging one, coupled with the fact that there are less sources at high redshift. Another reason is the use of ray bundles, which tend to smooth out the matter inhomogeneities on very small scales and thereby reduce the occurrence of strong lensing.

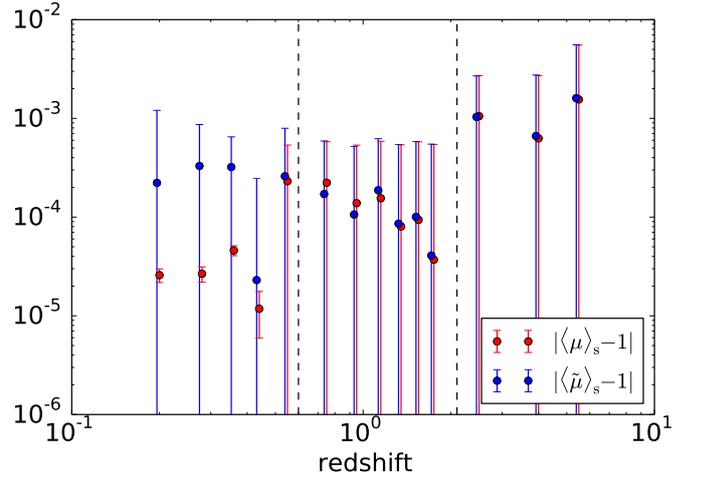


Fig. 10. Departures from 1 of the source-average of the geometric magnification $\langle\mu(z)\rangle_s$ and of the observable magnification $\langle\tilde{\mu}(z)\rangle_d$ at constant redshift z . Dots indicate numerical averages over binned halo catalogues, while error bars allow for super-sample and Poisson variance; see Appendix D.1 for their expressions. The full-sky variance on $\langle\mu(z)\rangle_s$ does not contain super-sample variance.

depicted in Fig. 10. Again, error bars account for both Poisson and super-sample (cosmic variance) whose expressions are given in Appendix D.1. Since the number of mock haloes per bin is smaller than the number of pixels of the maps used in Sect. 4.1.1, Poisson variance is larger here, especially at high redshift where it exceeds super-sample variance.

For the full-sky data ($z < 0.5$), we observe that $|\langle\mu(z)\rangle_s - 1| \approx 10^{-5}$, which is about 100 times larger than what was found for $|\langle\mu^{-1}(z)\rangle_d - 1|$. This discrepancy, which goes beyond the estimated uncertainty, is due to the fact that the latter does not properly account for the spatial clustering of haloes. Halo clustering is present in the $d^2N/d^2\theta$ kernel in the definition (22) of source averaging. The correlation between spatial clustering and lensing convergence was predicted to be on the order of 10^{-5} in Fleury et al. (2017a), Fanizza et al. (2020), which agrees with the present results. As for $\langle\tilde{\mu}\rangle_s$, just as in Fig. 7 the numerical results are in agreement with unity within the error bars dominated by peculiar velocities.

The interpretation of the results for the intermediate ($0.5 < z < 2$) cone is similar to the directional-averaging case. The relative effect of the shift, that is, the main difference between μ and $\tilde{\mu}$, reduces as the impact of super-sample variance increases. As for the deep cone, Poisson variance dominates due to the small number of haloes per tomographic bin. In both cases, $\langle\mu\rangle_s$ and $\langle\tilde{\mu}\rangle_s$ are compatible with unity within error bars, so that no unexpected bias arises.

We now turn to more observationally relevant biases. For SN surveys, the common practice consists in fitting the magnitude-redshift relation $m(z)$ with the FLRW prediction (Scolnic et al. 2018). Such a method is thus biased by $\langle\Delta m(z)\rangle_s = 5\langle\log_{10} d(z)\rangle_s$. For standard-siren Hubble diagrams, it may be more common to fit the luminosity distance-redshift relation, which would be biased by $\langle d(z)\rangle_s$. Numerical results on these biases are reported in Fig. 11, and compared with the theoretical predictions presented in Sect. 2.4. Similarly to Sect. 4.1.1, we estimate the variance of the convergence from the data itself. We note that $\langle\kappa^2(z)\rangle_s$ slightly differs from $\langle\kappa^2(z)\rangle_d$; their relation is essentially given by Eq. (24).

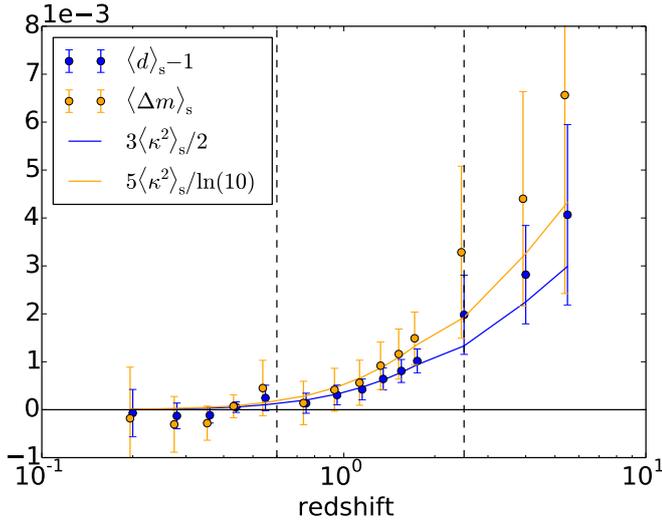


Fig. 11. Bias of the source-averaged distance-redshift $\langle d(z)_s \rangle$ (blue) and magnitude-redshift $\langle m(z)_s \rangle$ (yellow) relations. The associated solid lines are the theoretical predictions for these quantities as presented in Sect. 2.4. Error bars account for Poisson and super-sample variance; their expressions can be found in Appendix D.1.

Again, our results agree with theoretical predictions within the error bars. We note that $\langle d(z)_s \rangle$ reaches a few 10^{-3} at high redshift, which may become non-negligible for the standard-siren Hubble diagrams of the future LISA mission (Caprini & Tamanini 2016). Such a bias would be easily removed by fitting $D_L^{-2}(z)$ instead of $D_L(z)$, as pointed out by Fleury et al. (2017a).

We finally mention that, since d and m are non-linear functions of μ , theoretical predictions on their bias based on a second-order Taylor expansion are subject to the same small inaccuracies as displayed in Fig. 9. These inaccuracies are however much smaller than the uncertainty on those quantities, and hence may be safely neglected.

4.2. Focus on the shift correction

In Sect. 4.1, we have analysed the statistical biases to distance measures for both directional and source averaging. In particular, we have found no unexpected violation of $\langle \tilde{\mu}^{-1}(z) \rangle_d \approx 1$ within numerical uncertainties. As seen in Sect. 2.5, this relation may be understood in terms of the area of iso- z surfaces, namely $A(z) \approx \bar{A}(z)$ – the area is unaffected by inhomogeneities.

We now propose to further focus on the two subtle corrections that are making the above ‘ \approx ’ differ from equality, namely the shift and tilt effects (see Fig. 1). We start, in the sub-section, with the analysis of the shift, that is, the discrepancy between the mean radius of a given light-cone slice (such as iso- z or iso- η) and its value in the FLRW background.

4.2.1. Wiggly ray effect: Mean distance reached at fixed distance travelled

We first consider the mean comoving distance that is reached by a photon after it travelled over a given comoving distance. As briefly described in Sect. 2.5.3, because light rays do not travel in straight lines, the radius reached for a distance travelled s is shorter than $\bar{r} = s$. We may thus write $r(s) = \bar{r}(s) - \delta r_{\text{geo}}(s)$, where δr_{geo} encodes this wiggly-ray effect. This is illustrated with a HEALPIX map of $\delta r_{\text{geo}}(s)/\bar{r}(s)$ at $z = 0.2$ in Fig. 12. We

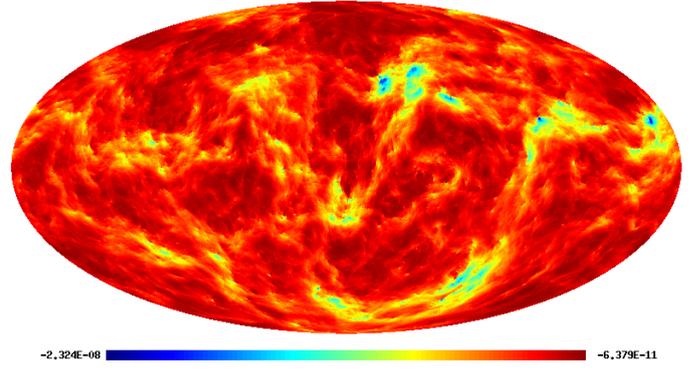


Fig. 12. Map of wiggly-ray effect, that is, the relative fluctuations of the comoving distance reached at fixed distance travelled, $\delta r_{\text{geo}}(s)/\bar{r}(s) = r(s)/\bar{r}(s) - 1$, for $s = \bar{s}(z = 0.2)$.

see that the fluctuations are very small, on the order of 10^{-8} , and vary on relatively small angular scales.

On average, the wiggly-ray correction is found to read (see Appendix C.1.2)

$$\frac{\langle \delta r_{\text{geo}}(s) \rangle}{\bar{r}(s)} = - \int_0^{\bar{r}} dr \frac{(\bar{r} - r)r}{\bar{r}^2} J(r). \quad (68)$$

It may be noted, however, that Eq. (68) was obtained with the help of a few approximations, among which is Limber’s. A slightly more accurate computation, in the spirit of Eq. (A.26) in KP16, would yield

$$\frac{\langle \delta r_{\text{geo}}(s) \rangle}{\bar{r}(s)} = -8 \int_0^{\bar{r}} dr \frac{\bar{r} - r}{\bar{r}^2} g(r) \int_0^r dR \frac{R}{r - R} g(R) \xi'_\phi(r - R), \quad (69)$$

with $g(r) \equiv D_+(\eta_0 - r)/a(\eta_0 - r)$ where D_+ is the linear growth factor, ξ_ϕ is the two-point correlation function of the gravitational potential at $z = 0$, and ξ'_ϕ is its derivative.

These theoretical predictions are successfully confronted with numerical results in Fig. 13. Error bars allow for both Poisson and correlated variance, whose expression is given in Appendix D.2. We note that the predictions of Eqs. (68) and (69) are very similar, but only the latter falls within error bars for the full-sky ($z < 0.5$) and intermediate cones ($0.5 < z < 2$).

4.2.2. Shapiro effect: Mean distance reached at fixed time

We now investigate the mean comoving distance reached at constant time $\langle r(\eta) \rangle$; in other words, we consider the shift of iso- η surfaces, which would be relevant for CMB-like observations. A HEALPIX map of $\delta r(\eta)/\bar{r}(\eta)$ is given in Fig. 14 for illustration. Compared to Fig. 12, we note that fluctuations are much larger (of order 10^{-4}) and take place on much larger angular scales.

As discussed in Sect. 2.5.3, in theory this shift may be decomposed in two components: (i) the wiggly-ray effect δr_{geo} considered above; and (ii) the Shapiro time-delay effect. The latter indeed changes the comoving distance travelled $s(\eta)$ during a given time, compared to its background counterpart $\bar{s}(\eta) = \eta_0 - \eta$, depending on the path-integrated gravitational potential experienced by light. Summarising,

$$r(\eta) = \bar{r}(\eta) + \delta r(\eta), \quad \delta r(\eta) = \delta r_{\text{geo}}(\eta) + \delta s(\eta), \quad (70)$$

and as shown in Appendix C we expect

$$\langle \delta r(\eta) \rangle = -\langle \delta r_{\text{geo}}(\eta) \rangle > 0, \quad (71)$$

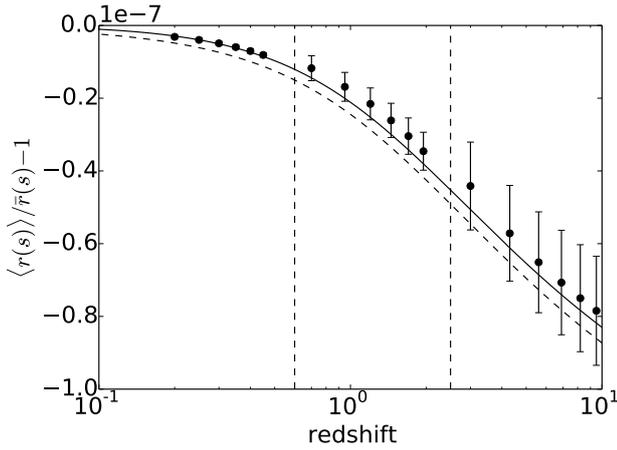


Fig. 13. Wiggly ray effect: Mean fractional reduction $\langle r(s) \rangle / \bar{r}(s) - 1$ of the comoving distance $r(s)$ reached after travelling a comoving distance s . The x -axis indicates the redshift $\bar{z}(s)$ instead of s for better readability. Dots indicate directional averages in the simulation $\langle r(s) \rangle_d$; the dashed and solid lines respectively indicate the theoretical predictions of Eqs. (68) and (69). Error bars are computed following Appendix D.2.

because of the post-Born corrections to δs which turn out to be minus twice the geometrical contribution.

Numerical results are confronted with theory in Fig. 15. The first striking feature is perhaps that error bars are about two orders of magnitude larger than the ones of Fig. 13. This is because the time-delay contribution $\delta s(\eta)$ is a first-order quantity, while $\delta r_{\text{geo}}(\eta)$ is second-order. Thus, the corresponding super-sample variance is much larger. As such, the prediction (71) naturally falls in the error bars, which is not particularly informative.

We may add an extra layer of refinement since we know the exact value of the gravitational potential in the simulation at the observer, $\phi_0 \approx -7 \times 10^{-6}$ (which is about -2 km s^{-1}). This information may be used to improve our prediction of $\delta s(\eta)$ at low redshift. Precisely, following the method outlined in Sect. 3.11, we find that the average Shapiro contribution under the constraint $\phi(0) = \phi_0$ totally overwhelms the geometrical and post-Born terms¹³, yielding

$$\langle \delta r(\eta) | \phi_0 \rangle = 2\phi_0 \int_0^{\bar{r}(\eta)} dr \frac{\xi_\phi(r)}{\xi_\phi(0)} \gg \langle \delta r_{\text{geo}}(\eta) \rangle. \quad (72)$$

The corresponding prediction is indicated by a solid line in Fig. 15, and is observed to reproduce the behaviour of numerical results at low redshift.

The fact that Fig. 15 is dominated by super-sample variance makes it a priori impossible to check the unconstrained prediction (71), which yet would be the most relevant at high redshift, where the effect of ϕ_0 becomes negligible. However, we may apply the following trick to numerically extract the second-order contribution of $\langle \delta s(\eta) \rangle$. We consider the following estimator

$$\widehat{\delta_{(2)}s(\eta)} = \frac{\langle \mu^{-1}(\eta) \delta r(\eta) \rangle_d}{\langle \mu^{-1}(\eta) \rangle_d} - \langle \delta r(\eta) \rangle_d. \quad (73)$$

By construction, this combination eliminates the wiggly-ray and first-order Shapiro contributions to $\delta r(\eta)$, while preserving the

¹³ This finding does not contradict the conclusions of Hall (2020) that the impact of our local gravitational potential is negligible in current weak-lensing and galaxy-clustering surveys, because they concern very different cosmological quantities.

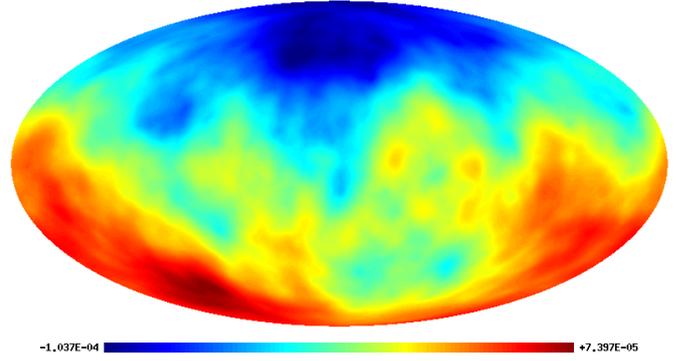


Fig. 14. Map of the relative fluctuations of the comoving distance reached at fixed time, $r(\eta) / \bar{r}(\eta) - 1$ for $\eta = \bar{\eta}(z = 0.2)$. These fluctuations are dominated by the first-order Shapiro time-delay effect.

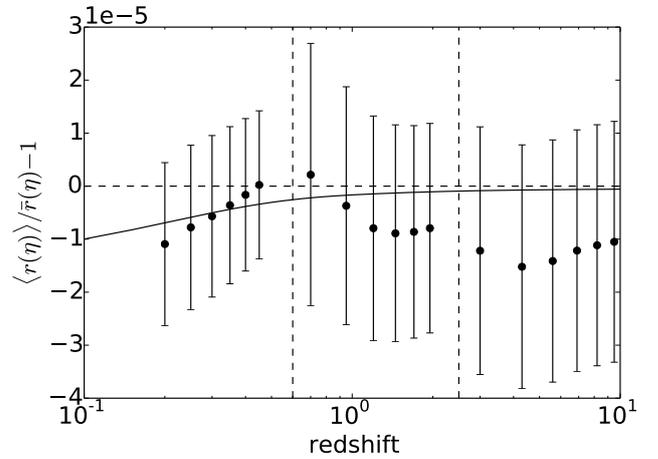


Fig. 15. Mean comoving distance reached at fixed time, $\langle r(\eta) \rangle$, compared to its background counterpart $\bar{r}(\eta) = \eta_0 - \eta$. The difference combines the wiggly-ray and Shapiro time-delay effects. The x -axis indicates $\bar{z}(\eta)$ instead of η for better readability. Dots indicate numerical results and error bars are computed according to Appendix D.4. The dashed line indicates the unconstrained theoretical prediction (71), while the solid line shows the prediction (72) that accounts for our knowledge of the gravitational potential ϕ_0 at the observer.

post-Born term of δs ; see Appendix C.2.2 for details. Thus, we expect $\widehat{\delta_{(2)}s} = -2\langle \delta r_{\text{geo}} \rangle$. This is indeed roughly what is observed in Fig. 16, although the error bars seem to be under-estimated. This confirms our understanding of the subtle behaviour of $r(\eta)$.

4.2.3. Doppler effect of peculiar velocities: Spatio-temporal shift at fixed observed redshift

We now turn to observations performed at fixed redshift z . In the inhomogeneous Universe, several phenomena may affect the observed redshift of a source at a given position: Doppler effect due to peculiar velocities, Sachs-Wolfe (SW) and ISW effects. These imply that for a given redshift, light may have been emitted slightly closer to the observer (and later), or slightly further (and earlier) compared to the background FLRW case. In other words, we have $r(z) = \bar{r}(z) + \delta r(z)$ and $\eta(z) = \bar{\eta}(z) + \delta \eta(z)$, with $\delta r = -\delta \eta$ the associated spatio-temporal shift of $\Sigma(z)$.

The Doppler effect of peculiar velocities is expected to dominate, especially at low- z . Assuming that the observer is comoving ($\mathbf{v}_0 = \mathbf{0}$), which is the case in the simulation, the shift due to

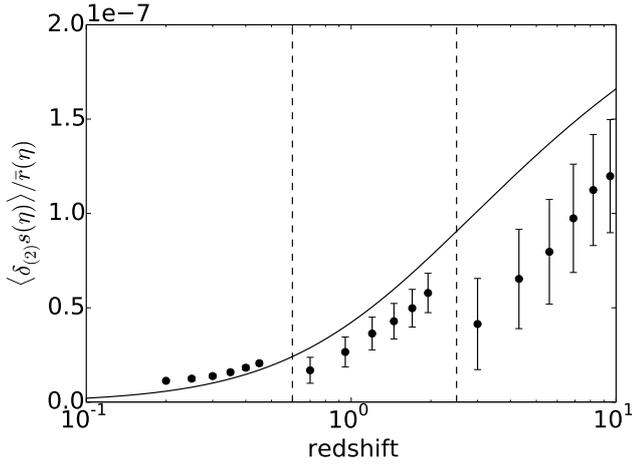


Fig. 16. Second-order (post-Born) contribution to $\delta s(\eta)$. Dots indicate the numerical estimate from Eq. (73), while the solid line is the theoretical prediction $-2\langle \delta r_{\text{geo}}(\eta) \rangle$. Error bars are twice those of Fig. 13.

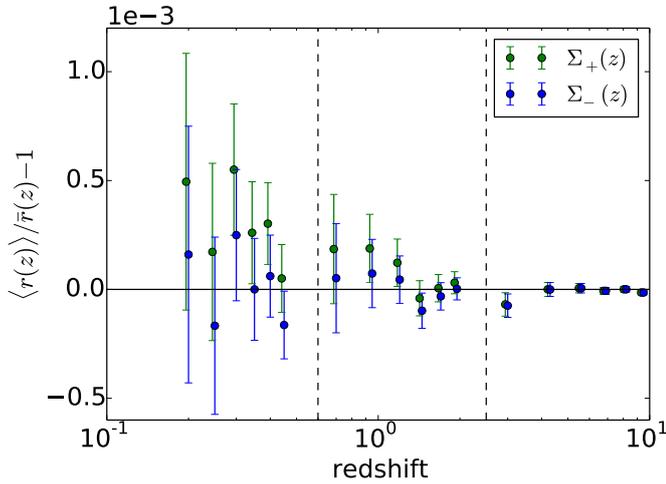


Fig. 17. Radial shift of surfaces of constant redshift. Dots indicate numerical results for $\langle \delta r(z) \rangle_d / \bar{r}(z)$ while error bars account for Poisson variance and for the super-sample variance due to peculiar velocities only. For each z , we provide two values of the mean shift: one across the closest iso- z surface [$\Sigma_-(z)$, blue] and the other across the farthest iso- z surface [$\Sigma_+(z)$, green].

the source's peculiar velocity \mathbf{v} reads

$$\frac{\delta r(\boldsymbol{\theta}, z)}{\bar{r}(z)} = -\frac{\boldsymbol{\theta} \cdot \mathbf{v}(\boldsymbol{\theta}, z)}{\mathcal{H}(z)\bar{r}(z)}, \quad (74)$$

at first order. We note that, combined with the (opposite) temporal shift, we find that the corresponding area perturbation reads $\delta[a^2(\eta)r^2]/[a^2(\bar{\eta})\bar{r}^2] = 2\bar{\kappa}_v$, in agreement with Eq. (15).

Numerical results for $\langle \delta r(z) \rangle_d$ are shown in Fig. 17. Two values of $\delta r(z)$ are provided for each z . This is because the fluctuations of peculiar velocities may be important on the light cone, so that several events may have the same redshift: $\Sigma(z)$ is not unique. We plotted here the smallest one, $\langle \min \delta r(z) \rangle_d$, and the largest one, $\langle \max \delta r(z) \rangle_d$. Numerical values for $\langle \delta r(z) \rangle_d$ account for all the redshift components at first order in metric perturbations; however, the super-sample variance contribution to the error bars only account for peculiar velocities, which are highly dominant. See Appendix D.3 for their expression.

The mean shift is compatible with zero within the large error bars, which is why we did not work on a more elaborate theoretical prediction for $\langle \delta r(z) \rangle$. We also note that the two extremal surfaces $\Sigma_-(z), \Sigma_+(z)$ that we have considered differ more as we get closer to the observer. We shall now explain this point: at high redshift there are few virialised objects, and hence velocity dispersion is small next to such objects; thus, at high z , $\Sigma(z) = \Sigma_-(z) = \Sigma_+(z)$ is typically unique. As z decreases, more haloes form, which implies that more regions have high velocity dispersion¹⁴, thereby de-multiplying $\Sigma(z)$ and spreading its occurrences. The $1/(\mathcal{H}\bar{r})$ factor in Eq. (74) further enhances that spread.

4.2.4. Affine parameter at constant time

While surfaces of constant redshift $\Sigma(z)$ or time $\Sigma(\eta)$ are observationally relevant, we may also consider surfaces of constant affine parameter $\Sigma(\lambda)$, whose interest is strictly theoretical. Such light-cone slices have the advantage of being defined regardless of any specific model for the space-time metric. The analysis of Kibble & Lieu (2005) was indeed conducted on $\Sigma(\lambda)$. We shall examine the time (or equivalently radial) shift at fixed affine parameter, $\delta\eta(\lambda)$, in the next sub-section. Before that, we propose to first consider the converse shift $\delta\lambda(\eta)$ for pedagogical reasons.

In the background FLRW model, the 0th component of the geodesic Eq. (2) is integrated as

$$\bar{\lambda}(\eta) = \int_{\eta}^{\eta_0} d\eta' a^2(\eta'). \quad (75)$$

In the presence of perturbations, this becomes

$$\lambda(\eta) = \int_{\eta}^{\eta_0} d\eta' a^2(\eta') \left[1 + 2(\phi - \phi_0) - 2 \int_0^{\eta'} d\eta'' \frac{\partial\phi}{\partial\eta''} \right]. \quad (76)$$

at first order (higher-order terms are negligible here). We note that since $\langle \phi \rangle = 0$, the mean correction to the affine parameter is simply $\langle \delta\lambda(\eta) \rangle / \bar{\lambda}(\eta) = -2\phi_0$. However, we may account for the fact that, in our simulation just as in observations, the gravitational potential at the observer is fixed, and estimate $\langle \delta\lambda(\eta) \rangle_d$ with the following constrained ensemble average

$$\langle \delta\lambda(\eta) | \phi_0 \rangle = -2\phi_0 \int_0^{\bar{r}(\eta)} dr a^2(\eta_0 - r) \left[1 - \frac{\xi_\phi(r)}{\xi_\phi(0)} \right] \neq 0, \quad (77)$$

where we chose to neglect the small ISW term of Eq. (76). In the high- z limit, this result may be approximated as

$$\frac{\langle \delta\lambda(\eta) | \phi_0 \rangle}{\bar{\lambda}(\eta)} \approx -2(1 - \Xi_\infty)\phi_0 = \text{const.}, \quad (78)$$

with

$$\Xi_\infty \equiv \lim_{R \rightarrow \infty} \frac{1}{\bar{\lambda}(R)} \int_0^R dr a^2(\eta_0 - r) \frac{\xi_\phi(r)}{\xi_\phi(0)} \approx 0.15. \quad (79)$$

As shown in Fig. 18, the numerical results do follow the theoretical prediction of Eq. (77) within the error bars, although the unconstrained prediction $\langle \delta\lambda(\eta) \rangle / \bar{\lambda} = -2\phi_0$ at first order would also be validated by the data. In particular, $\langle \delta\lambda \rangle / \bar{\lambda}$ is indeed observed to converge to a constant at high z . For the deep cone ($z > 2$) the error bars are dominated by super-sample variance, but it seems that by chance the mean gravitational potential in these 400 deg^2 is very close to zero.

¹⁴ This is known as the Finger-of-God effect in galaxy clustering and redshift-space distortions analysis.

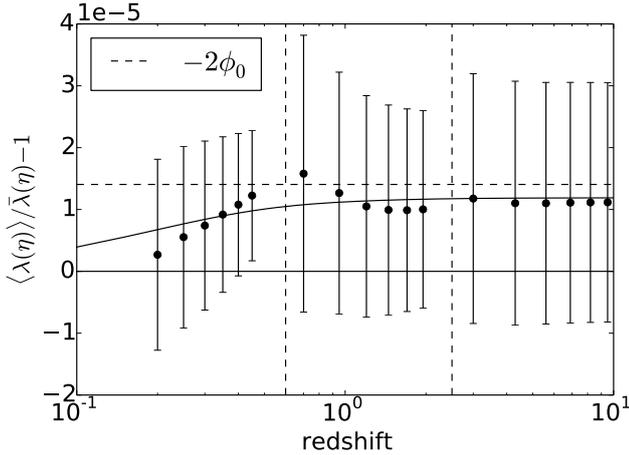


Fig. 18. Perturbation of the affine parameter at fixed time η . Dots are numerical results obtained by directional averaging; the solid line indicates the theoretical prediction (77); the dashed line shows $-2\phi_0$. Hence, the fractional difference between the dashed line and the asymptote of the solid line indicates Ξ_∞ . Error bars account for Poisson variance and constrained super-sample variance; see Appendix D.4 for details.

We stress that error bars of Fig. 18 were computed in configuration space to account for the constraint $\phi(0) = \phi_0$ on the variance (see Appendix D.4). For comparison, we also computed the variance in harmonic space using Eq. (63). Results were similar for the intermediate and deep cones, while for the full-sky cone the unconstrained error bars were twice larger than the constrained one. Had we disposed of full-sky data up to $z = 10$, the error bars at high- z would have been smaller than the bias of $\langle \lambda(\eta) \rangle$. In other words, $\langle \delta\lambda(\eta) \rangle = 0$ would have been excluded.

4.2.5. Spatio-temporal shift at constant affine parameter

We now consider the time and radial shifts $\delta\eta(\lambda) = -\delta r(\lambda)$ at fixed affine parameter, which is the converse of the above. For that operation, one would typically adopt a linearised approach and write $\delta\eta = (d\bar{\eta}/d\lambda)\delta\lambda$. However, since $d\bar{\eta}/d\lambda = 1/a^2 = (1 + \bar{z})^2$, this derivative can become quite large as the redshift increases, so that $\delta\lambda$ may actually get out of the linear behaviour of $\bar{\eta}(\lambda)$. We may thus adopt a more accurate inversion that could be extrapolated up to the LSS. Specifically, we use

$$\eta(\lambda) = \bar{\eta} \left\{ \bar{\lambda}[\eta(\lambda)] \right\}, \quad (80)$$

$$= \bar{\eta} \left\{ \lambda[\eta(\lambda)] - \delta\lambda[\eta(\lambda)] \right\}, \quad (81)$$

$$\approx \bar{\eta} \left\{ \lambda - \delta\lambda[\bar{\eta}(\lambda)] \right\}, \quad (82)$$

which we shall refrain from further expanding.

We have checked numerically that the fluctuations of $\lambda(\eta)$ about its mean are much smaller than its bias $\langle \delta\lambda(\eta) \rangle$. This allows us to express the average emission time at λ as

$$\langle \eta(\lambda) \rangle \approx \bar{\eta} \left\{ \lambda - \langle \delta\lambda[\bar{\eta}(\lambda)] \rangle \right\} \equiv \bar{\eta}(\lambda) + \langle \delta\eta(\lambda) \rangle, \quad (83)$$

where $\langle \delta\lambda \rangle$ is given by Eq. (77).

As shown in Fig. 19, the prediction (83) interpreted as a radial shift $\langle \delta r(\lambda) \rangle = -\langle \delta\eta(\lambda) \rangle$ is in excellent agreement with the simulation data $\langle \delta r(\lambda) \rangle_d$, although again 0 remains within the error bars dominated by super-sample variance at high- z . As expected, the amplitude of the shift blows up as λ increases.

In fact, a linear expansion of Eq. (77) would also match our numerical results, whose redshift range is not large enough

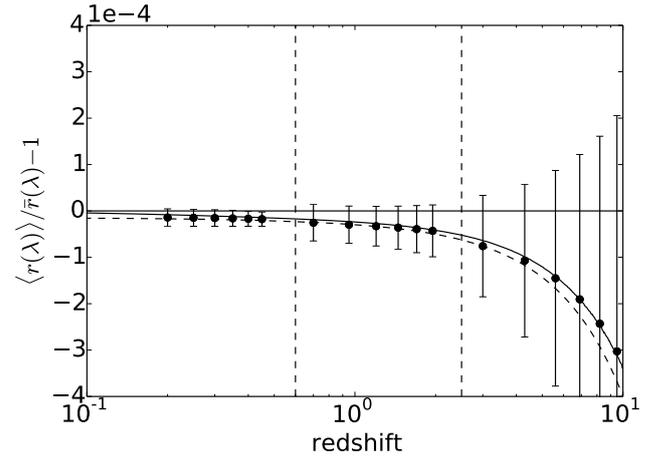


Fig. 19. Radial (or equivalently temporal) shift $\langle \delta r(\lambda) \rangle = -\langle \delta\eta(\lambda) \rangle$ at fixed affine parameter λ . Dots indicate numerical results from directional averaging; the solid line shows the theoretical prediction (83), while the dashed line indicates $\bar{r}[(1 - 2\phi_0)\lambda] - \bar{r}(\lambda)$ for comparison. Error bars account for Poisson and super-sample uncertainties; see Appendix D.4.

for the non-linear prescription to be critical. However, it does matter for the surface $\lambda = \lambda_*$, that is the surface of constant affine parameter dictated by the background affine parameter of the LSS. In that case the linearised version of Eq. (77) would highly over-estimate the mean shift, which is already very large, $\langle \delta r(\lambda_*) \rangle / \bar{r}(\lambda_*) \sim 10\%$ with $\phi_0 = -2 \text{ km s}^{-1}$ of our simulation.

We now need to investigate the consequences of such a large shift on the area $A(\lambda_*)$ of the surface of affine parameter λ_* . Just as for iso- z surfaces, the time and radial shifts of iso- λ surfaces both contribute to the perturbation of their area, or equivalently to $\langle \bar{\mu}^{-1}(\lambda) \rangle$. Accounting for that shift only (the tilt being second-order, it would be sub-dominant), we have

$$\frac{\delta A(\lambda_*)}{\bar{A}(\lambda_*)} = \left\langle \frac{\delta \left\{ a^2[\eta(\lambda_*)] r^2(\lambda_*) \right\}}{a^2[\bar{\eta}(\lambda_*)] \bar{r}^2(\lambda_*)} \right\rangle \approx \frac{a^2[\langle \eta(\lambda_*) \rangle] - a^2[\bar{\eta}(\lambda_*)]}{a^2[\bar{\eta}(\lambda_*)]}, \quad (84)$$

which is dominated by the difference in scale factor¹⁵. Again, we shall refrain from linearising Eq. (84) because the time shift is large enough to invalidate a first-order Taylor expansion of $a(\eta)$.

We illustrate the behaviour of Eq. (84) in Fig. 20 (black solid line) as a function of the observer's potential. Even for reasonable values of that quantity, such as $\phi_0 \sim -10 \text{ km s}^{-1}$, which is the order of magnitude expected within galaxy clusters (Wojtak et al. 2011, 2015), we find an astonishing $\delta A(\lambda_*) / \bar{A}(\lambda_*) \sim 3000\%$. We note that the black solid line of Fig. 20 is only plotted for negative values of ϕ_0 . This is because for $\phi_0 > 0$, $\langle \delta\eta(\lambda_*) \rangle < 0$ quickly diverges as ϕ_0 increases; since $\lim_{\eta \rightarrow -\infty} \bar{\lambda}(\eta) < \infty$, it may not even be defined.

At that point, we may note that the affine parameter is defined up to an arbitrary normalisation. Here, this normalisation has been chosen so that $k_0^0 = 1$ at the observer, hence λ coincides with cosmic time at that point. Another sensible choice consists in setting the observed frequency to unity, $\omega_o = (1 + 2\phi_0)k_0^0 = 1$. This corresponds to a re-normalisation $\lambda \mapsto \lambda' = (1 + 2\phi_0)\lambda$, where now λ' represents proper time at the observer. This alternative normalisation eliminates the ϕ_0 term in Eq. (76) and yields $\langle \delta\lambda'(\eta) \rangle / \bar{\lambda}(\eta) = 2\Xi_\infty\phi_0$ instead of Eq. (78). However, the bias on the area $A(\lambda'_*)$ remains huge, as seen with the red line of

¹⁵ This is because the surface $\lambda = \lambda_*$ is very far from the observer. For closer surfaces the radial shift would dominate.

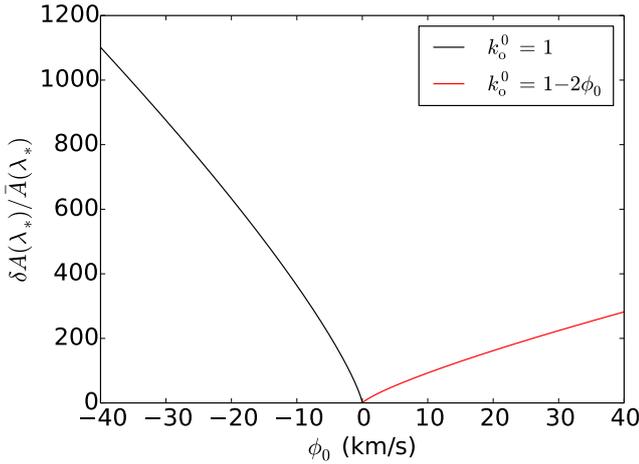


Fig. 20. Fractional change of the area $A(\lambda_*)$ of the surface of constant affine parameter $\lambda = \lambda_*$, as a function of the gravitational potential ϕ_0 at the observer. The black and red lines correspond to two different normalisations for the affine parameter.

Fig. 20¹⁶, which corresponds to that alternative normalisation for the affine parameter. The only way to avoid such a huge bias would be to define λ'' with the quite un-natural normalisation $k_o^0 = 1 - 2(1 - \Xi_\infty)\phi_0$.

The perhaps surprising results reported in this section must be attributed to the fact that $d\bar{\lambda}/d\eta = a^2(\eta) \rightarrow 0$ when $\eta \rightarrow -\infty$, so that $\bar{\lambda}(\eta)$ flattens out and saturates in the early Universe. This makes η extremely sensitive to even tiny changes in λ . All in all, this indicates that the affine parameter is a rather dangerous quantity to be used in theoretical analyses of light propagation down to the early Universe. The huge bias on $A(\lambda)$ shown in Fig. 20 must be considered a theoretical hiccup with no observational consequences.

4.3. Focus on the tilt correction

We may close this section on numerical results by evaluating the tilt, or wrinkly-surface effect, which tends to increase the inverse amplification $1/\bar{\mu}$ and the area of light-cone slices. As outlined in Sect. 2.2.3 and illustrated in Fig. 1, the tilt is defined as the angle ι between the radial direction β of a point and the local direction of light propagation at that point.

The contribution of the tilt to the area $A(p)$ of an iso- p surface, or equivalently to the directional average of the inverse amplification $\langle \bar{\mu}^{-1}(p) \rangle_d$, where p is any relevant parameter, goes as $\iota^2(p)$; it is always a second-order quantity. As such, it does not matter much which parameter p is actually considered here. Indeed, for any two parameters p, q (which can stand for z, η, λ, \dots) we have $\iota(p)/\iota(q) - 1 \sim |\text{shift}(p) - \text{shift}(q)| \ll 1$. In other words, the tilt can be considered universal in first approximation.

In light of the above discussion, we shall consider the tilt over surfaces of constant time, keeping in mind that the result would equally apply to other ways to slice the light cone. The choice of iso- η surfaces is also motivated by the fact that, in this case¹⁷, ι coincides with the angle formed by the normal \mathbf{n} to the

¹⁶ This red line is only plotted for $\phi_0 \geq 0$ for the same reason as the black line is only plotted for $\phi_0 \leq 0$. Namely, $\langle \delta\eta(\lambda_*) \rangle < 0$ quickly diverges and may not even exist when $\phi_0 < 0$.

¹⁷ This property would apply to other surfaces in their natural frame, which is not necessarily the comoving frame.

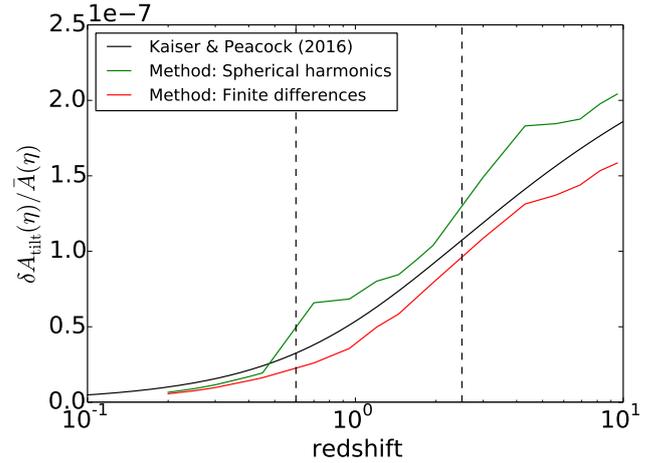


Fig. 21. Wrinkly-surface effect: relative increase of the area of surfaces of constant time due to their non-sphericity. The black solid line indicates the theoretical prediction (86) as originally found by KP16. The green and red lines indicate numerical results obtained from the two different methods outlined in Sect. 3.8.

surface and the radial direction β , as seen in Sect. 2.5.3. Thus,

$$\cos \iota(\eta, \beta) = \beta \cdot \mathbf{n}(\eta, \beta) = \left[1 + \left| \frac{1}{r(\eta, \beta)} \frac{\partial r}{\partial \beta} \right|^2 \right]^{-1/2}. \quad (85)$$

At second order, the area increase due to the tilt is predicted to be (see Appendix C.4 for details),

$$\frac{\delta A_{\text{tilt}}(\eta)}{\bar{A}(\eta)} = -\frac{1}{2} \langle \iota^2(\eta) \rangle = \frac{1}{\bar{r}^2(\eta)} \int_0^{\bar{r}(\eta)} dr r^2 J(r), \quad (86)$$

with $J(r)$ given in Eq. (49).

This prediction is successfully confronted with numerical results following Sect. 3.8 in Fig. 21. This ends our series of numerical checks of the remarkably accurate predictions of KP16.

5. Conclusion

The general topic of this article was the effect of inhomogeneities on cosmological observables, notably the average result of cosmic distance measurements. In the theoretical part (Sect. 2), we started by emphasising the subtle difference between the notions of geometric magnification μ and observable magnification $\bar{\mu}$. We interpreted that difference in terms of shifts and tilts associated with light propagation in the inhomogeneous Universe. We then reviewed and compared various notions of averaging involved in cosmology, in particular directional averaging $\langle \dots \rangle_d$ and source-averaging $\langle \dots \rangle_s$.

We argued that, because of the exact identity $\langle \mu^{-1} \rangle_d = 1$, one may expect the approximate relations $\langle \bar{\mu}^{-1} \rangle_d \approx \langle \bar{\mu} \rangle_s \approx 1$. Departures from the exact $\langle \bar{\mu}^{-1} \rangle_d = 1$ are due to the shift and tilt effects. We rigorously showed that such statements may be reformulated in terms of the area A of slices of the light cone; namely A should be almost unaffected by cosmological inhomogeneities, which is the conjecture of Weinberg (1976). Finally, we have reviewed how $\langle \bar{\mu}^{-1} \rangle_d \approx \langle \bar{\mu} \rangle_s \approx 1$ allows one to predict the statistical bias of any distance measure, such as the magnitude in SN surveys.

Most of the above theoretical considerations had been investigated in the past, notably in KP16, within the framework of cosmological perturbation theory at second order. The main addition

of this article is their thorough analysis via ray tracing in a high-resolution N -body simulation up to $z = 10$. This tool (Sect. 3) allowed us to account for the inhomogeneity of the Universe down to very small scales, where the perturbation theory fails. We produced HEALPIX maps and halo catalogues to generate direction-averaged and source-averaged mock observations respectively. Our main results are the following:

- (i) At all redshifts, we confirmed that $\langle \mu^{-1}(z) \rangle_d \approx \langle \tilde{\mu}^{-1}(z) \rangle_d \approx \langle \mu(z) \rangle_s \approx \langle \tilde{\mu}(z) \rangle_s \approx 1$ within our error bars, which account for both Poisson and super-sample variance. Source-averaged quantities were found to be more biased due to real-space clustering.
- (ii) Still within error bars, the bias on the direction-averaged distance is $\langle \delta d(z) \rangle_d = -\langle \kappa^2 \rangle / 2$, while the source-averaged distance and distance modulus are biased as $\langle \delta d(z) \rangle_s = 3\langle \kappa^2 \rangle / 2$ and $\langle \Delta m(z) \rangle_s = 5\langle \kappa^2 \rangle / \ln 10$. All these numerical results thus agree very well (within error bars) with theoretical predictions. A large scatter is observed at low redshift due to peculiar velocities.
- (iii) In order to further test the accuracy of Weinberg's conjecture, we investigated in detail the discrepancy between geometric and observable magnifications $\mu, \tilde{\mu}$, that is, the effects of shift and tilt. We found that the fractional area perturbations of $\Sigma(\eta)$ are well recovered by the predictions of KP16, of the order of order 10^{-7} . However, this bias turned out to be much smaller than the super-sample variance on time delays, which is on the order of 10^{-5} . Furthermore, we checked that propagating light rays on a coarse grid (rather than the AMR grid) does not impact the discrepancy between μ and $\tilde{\mu}$, which shows that the latter is relatively insensitive to very small scales.

Summarising, our results show no unexpectedly large bias for source and direction-averaged observables. They also confirm Weinberg's conjecture that the area of surfaces of constant redshift, or constant time, are almost unaffected by inhomogeneities, with corrections remaining below a part in a million for the latter (while the former might be slightly larger at low z due to peculiar velocities, but so is the associated variance).

As a theoretical curiosity, we also considered the area bias of surfaces of constant affine parameter, $\Sigma(\lambda)$. We found that at very high redshift, the bias grows quickly and eventually diverges, because $\eta \mapsto \lambda(\eta)$ reaches a flat asymptote for $\eta \rightarrow -\infty$. In practice, this leads to absurdly large corrections to $A(\lambda)$, and shows that one should avoid the use of the affine parameter to describe light rays in the early Universe.

For the present analysis we used finite ray bundles to compute the lensing distortion matrix \mathcal{A} . This allowed us, as a side product, to test the predictions of the finite-beam formalism developed by Fleury et al. (2017b, 2019a,b). Specifically, we checked that the convergence and shear power spectrum for finite ray bundles were suppressed for scales smaller than the bundle's width, in excellent agreement with Fleury et al. (2019a). Last, we found that treating the gravitational potential field ϕ as a constrained field near the observer, improves the agreement between numerical data and theoretical predictions.

Several extensions are possible for this work: at very large scale it could be interesting to see the impact of a fully general relativistic treatment, either by correcting the results from a Newtonian N -body code (Chisari & Zaldarriaga 2011; Fidler et al. 2015) or by directly using GR simulations (Adamek et al. 2016; Barrera-Hinojosa & Li 2020). At smaller scales, one could investigate the effect of strong lensing, allowing for multiple images for a single source. Also, we studied the bias on the distance-redshift relation within the Λ CDM framework. One could perform a similar analysis using different cosmologies, for

example by changing the nature of the dark sector or departing from GR, to see how such new physics may be degenerate with observational biases.

Acknowledgements. This paper is the continuation of a work that started during Vincent Reverdy's PhD thesis (Reverdy 2014). MAB thanks Sylvain de la Torre for pointing out the paper of Desjacques et al. (2021). We thank the referee John Peacock for many relevant comments which significantly improved the quality of this manuscript, especially in Sect. 2. This work was granted access to HPC resources of TGCC through allocations made by GENCI (Grand Equipement National de Calcul Intensif) under the allocations A0050402287 and A0070402287. PF received the support of a fellowship from "la Caixa" Foundation (ID 100010434). The fellowship code is LCF/BQ/PI19/11690018.

References

- Abbott, T., Allam, S., Andersen, P., et al. 2019, *ApJ*, **872**, L30
Adamek, J., Daverio, D., Durrer, R., & Kunz, M. 2016, *JCAP*, **7**, 053
Adamek, J., Clarkson, C., Coates, L., Durrer, R., & Kunz, M. 2019, *Phys. Rev. D*, **100**, 021301
Alam, S., Aubert, M., Avila, S., et al. 2021, *Phys. Rev. D*, **103**, 083533
Bacon, D. J., Andrianomena, S., Clarkson, C., Bolejko, K., & Maartens, R. 2014, *MNRAS*, **443**, 1900
Bagla, J. S., & Ray, S. 2005, *MNRAS*, **358**, 1076
Bardeen, J. M. 1980, *Phys. Rev. D*, **22**, 1882
Barreira, A., Llinares, C., Bose, S., & Li, B. 2016, *JCAP*, **5**, 001
Barrera-Hinojosa, C., & Li, B. 2020, *JCAP*, **01**, 007
Ben-Dayan, I., Marozzi, G., Nugier, F., & Veneziano, G. 2012, *JCAP*, **11**, 045
Ben-Dayan, I., Durrer, R., Marozzi, G., & Schwarz, D. J. 2014, *Phys. Rev. Lett.*, **112**, 221301
Bentivegna, E., Korzyński, M., Hinder, I., & Gerlicher, D. 2017, *JCAP*, **03**, 014
Biswas, T., & Notari, A. 2008, *JCAP*, **0806**, 021
Blandford, R., & Narayan, R. 1986, *ApJ*, **310**, 568
Bolejko, K. 2009, *Gen. Rel. Grav.*, **41**, 1737
Bolejko, K. 2011, *JCAP*, **2**, 25
Bolejko, K., & Célérier, M.-N. 2010, *Phys. Rev. D*, **82**, 103510
Bolejko, K., Clarkson, C., Maartens, R., et al. 2013, *Phys. Rev. Lett.*, **110**, 021302
Bonvin, C. 2008, *Phys. Rev. D*, **78**, 123530
Bonvin, C., Durrer, R., & Gasparini, M. A. 2006, *Phys. Rev. D*, **73**, 023523
Bonvin, C., Clarkson, C., Durrer, R., Maartens, R., & Umeh, O. 2015a, *JCAP*, **06**, 050
Bonvin, C., Clarkson, C., Durrer, R., Maartens, R., & Umeh, O. 2015b, *JCAP*, **07**, 040
Breton, M.-A., Rasera, Y., Taruya, A., Lacombe, O., & Saga, S. 2019, *MNRAS*, **483**, 2671
Brouzakis, N., Tetradis, N., & Tzavara, E. 2007, *JCAP*, **2**, 13
Brouzakis, N., Tetradis, N., & Tzavara, E. 2008, *JCAP*, **0804**, 008
Bruneton, J.-P., & Larena, J. 2013, *Class. Quant. Grav.*, **30**, 025002
Burke, W. L. 1981, *ApJ*, **244**, L1
Caprini, C., & Tamanini, N. 2016, *JCAP*, **10**, 006
Chisari, N. E., & Zaldarriaga, M. 2011, *Phys. Rev. D*, **83**, 123505
Chon, G., Challinor, A., Prunet, S., Hivon, E., & Szapudi, I. 2004, *MNRAS*, **350**, 914
Clarkson, C., Umeh, O., Maartens, R., & Durrer, R. 2014, *JCAP*, **11**, 036
Clifton, T., & Ferreira, P. G. 2009a, *Phys. Rev. D*, **80**, 103503
Clifton, T., & Ferreira, P. G. 2009b, *JCAP*, **10**, 026
Clifton, T., & Ferreira, P. G. 2011, *Phys. Rev. D*, **84**, 109902
Clifton, T., & Zuntz, J. 2009, *MNRAS*, **400**, 2185
Clifton, T., Ferreira, P. G., & O'Donnell, K. 2012, *Phys. Rev. D*, **85**, 023502
Corasaniti, P. S., Ettori, S., Rasera, Y., et al. 2018, *ApJ*, **862**, 40
Dashevskii, V. M., & Slysh, V. I. 1966, *Sov. Astron.*, **9**, 671
Davis, T. M., Hui, L., Frieman, J. A., et al. 2011, *ApJ*, **741**, 67
Desjacques, V., Ginat, Y. B., & Reischke, R. 2021, *MNRAS*, **504**, 5612
Di Dio, E., Vonlanthen, M., & Durrer, R. 2012, *JCAP*, **02**, 036
Dyer, C. C., & Roeder, R. C. 1974, *ApJ*, **189**, 167
Ellis, G. F. R., & Durrer, R. 2018, ArXiv e-prints [arXiv:1806.09530]
Ellis, G. F. R., Bassett, B. A. C. C., & Dunsby, P. K. S. 1998, *Class. Quant. Grav.*, **15**, 2345
Etherington, I. M. H. 1933, *Phil. Mag.*, **15**, 761
Fanizza, G., Gasperini, M., Marozzi, G., & Veneziano, G. 2020, *JCAP*, **02**, 017
Fidler, C., Rampf, C., Tram, T., et al. 2015, *Phys. Rev. D*, **92**, 123517
Fidler, C., Tram, T., Rampf, C., et al. 2016, *JCAP*, **9**, 031
Flanagan, E. E., Kumar, N., & Wasserman, I. 2013, *Phys. Rev. D*, **88**, 043004
Fleury, P. 2014, *JCAP*, **06**, 054
Fleury, P. 2015, PhD Thesis, Paris U., VI, IAP

- Fleury, P., & García-Bellido, J. 2020, *Phys. Dark Univ.*, **29**, 100567
- Fleury, P., Dupuy, H., & Uzan, J.-P. 2013, *Phys. Rev. D*, **87**, 123526
- Fleury, P., Clarkson, C., & Maartens, R. 2017a, *JCAP*, **3**, 062
- Fleury, P., Larena, J., & Uzan, J.-P. 2017b, *Phys. Rev. Lett.*, **119**, 191101
- Fleury, P., Larena, J., & Uzan, J.-P. 2019a, *Phys. Rev. D*, **99**, 023525
- Fleury, P., Larena, J., & Uzan, J.-P. 2019b, *Phys. Rev. D*, **99**, 023526
- Fluke, C. J., & Lasky, P. D. 2011, *MNRAS*, **416**, 1616
- Fluke, C. J., Webster, R. L., & Mortlock, D. J. 1999, *MNRAS*, **306**, 567
- Fosalba, P., Gaztañaga, E., Castander, F. J., & Manera, M. 2008, *MNRAS*, **391**, 435
- Freedman, W. L., Madore, B. F., Hatt, D., et al. 2019, *ApJ*, **882**, 34
- Gelb, J. M., & Bertschinger, E. 1994, *ApJ*, **436**, 491
- Giblin, J. T., Jr., Mertens, J. B., & Starkman, G. D. 2016, *ApJ*, **833**, 247
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, **622**, 759
- Green, S. R., & Wald, R. M. 2014, *Class. Quant. Grav.*, **31**, 234003
- Guillet, T., & Teyssier, R. 2011, *J. Comput. Phys.*, **230**, 4756
- Gunn, J. E. 1967, *ApJ*, **150**, 737
- Hall, A. 2020, *Phys. Rev. D*, **101**, 043519
- Harnois-Déraps, J., & van Waerbeke, L. 2015, *MNRAS*, **450**, 2857
- Helbig, P. 2020, *Open J. Astrophys.*, **3**, 1
- Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2009, *A&A*, **499**, 31
- Hockney, R. W., & Eastwood, J. W. 1981, *Computer Simulation Using Particles* (New York: McGraw-Hill)
- Hoffman, Y., & Ribak, E. 1991, *ApJ*, **380**, L5
- Holz, D. E., & Hughes, S. A. 2005, *ApJ*, **629**, 15
- Hui, L., & Greene, P. B. 2006, *Phys. Rev. D*, **73**, 123526
- Jain, B., Seljak, U., & White, S. 2000, *ApJ*, **530**, 547
- Kaiser, N. 1987, *MNRAS*, **227**, 1
- Kaiser, N., & Hudson, M. J. 2015, *MNRAS*, **454**, 280
- Kaiser, N., & Peacock, J. A. 2016, *MNRAS*, **455**, 4518
- Kantowski, R. 1969, *ApJ*, **155**, 89
- Kibble, T. W. B., & Lieu, R. 2005, *ApJ*, **632**, 718
- Kilbinger, M., Heymans, C., Asgari, M., et al. 2017, *MNRAS*, **472**, 2126
- Kocsbang, S. M. 2017, *Phys. Rev. D*, **95**, 063532
- Kocsbang, S. M. 2019a, *Class. Quant. Grav.*, **36**, 185004
- Kocsbang, S. M. 2019b, *Phys. Rev. D*, **100**, 063533
- Kocsbang, S. 2020a, *JCAP*, **11**, 061
- Kocsbang, S. M. 2020b, *MNRAS*, **498**, L135
- Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, *ApJS*, **192**, 18
- Lavinto, M., & Rasanen, S. 2015, *JCAP*, **10**, 057
- Lepori, F., Adamek, J., Durrer, R., Clarkson, C., & Coates, L. 2020, *MNRAS*, **497**, 2078
- Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, **538**, 473
- Liu, R. G. 2015, *Phys. Rev. D*, **92**, 063529
- Marra, V., Kolb, E. W., & Matarrese, S. 2008, *Phys. Rev. D*, **77**, 023003
- Mitsou, E., Yoo, J., Durrer, R., Scaccabarozzi, F., & Tansella, V. 2020, *Phys. Rev. Res.*, **2**, 033004
- Odderskov, I., Kocsbang, S. M., & Hannestad, S. 2016, *JCAP*, **2**, 001
- Okamura, T., & Futamase, T. 2009, *Prog. Theor. Phys.*, **122**, 511
- Peel, A., Troxel, M. A., & Ishak, M. 2014, *Phys. Rev. D*, **90**, 123536
- Pen, U.-L. 1997, *ApJ*, **490**, L127
- Perlmutter, S., Aldering, G., della Valle, M., et al. 1998, *Nature*, **391**, 51
- Peter, P., & Uzan, J.-P. 2013, *Primordial Cosmology. Oxford Graduate Texts* (Oxford: Oxford University Press)
- Planck Collaboration VI. 2020, *A&A*, **641**, A6
- Power, C., & Knebe, A. 2006, *MNRAS*, **370**, 691
- Prunet, S., Pichon, C., Aubert, D., et al. 2008, *ApJS*, **178**, 179
- Reverdy, V. 2014, PhD Thesis, Laboratoire Univers et Théories
- Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *AJ*, **116**, 1009
- Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019, *ApJ*, **876**, 85
- Roy, F., Bouillot, V. R., & Rasera, Y. 2014, *A&A*, **564**, A13
- Sachs, R. 1961, *Proc. R. Soc. London Ser. A*, **264**, 309
- Sanghai, V. A. A., Fleury, P., & Clifton, T. 2017, *JCAP*, **7**, 028
- Sasaki, M. 1987, *MNRAS*, **228**, 653
- Schneider, P., Ehlers, J., & Falco, E. E. 1992, *Gravitational Lenses* (Berlin: Springer-Verlag)
- Scoccamarro, R. 1998, *MNRAS*, **299**, 1097
- Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, *ApJ*, **859**, 101
- Sheth, R. K., & Tormen, G. 1999, *MNRAS*, **308**, 119
- Sirko, E. 2005, *ApJ*, **634**, 728
- Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, *MNRAS*, **341**, 1311
- Szapudi, I., Prunet, S., Pogosyan, D., Szalay, A. S., & Bond, J. R. 2001, *ApJ*, **548**, L115
- Szybka, S. J. 2011, *Phys. Rev. D*, **84**, 044011
- Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, *ApJ*, **761**, 152
- Teyssier, R. 2002, *A&A*, **385**, 337
- Teyssier, R., Pires, S., Prunet, S., et al. 2009, *A&A*, **497**, 335
- Troxel, M. A., Ishak, M., & Peel, A. 2014, *JCAP*, **3**, 40
- Umeh, O., Clarkson, C., & Maartens, R. 2014, *Class. Quant. Grav.*, **31**, 202001
- Valkenburg, W. 2009, *JCAP*, **0906**, 010
- van de Weygaert, R., & Bertschinger, E. 1996, *MNRAS*, **281**, 84
- Vanderveld, R. A., Flanagan, E. E., & Wasserman, I. 2008, *Phys. Rev. D*, **78**, 083511
- Weinberg, S. 1976, *ApJ*, **208**, L1
- Wojtak, R., Hansen, S. H., & Hjorth, J. 2011, *Nature*, **477**, 567
- Wojtak, R., Davis, T. M., & Wiis, J. 2015, *JCAP*, **07**, 025
- Wong, K. C., Suyu, S. H., Chen, G. C. F., et al. 2019, *MNRAS*, **498**, 1420
- Wucknitz, O. 2008, *MNRAS*, **386**, 230
- Yoo, J., & Scaccabarozzi, F. 2016, *JCAP*, **9**, 046
- Zel'dovich, Y. B. 1964, *Sov. Astron.*, **8**, 13
- Zonca, A., Singer, L., Lenz, D., et al. 2019, *J. Open Sour. Softw.*, **4**, 1298

Appendix A: Alternative approach to the shift and tilt corrections

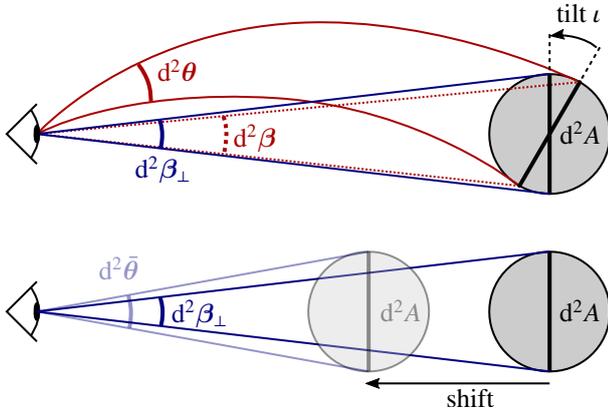


Fig. A.1. Illustration of the shift and tilt effects, implying that the observable magnification differs from the geometric magnification, in the case of an isotropic spherical source.

In Sect. 2.2.3 we have presented the shift and tilt corrections to the magnification by assuming that the source was an infinitesimal element of the iso- z surface. This choice of formulation was justified by its tight connection with the subsequent considerations on the area of light-cone slices. Having said that, the curious reader may wonder about the generality of the shift and tilt corrections as defined therein. We may now consider an individual source instead of an element of iso- z surface. This appendix aims to show that the results of Sect. 2.2.3 equally applies to this case.

We assume, for simplicity, that the source at redshift z is spherical with isotropic emission¹⁸. Recall that the geometric and observational magnifications are respectively defined as

$$\mu = \pm \frac{d^2\theta}{d^2\beta}, \quad (\text{A.1})$$

$$\tilde{\mu} = \pm \frac{d^2\theta}{d^2\bar{\theta}}, \quad (\text{A.2})$$

where $d^2\theta$ is the angular size of an image, $d^2\beta$ the coordinate solid angle covered by the source, and $d^2\bar{\theta}$ the apparent size of that source if it were placed at the same redshift in an FLRW universe (see Fig. A.1). The difference between μ and $\tilde{\mu}$ lies in the subtle difference between $d^2\beta$ and $d^2\bar{\theta}$.

Since the source is spherical, it is equivalent to a disk of area d^2A orthogonal to the direction of light propagation at emission. However, due to light deflection, this disk is tilted with respect to the radial direction. As a consequence, the disk covers a smaller coordinate solid angle $d^2\beta$ than its non-tilted counterpart $d^2\beta_\perp$ (top of Fig. A.1). Both are related by $d^2\beta = d^2\beta_\perp \cos \iota$.

Besides, due to redshift corrections in the inhomogeneous Universe (peculiar velocities, Sachs-Wolfe effect, etc.) the time and radial coordinates of the source are shifted with respect to their background counterparts. As a result, a source with a given size is seen under a different solid angle in both cases (bottom of Fig. A.1), namely

$$\frac{d^2\bar{\theta}}{d^2\beta_\perp} = \frac{a^2[\eta(z)] r^2(z)}{a^2[\bar{\eta}(z)] \bar{r}^2(z)}. \quad (\text{A.3})$$

¹⁸ Anisotropic emission would bring additional effects that go beyond the problematic of this article.

Summarising, the geometric and observational magnifications are related by

$$\frac{\tilde{\mu}(z, \theta)}{\mu(z, \theta)} = \frac{d^2\beta}{d^2\bar{\theta}} = \underbrace{\frac{d^2\beta_\perp}{d^2\bar{\theta}}}_{\text{shift}} \underbrace{\frac{d^2\beta}{d^2\beta_\perp}}_{\text{tilt}} = \frac{a^2[\bar{\eta}(z)] \bar{r}^2(z)}{a^2[\eta(z)] r^2(z)} \cos \iota, \quad (\text{A.4})$$

which is indeed equivalent to Eq. (14).

Appendix B: Finite-beam corrections

In this appendix we derive the finite-beam corrections (52), (53) to the power spectra of convergence and shear. The computation will follow the general philosophy of Fleury et al. (2017b, 2019a,b), but it will differ in the details, due to the specific four-ray set-up used to compute κ and γ in this article.

B.1. Estimators of convergence and shear

As shown by Fleury et al. (2019a), finite-beam corrections to cosmic convergence and shear occur on very small scales. Thus, we can safely work in the flat-sky approximation in the following. In that context, the lens equation reads

$$\beta = \theta - \alpha(\theta), \quad (\text{B.1})$$

where β is the position of a point source, θ the position of its image, and α the displacement angle. An infinitesimal image is a collection of points θ whose separation is much smaller than the typical angular scale over which $\alpha(\theta)$ varies appreciably. In that case, one Taylor-expands $\alpha(\theta)$ at first order and gets

$$\alpha(\theta) = \alpha(\mathbf{0}) + \begin{pmatrix} \kappa + \gamma_1 & \gamma_2 \\ \gamma_2 & \kappa - \gamma_1 \end{pmatrix} \theta, \quad (\text{B.2})$$

which defines the convergence κ and shear γ . We note that we neglected the rotation $\omega \sim |\gamma|^2$ [see Fleury (2015)] for simplicity. Since $\alpha(\mathbf{0})$ could be absorbed in a re-definition of the origin of the source plane, we set it to zero for convenience, $\alpha(\mathbf{0}) = \mathbf{0}$.

In the following calculation, it will be very convenient to associate a complex number $\underline{\theta} = \theta_x + i\theta_y$ with any 2-dimensional vector $\theta = \theta_x \mathbf{e}_x + \theta_y \mathbf{e}_y$. The lens equation then reads $\underline{\beta} = \underline{\theta} - \underline{\alpha}(\underline{\theta})$, and for infinitesimal images

$$\underline{\alpha}(\underline{\theta}) = \kappa \underline{\theta} + \gamma \underline{\theta}^*, \quad (\text{B.3})$$

where a star denotes complex conjugation, and $\gamma = \gamma_1 + i\gamma_2$ is the complex shear.

In this article, as explained in Sect. 3.5, the distortion matrix, and hence convergence and shear, are estimated using a four-ray-bundle method. About a direction θ , four rays are shot in the directions $\theta \pm \varepsilon \mathbf{e}_x$, $\theta \pm \varepsilon \mathbf{e}_y$, and traced to get the associated four source positions. With complex notations, the corresponding estimators of convergence and shear are found to read

$$\kappa(\underline{\theta}; \varepsilon) = \frac{1}{4\varepsilon^2} \sum_{p=0}^3 \Re \left[\varepsilon_p^* \underline{\alpha}(\underline{\theta} + \varepsilon_p) \right], \quad (\text{B.4})$$

$$\gamma(\underline{\theta}; \varepsilon) = \frac{1}{4\varepsilon^2} \sum_{p=0}^3 \varepsilon_p \underline{\alpha}(\underline{\theta} + \varepsilon_p), \quad (\text{B.5})$$

where

$$\varepsilon_p \equiv \varepsilon e^{ip\frac{\pi}{2}} = \varepsilon i^p \quad p = 0, 1, 2, 3, \quad (\text{B.6})$$

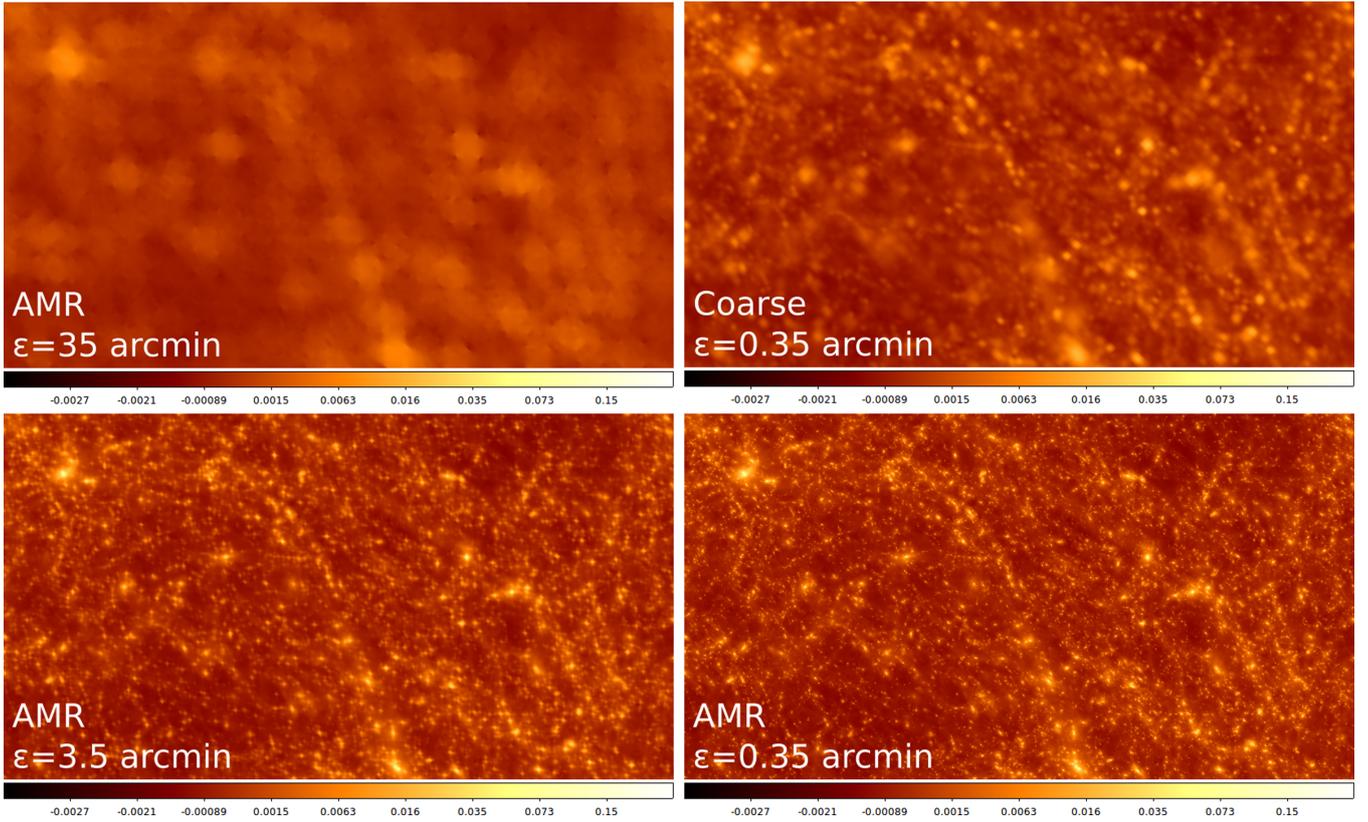


Fig. B.1. Convergence maps at $z = 0.2$ using HEALPIX with $n_{\text{side}} = 2048$. *Top right panel:* coarse grid and $\varepsilon = 0.35$ arcmin. *Top left panel:* AMR grid and $\varepsilon = 35$ arcmin. *Bottom left panel:* AMR grid and $\varepsilon = 3.5$ arcmin. *Bottom right panel:* AMR grid and $\varepsilon = 0.35$ arcmin.

are the four shifts with respect to the central ray θ used here.

We note that the infinitesimal-beam case is recovered as $\varepsilon \rightarrow 0$,

$$\lim_{\varepsilon \rightarrow 0} \kappa(\underline{\theta}; \varepsilon) = \Re(\underline{\partial}\alpha) = \underline{\partial}\alpha = \kappa(\underline{\theta}; 0), \quad (\text{B.7})$$

$$\lim_{\varepsilon \rightarrow 0} \gamma(\underline{\theta}; \varepsilon) = \underline{\partial}^* \alpha = \gamma(\underline{\theta}; 0), \quad (\text{B.8})$$

where we introduced the complex derivative

$$\underline{\partial} \equiv \frac{\partial}{\partial \underline{\theta}} = \frac{1}{2} \left(\frac{\partial}{\partial \theta_x} - i \frac{\partial}{\partial \theta_y} \right). \quad (\text{B.9})$$

B.2. Fourier transform

Before moving to the actual computation of the power spectra if κ , γ , it is useful to express their Fourier transforms, as a function of their infinitesimal-beam counterparts. We use the convention

$$\tilde{f}(\underline{\ell}) = \int d^2\theta e^{-i\underline{\ell}\cdot\theta} f(\theta), \quad (\text{B.10})$$

$$f(\theta) = \int \frac{d^2\ell}{(2\pi)^2} e^{i\underline{\ell}\cdot\theta} \tilde{f}(\underline{\ell}). \quad (\text{B.11})$$

In the infinitesimal beam case, this implies

$$\tilde{\kappa}(\underline{\ell}; 0) = \frac{1}{2} i \underline{\ell}^* \underline{\tilde{\alpha}}(\underline{\ell}), \quad (\text{B.12})$$

$$\tilde{\gamma}(\underline{\ell}; 0) = \frac{1}{2} i \underline{\ell} \underline{\tilde{\alpha}}(\underline{\ell}). \quad (\text{B.13})$$

We start with convergence. Taking the Fourier transform of Eq. (B.4), and using $\tilde{\kappa}(-\underline{\ell}; 0) = [\tilde{\kappa}(\underline{\ell}; 0)]^*$, we find

$$\tilde{\kappa}(\underline{\ell}; \varepsilon) = C(\underline{\ell}; \varepsilon) \tilde{\kappa}(\underline{\ell}; 0), \quad (\text{B.14})$$

with the finite-beam filter

$$C(\underline{\ell}; \varepsilon) \equiv \frac{1}{2\varepsilon^2} \sum_{p=0}^3 \Im \left[\frac{\varepsilon_p}{\underline{\ell}} \right] e^{i\varepsilon_p \underline{\ell}}. \quad (\text{B.15})$$

Similarly, the Fourier transform of shear reads

$$\tilde{\gamma}(\underline{\ell}; \varepsilon) = S(\underline{\ell}; \varepsilon) \tilde{\gamma}(\underline{\ell}; 0), \quad (\text{B.16})$$

with the finite-beam filter

$$S(\underline{\ell}; \varepsilon) = \frac{1}{2i\varepsilon^2} \sum_{p=0}^3 \frac{\varepsilon_p}{\underline{\ell}} e^{i\varepsilon_p \underline{\ell}}. \quad (\text{B.17})$$

B.3. Power spectra

We are now ready to compute the power spectra of $\kappa(\underline{\theta}; \varepsilon)$, $\gamma(\underline{\theta}; \varepsilon)$, and in particular to evaluate how the results of the four-ray set-up may differ from the theoretical predictions with infinitesimal beams. The convergence power spectrum can be defined via

$$\langle \tilde{\kappa}(\underline{\ell}_1; \varepsilon) \tilde{\kappa}(\underline{\ell}_2; \varepsilon) \rangle = (2\pi)^2 \delta_{\text{D}}(\underline{\ell}_1 + \underline{\ell}_2) P_{\kappa}(\underline{\ell}_1; \varepsilon), \quad (\text{B.18})$$

from which we deduce that

$$P_{\kappa}(\underline{\ell}; \varepsilon) = C(\underline{\ell}; \varepsilon) C(-\underline{\ell}; \varepsilon) P_{\kappa}(\underline{\ell}; 0). \quad (\text{B.19})$$

We note that, contrary to $P_{\kappa}(\underline{\ell}; 0)$, $P_{\kappa}(\underline{\ell}; \varepsilon)$ depends on the orientation of $\underline{\ell}$. This is due to the anisotropic square-like geometry of the four-beam set-up. However, in practice we effectively calculate the isotropic part from ray tracing,

$$P_{\kappa}(\underline{\ell}; \varepsilon) \equiv \int_0^{2\pi} \frac{d\psi}{2\pi} P_{\kappa}(\underline{\ell}; \varepsilon). \quad (\text{B.20})$$

where ψ denotes the polar angle of $\boldsymbol{\ell} = \ell(\cos \psi, \sin \psi)$. After a tedious but straightforward calculation, we finally get

$$\frac{P_\kappa(\boldsymbol{\ell}; \varepsilon)}{P_\kappa(\boldsymbol{\ell}; 0)} = \int_0^{2\pi} \frac{d\psi}{2\pi} C(\boldsymbol{\ell}; \varepsilon) C(-\boldsymbol{\ell}; \varepsilon), \quad (\text{B.21})$$

$$= \frac{1 + J_2(2\varepsilon\ell) + 2J_2(\sqrt{2}\varepsilon\ell) - J_0(2\varepsilon\ell)}{2(\varepsilon\ell)^2}. \quad (\text{B.22})$$

The calculation for shear is slightly different but technically simpler. The power spectrum can be defined via

$$\langle \tilde{\gamma}(\boldsymbol{\ell}_1; \varepsilon) \tilde{\gamma}^*(-\boldsymbol{\ell}_2; \varepsilon) \rangle = (2\pi)^2 \delta_{\text{D}}(\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2) P_\gamma(\boldsymbol{\ell}_1; \varepsilon), \quad (\text{B.23})$$

so that

$$P_\gamma(\boldsymbol{\ell}; \varepsilon) = |S(\boldsymbol{\ell}; \varepsilon)|^2 P_\gamma(\boldsymbol{\ell}; 0). \quad (\text{B.24})$$

Taking the isotropic part then yields

$$\frac{P_\gamma(\ell; \varepsilon)}{P_\gamma(\ell; 0)} = \int_0^{2\pi} \frac{d\psi}{2\pi} |S(\boldsymbol{\ell}; \varepsilon)|^2, \quad (\text{B.25})$$

$$= \frac{1 - J_0(2\varepsilon\ell)^2}{(\varepsilon\ell)^2}. \quad (\text{B.26})$$

As we can see in Fig. B.1, similarly to resolution effects on simulations (Lepori et al. 2020), the finite-beam effect acts as a smoothing on weak lensing maps.

Appendix C: Details on the area of constant-time surfaces: shift and tilt

In this appendix, we derive the theoretical predictions for the perturbation of the area of surfaces of constant time. As shown in Sect. 2.5.3, at second order, we have

$$\frac{\delta A(\eta)}{\bar{A}(\eta)} = \int_{\mathbb{S}^2} \frac{d^2\boldsymbol{\beta}}{4\pi} \left[\frac{2\delta r(\eta, \boldsymbol{\beta})}{\bar{r}(\eta)} + \frac{\delta r^2(\eta, \boldsymbol{\beta})}{\bar{r}^2(\eta)} + \frac{1}{2} \iota^2(\eta, \boldsymbol{\beta}) \right], \quad (\text{C.1})$$

where $\delta r(\eta, \boldsymbol{\beta}) = r(\eta, \boldsymbol{\beta}) - \bar{r}(\eta)$ is the radial perturbation of the iso- η surface, which may be decomposed into a geometrical and a time-delay (Shapiro) contribution as $\delta r = \delta r_{\text{geo}} + \delta s$; ι denotes the tilt between the normal iso- η surface and the radial direction.

We compute the ensemble average of δr_{geo} , δs , ι in the spirit of KP16. Although our final result (48) agrees with KP16, our derivations slightly differ. We also derive the angular correlation functions for δr_{geo} and δs , which is necessary to estimate their super-sample variance, and hence to allow a consistent confrontation of theory and numerical results.

C.1. Comoving radius reached at fixed comoving distance travelled

Due to gravitational lensing, light rays are wiggly, and hence the comoving radius that they reach is smaller than their comoving distance travelled. For a given comoving distance travelled s , the radius thus reads $r = \bar{r} + \delta r_{\text{geo}}(s)$ with $\bar{r} = s$. Here we carefully derive the expression of $\delta r_{\text{geo}} < 0$ and its statistical properties.

C.1.1. Expression of δr_{geo}

We consider a photon observed in the direction $\boldsymbol{\theta}$ from a source in the direction $\boldsymbol{\beta}$; we may write its trajectory as $\mathbf{x}(r) = r\boldsymbol{\beta} + \boldsymbol{\xi}(r)$, where $\boldsymbol{\xi} \perp \boldsymbol{\beta}$ is the transverse displacement with respect to the axis spanned by $\boldsymbol{\beta}$. If the source is located at r_s , then by definition

$\boldsymbol{\xi}(0) = \boldsymbol{\xi}(r_s) = \mathbf{0}$ and $\dot{\boldsymbol{\xi}}(0) = \boldsymbol{\theta} - \boldsymbol{\beta} \equiv \boldsymbol{\alpha}$ in the limit of small angles, where a dot denotes a derivative with respect to r .

With such conditions, the equation of motion $\ddot{\boldsymbol{\xi}} = -2\nabla_\perp \phi$ is solved as

$$\dot{\boldsymbol{\xi}}(r) = \boldsymbol{\alpha} - 2 \int_0^r dr' \nabla_\perp \phi(r'), \quad (\text{C.2})$$

$$\boldsymbol{\xi}(r) = r\boldsymbol{\alpha} - 2 \int_0^r dr' (r - r') \nabla_\perp \phi(r'). \quad (\text{C.3})$$

The condition $\boldsymbol{\xi}(r_s) = \mathbf{0}$ then implies the familiar

$$\boldsymbol{\alpha} = 2 \int_0^{r_s} dr \frac{r_s - r}{r_s} \nabla_\perp \phi(r). \quad (\text{C.4})$$

We note that the presence of $\boldsymbol{\alpha}$ in Eqs. (C.2) and (C.3) comes from our definition of $\boldsymbol{\xi}$ as the transverse displacement with respect to the $\boldsymbol{\beta}$ -axis; had we considered, as KP16, transverse displacement with respect to the $\boldsymbol{\theta}$ -axis, $\boldsymbol{\alpha}$ would not be present. Our convention is more adapted (i) to the geometry of the problem, and (ii) to the fact that we will eventually consider ensemble averages where $\boldsymbol{\beta}$ is fixed.

The total path length depends on the local inclination $\iota(r) = |\dot{\boldsymbol{\xi}}(r)|$ of the ray with respect to the axis spanned by $\boldsymbol{\beta}$. Specifically, we have, at second order,

$$s(r_s) = \int_0^{r_s} \frac{dr}{\cos \iota(r)} = r_s + \underbrace{\frac{1}{2} \int_0^{r_s} dr |\dot{\boldsymbol{\xi}}(r)|^2}_{-\delta r_{\text{geo}}(r_s)}, \quad (\text{C.5})$$

which defines the wiggly-ray correction $\delta r_{\text{geo}} < 0$ to the radius reached after travelling s ¹⁹. Substituting the expression (C.2) of $\dot{\boldsymbol{\xi}}$ then yields

$$\begin{aligned} \delta r_{\text{geo}} &\equiv -\frac{1}{2} \int_0^{r_s} dr |\dot{\boldsymbol{\xi}}(r)|^2, \quad (\text{C.6}) \\ &= -\frac{1}{2} r_s |\boldsymbol{\alpha}|^2 + 2\boldsymbol{\alpha} \cdot \int_0^{r_s} dr \int_0^r dr' \nabla_\perp \phi(r'), \\ &\quad - 2 \int_0^{r_s} dr \int_0^r dr' \int_0^r dr'' \nabla_\perp \phi(r') \cdot \nabla_\perp \phi(r''). \quad (\text{C.7}) \end{aligned}$$

We may now simplify this expression by playing with the order of integration. We recommend the reader to draw the various integration domains in order to make sense of the following operations. The integrals in the second term of Eq. (C.7) read

$$\int_0^{r_s} dr \int_0^r dr' \nabla_\perp \phi(r') = \int_0^{r_s} dr' \int_{r'}^{r_s} dr \nabla_\perp \phi(r'), \quad (\text{C.8})$$

$$= \int_0^{r_s} dr' (r_s - r') \nabla_\perp \phi(r'), \quad (\text{C.9})$$

$$= \frac{1}{2} r_s \boldsymbol{\alpha}, \quad (\text{C.10})$$

so that the whole second term is simply $r_s |\boldsymbol{\alpha}|^2$. As for the third term of Eq. (C.7), we may perform the following operations, aiming at moving the integration over r to the right. The first step is identical as above, and reads

$$\int_0^{r_s} dr \int_0^r dr' \int_0^r dr'' = \int_0^{r_s} dr' \int_{r'}^{r_s} dr \int_0^r dr''. \quad (\text{C.11})$$

¹⁹ The cautious reader may note that our definition of δr_{geo} differs from (A21) of KP16. Specifically, our approach involves the angle ι between the photon direction and $\boldsymbol{\beta}$, while KP16 consider instead the angle between photon propagation and the instantaneous photon position $\boldsymbol{\beta} + \boldsymbol{\xi}(r)/r$. Our methods yield the same final result for δr_{geo} .

The second inversion is more involved because it features an integration region made of right-angled triangle and a rectangle. A possible operation is

$$\int_{r'}^{r_s} dr \int_0^r dr'' = \int_0^{r'} dr'' \int_{r'}^{r_s} dr + \int_{r'}^{r_s} dr'' \int_{r'}^{r_s} dr, \quad (\text{C.12})$$

so that after a couple of additional manipulations we have

$$\begin{aligned} & \int_0^{r_s} dr \int_0^r dr' \int_0^r dr'' \nabla_{\perp} \phi(r') \cdot \nabla_{\perp} \phi(r'') \\ &= 2 \int_0^{r_s} dr' \int_0^{r'} dr'' (r_s - r') \nabla_{\perp} \phi(r') \cdot \nabla_{\perp} \phi(r''). \end{aligned} \quad (\text{C.13})$$

Gathering all three terms of Eq. (C.7) and using the symmetry of the double integration in $|\alpha|^2$, we finally obtain

$$\delta r_{\text{geo}} = -4 \int_0^{r_s} dr \int_0^r dr' \frac{(r_s - r)r'}{r_s} \nabla_{\perp} \phi(r) \cdot \nabla_{\perp} \phi(r'), \quad (\text{C.14})$$

in agreement with the third line of (A21) in KP16.

C.1.2. Ensemble average of δr_{geo}

For $\delta A(\eta)$ we need to evaluate the average of δr_{geo} over β . Following Sect. 2.3.3, we apply the ergodicity principle to translate this into an ensemble average,

$$\int_{\mathbb{S}^2} \frac{d^2 \beta}{4\pi} \delta r_{\text{geo}}(\eta, \beta) = \langle \delta r_{\text{geo}}(\eta, \beta) \rangle, \quad (\text{C.15})$$

which holds up to super-sample variance.

Introducing the Fourier transform $\tilde{\phi}$ of ϕ , we have

$$\nabla_{\perp} \phi = \int \frac{d^3 \mathbf{k}}{(2\pi)^3} e^{i\mathbf{k} \cdot \mathbf{x}} (i\mathbf{k}_{\perp}) \tilde{\phi}(\mathbf{k}), \quad (\text{C.16})$$

and using Limber's approximation we may proceed as

$$\begin{aligned} \langle \nabla_{\perp} \phi \cdot \nabla_{\perp} \phi' \rangle &= \int \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{d^3 \mathbf{k}'}{(2\pi)^3} e^{i(\mathbf{k} \cdot \mathbf{x} + \mathbf{k}' \cdot \mathbf{x}')} \\ &\quad \times (-\mathbf{k}_{\perp} \cdot \mathbf{k}'_{\perp}) \langle \tilde{\phi}(\mathbf{k}) \tilde{\phi}(\mathbf{k}') \rangle, \end{aligned} \quad (\text{C.17})$$

$$= \int \frac{d^3 \mathbf{k}}{(2\pi)^3} e^{i\mathbf{k} \cdot (\mathbf{x} - \mathbf{x}')} k_{\perp}^2 P_{\phi}(\eta, \eta', \mathbf{k}), \quad (\text{C.18})$$

$$\approx \delta_{\text{D}}(r - r') \underbrace{\int \frac{d^2 \ell}{(2\pi r)^2} \left(\frac{\ell}{r}\right)^2 P_{\phi}\left(\eta, \frac{\ell}{r}\right)}_{\equiv J(r)/2}, \quad (\text{C.19})$$

where we introduced the power spectrum P_{ϕ} of the gravitational potential, which is evaluated down the background light cone with $\eta = \eta_0 - r$. In the last line we adopted the notation of KP16 and identified the integral

$$J(r) \equiv \frac{2}{r^4} \int_0^{\infty} \frac{d\ell}{2\pi} \ell^3 P_{\phi}\left(\eta_0 - r, \frac{\ell}{r}\right), \quad (\text{C.20})$$

$$= 2 \int_0^{\infty} \frac{dk}{2\pi} k^3 P_{\phi}(\eta_0 - r, k), \quad (\text{C.21})$$

$$= 2\pi \int_0^{\infty} k \mathcal{P}_{\phi}(\eta_0 - r, k) d \ln k, \quad (\text{C.22})$$

where \mathcal{P}_{ϕ} denotes the dimensionless power spectrum of ϕ , which is related to the standard one by $2\pi^2 k^{-3} \mathcal{P}_{\phi}(k) = P_{\phi}(k)$.

Substituting the above in the expression of δr_{geo} , we find

$$\langle \delta r_{\text{geo}} \rangle = - \int_0^{r_s} dr \frac{(r_s - r)r}{r_s} J(r). \quad (\text{C.23})$$

Importantly, during that last step a factor 1/2 appears, due to the integration of a Dirac delta on half of its domain:

$$\int_0^r dr' \delta_{\text{D}}(r' - r) f(r') = \frac{1}{2} f(r). \quad (\text{C.24})$$

C.2. Distance travelled at fixed time

We now turn to the effect of time delays, which implies that during a time $\eta_0 - \eta$, a photon travels a comoving distance $s = \eta_0 - \eta + \delta s$, with

$$\delta s(\eta) = \int_{\eta}^{\eta_0} d\eta' 2\phi[\eta', \mathbf{x}(\eta')]. \quad (\text{C.25})$$

Combined with the wiggly-ray effect, this implies that the radius reached at fixed time reads $r(\eta) = \bar{r}(\eta) + \delta r_{\text{geo}}(\eta) + \delta s(\eta)$.

C.2.1. Post-Born expansion of δs

We note that δs is a first-order quantity evaluated on the perturbed trajectory \mathbf{x} of the photon. Since it is involved in δr together with δr_{geo} which is a second-order quantity, we must account for post-Born corrections for consistency. These will turn out to be exactly minus twice δr_{geo} .

Just as in Sect. C.1, we write the perturbed photon path as $\mathbf{x}(r) = r\beta + \xi(r)$, so that δs becomes

$$\delta s = \int_0^{r_s} dr 2\phi(r) + \underbrace{\int_0^{r_s} dr 2\xi(r) \cdot \nabla_{\perp} \phi(r)}_{\equiv \delta s_{\text{pB}}}, \quad (\text{C.26})$$

where we changed to an integration over comoving radius without any loss of generality. The second term, δs_{pB} , in Eq. (C.26) encodes post-Born corrections. Substituting the expression (C.3) of ξ we may rewrite it as

$$\begin{aligned} \delta s_{\text{pB}} &= 4 \int_0^{r_s} dr \int_0^r dr' \frac{r(r_s - r')}{r_s} \nabla_{\perp} \phi(r') \cdot \nabla_{\perp} \phi(r) \\ &\quad - 4 \int_0^{r_s} dr \int_0^r dr' (r - r') \nabla_{\perp} \phi(r') \cdot \nabla_{\perp} \phi(r), \end{aligned} \quad (\text{C.27})$$

$$\begin{aligned} &= 4 \int_0^{r_s} dr \int_r^{r_s} dr' \frac{r(r_s - r')}{r_s} \nabla_{\perp} \phi(r') \cdot \nabla_{\perp} \phi(r) \\ &\quad + 4 \int_0^{r_s} dr \int_0^r dr' \frac{r'(r_s - r)}{r_s} \nabla_{\perp} \phi(r') \cdot \nabla_{\perp} \phi(r), \end{aligned} \quad (\text{C.28})$$

$$= 8 \int_0^{r_s} dr \int_0^r dr' \frac{r'(r_s - r)}{r_s} \nabla_{\perp} \phi(r') \cdot \nabla_{\perp} \phi(r), \quad (\text{C.29})$$

$$\delta s_{\text{pB}} = -2\delta r_{\text{geo}}. \quad (\text{C.30})$$

C.2.2. Ensemble average of δs

The ensemble average of ϕ being zero, the only contribution from the ensemble average of δs comes from the post-Born term,

$$\langle \delta s \rangle = \langle \delta s_{\text{pB}} \rangle = -2\langle \delta r_{\text{geo}} \rangle > 0. \quad (\text{C.31})$$

The expression of $\langle \delta r_{\text{geo}} \rangle$ is given in Eq. (C.23).

In Sect. 4.2.2 we find that numerical estimates of $\langle \delta s \rangle$ are overwhelmed by the super-sample variance of its first-order contribution. In order to check Eq. (C.31) numerically, it would be convenient to extract its post-Born term. This can actually be done with the following estimator for the average of δs_{pB} :

$$\widehat{\delta s_{\text{pB}}} \equiv \frac{\langle \mu^{-1} \delta s \rangle_{\text{d}}}{\langle \mu^{-1} \rangle_{\text{d}}} - \langle \delta s \rangle_{\text{d}} = \langle \delta s(\boldsymbol{\beta}) \rangle - \langle \delta s(\boldsymbol{\theta}) \rangle. \quad (\text{C.32})$$

A calculation along the same lines as before indeed shows that the post-Born contribution drops from $\langle \delta s(\boldsymbol{\theta}) \rangle$, that is when ensemble averaging is taken whilst $\boldsymbol{\theta}$ is kept fixed. Since, however, the Born contribution to both $\langle \delta s(\boldsymbol{\beta}) \rangle$, $\langle \delta s(\boldsymbol{\theta}) \rangle$ is identical, their difference eliminates it from the final result.

C.2.3. Ensemble average of δs^2

Since δs is a first-order quantity, we must also evaluate its mean square as a contribution to the term $\langle (\delta r / \bar{r})^2 \rangle$ in Eq. (C.1). The computation is straightforward and yields

$$\langle \delta s^2 \rangle \approx 4 \int_0^{r_s} dr \xi_\phi(r), \quad (\text{C.33})$$

in Limber's approximation, where ξ_ϕ is the two-point correlation function of ϕ .

C.3. Total effect of the shift

Gathering the geometric and time-delay effects, we conclude that, at second order

$$\frac{\delta A_{\text{shift}}(\eta)}{\bar{A}(\eta)} \equiv \left\langle \frac{2\delta r(\eta, \boldsymbol{\beta})}{\bar{r}(\eta)} + \frac{\delta r^2(\eta, \boldsymbol{\beta})}{\bar{r}^2(\eta)} \right\rangle, \quad (\text{C.34})$$

$$= \left\langle \frac{2\delta r_{\text{geo}}}{r_s} \right\rangle + \left\langle \frac{2\delta s}{r_s} \right\rangle + \left\langle \frac{\delta s^2}{r_s^2} \right\rangle, \quad (\text{C.35})$$

$$= 2 \int_0^{r_s} dr \frac{(r_s - r)r}{r_s^2} J(r) + 4 \int_0^{r_s} dr \xi_\phi(r). \quad (\text{C.36})$$

In practice, the second term is negligible with respect to the first one and it can be discarded.

C.4. Tilt or wrinkly-surface effect

We now turn to the increase of the area due to its wrinkles. As seen in Eq. (C.1), this effect is controlled by the angle $\iota(\eta, \boldsymbol{\beta}) = |\dot{\boldsymbol{\xi}}(\eta, \boldsymbol{\beta})|$ formed by the normal to the iso- η surface and the direction spanned by $\boldsymbol{\beta}$. Specifically,

$$\frac{\delta A_{\text{tilt}}(\eta)}{\bar{A}(\eta)} \equiv \frac{1}{2} \int_{\mathbb{S}^2} \frac{d\boldsymbol{\beta}}{4\pi} \iota^2(\eta, \boldsymbol{\beta}) = \frac{1}{2} \langle |\dot{\boldsymbol{\xi}}(\eta, \boldsymbol{\beta})|^2 \rangle, \quad (\text{C.37})$$

From the expression (C.2) of $\dot{\boldsymbol{\xi}}$, we find

$$|\dot{\boldsymbol{\xi}}|^2 = 4 \int_0^{r_s} dr \int_0^{r_s} dr' \frac{rr'}{r_s^2} \nabla_\perp \phi(r) \cdot \nabla_\perp \phi(r'), \quad (\text{C.38})$$

and hence, in Limber's approximation,

$$\frac{1}{2} \langle |\dot{\boldsymbol{\xi}}(\eta, \boldsymbol{\beta})|^2 \rangle = \int_0^{r_s} dr \frac{r^2}{r_s^2} J(r), \quad (\text{C.39})$$

in agreement with Eq. (A41) of KP16.

Combining this wrinkly surface contribution with the total contribution of the shift then yields the final result

$$\frac{\delta A(\eta)}{\bar{A}(\eta)} = \int_0^{r_s} dr \frac{(2r_s - r)r}{r_s^2} J(r), \quad (\text{C.40})$$

up to the negligible δs^2 term.

Appendix D: Variance calculations

For any statistical quantity X , its variance is given by Eq. (57), which depends on the angular power spectra of X . In what follows, we give C_ℓ^X for the quantities in Sect. 4.

D.1. Variance for source and directional averages

The quantities of interest for source and directional averages are functions of μ and $\bar{\mu}$, which can be rewritten in terms of κ and $\bar{\kappa}$. Their angular power spectra can be rewritten as

$$C_\ell^\mu = C_\ell^{1/\mu} = 4C_\ell^\kappa, \quad (\text{D.1})$$

$$C_\ell^{\bar{\mu}} = C_\ell^{1/\bar{\mu}} = 4C_\ell^{\bar{\kappa}}, \quad (\text{D.2})$$

$$C_\ell^d = C_\ell^{\bar{\kappa}}, \quad (\text{D.3})$$

$$C_\ell^{\Delta m} = [5/\ln 10]^2 C_\ell^{\bar{\kappa}}. \quad (\text{D.4})$$

Therefore, to estimate the variance we need to compute the angular power spectra of κ and $\bar{\kappa}$.

As discussed in Eq. (15), $\bar{\kappa}$ may be approximated as

$$\bar{\kappa} = \kappa + \bar{\kappa}_v, \quad (\text{D.5})$$

where the contribution from redshift perturbations $\bar{\kappa}_v$ only contain the effect of peculiar velocities. For simplicity, we only account for the auto-correlation of κ and $\bar{\kappa}_v$, so that

$$C_\ell^{\bar{\kappa}} = C_\ell^\kappa + C_\ell^{\bar{\kappa}_v}. \quad (\text{D.6})$$

We generate C_ℓ^κ with NICAIA (Kilbinger et al. 2017) which uses the non-linear prescription from HALOFIT (Smith et al. 2003; Takahashi et al. 2012).

The angular power spectrum of $\bar{\kappa}_v$ is simply

$$C_\ell^{\bar{\kappa}_v} = 4\pi \left(1 - \frac{1}{\mathcal{H}r}\right)^2 \int \frac{dk}{k} j_\ell^2(kr) \frac{k^3 P_v(k)}{2\pi^2}, \quad (\text{D.7})$$

with $P_v(k)$ the peculiar-velocity power spectrum.

We only consider the variance from κ and $\bar{\kappa}_v$ for simplicity. In principle, one should also account for cross-terms as well as additional variance terms, for example due to real-space clustering for source-averaged quantities, see (Fleury et al. 2017a).

We note that we do not account for the constraint at the observer (see Sect. 3.11), because this effect is only relevant when the field of interest is very correlated at large scales, such as the gravitational potential whose power spectrum scales as $P(k)/k^4$. This is not the case for the quantities investigated here, hence we do not need to account for the constraint at the observer.

D.2. Angular power spectrum of δr_{geo}

The geometrical shift δr_{geo} at fixed distance travelled is a second-order quantity, which depends on the power spectrum of ϕ . Hence its own angular power spectrum will depend on the four-point function of ϕ . In that context, a full spherical-harmonic computation would be rather involved; furthermore, expect most of the power to be held by small scales. Hence, we choose to compute a flat-sky power spectrum $P_{\text{geo}}(\ell) \approx C_\ell[\delta r_{\text{geo}}]$.

We start from the definition of the two-point correlation function as

$$\xi_{\text{geo}}(\theta) \equiv \langle \delta r_{\text{geo}}(\boldsymbol{\theta}_1) \delta r_{\text{geo}}(\boldsymbol{\theta}_2) \rangle - \langle \delta r_{\text{geo}}(\boldsymbol{\theta}_1) \rangle \langle \delta r_{\text{geo}}(\boldsymbol{\theta}_2) \rangle, \quad (\text{D.8})$$

Table D.1. Expressions used in Eq. (D.18) for various quantities studied in Sect. 4.2, with $\mathcal{K}_X \equiv \mathcal{K}_X(r)$ and $H_X \equiv H_X(r\theta, r'\theta')$. For H_X we use the relations in Sect. 3.11.

X	\mathcal{K}_X	H_X
$\delta r(\eta)/\bar{r}(\eta)$	$2/\bar{r}(\eta)$	$\langle \phi(r\theta)\phi(r'\theta') \phi_0 \rangle$
$\delta\lambda(\eta)/\bar{\lambda}(\eta)$	$2a^2(\eta)/\bar{\lambda}(\eta)$	$\langle (\phi(r\theta) - \phi_0)(\phi(r'\theta') - \phi_0) \phi_0 \rangle$
$\delta r(\lambda)/\bar{r}(\lambda)$	$2a^2(\lambda)/(\bar{r}(\lambda)\bar{a}^2(\lambda))$	$\langle (\phi(r\theta) - \phi_0)(\phi(r'\theta') - \phi_0) \phi_0 \rangle$

with $\theta \equiv |\theta_1 - \theta_2|$. We note that since δr_{geo} is second-order, it does not really matter whether we are considering observed or ‘true’ angular positions here.

Substituting the expression (C.14) of δr_{geo} in the first term of Eq. (D.8) and using its Fourier transform yields

$$\begin{aligned} \langle \delta r_{\text{geo}}(\theta_1) \delta r_{\text{geo}}(\theta_2) \rangle &= 16 \int_0^{r_s} dr_1 \int_0^{r_1} dr'_1 \int_0^{r_s} dr_2 \int_0^{r_2} dr'_2 \\ &\times \left[\frac{(r_s - r_1)r'_1}{r_s} \right] \left[\frac{(r_s - r_2)r'_2}{r_s} \right] \int \frac{d^3 \mathbf{k}_1}{(2\pi)^3} \frac{d^3 \mathbf{k}'_1}{(2\pi)^3} \frac{d^3 \mathbf{k}_2}{(2\pi)^3} \frac{d^3 \mathbf{k}'_2}{(2\pi)^3} \\ &\times \exp i[(r_1 \mathbf{k}_1 + r'_1 \mathbf{k}'_1) \cdot \boldsymbol{\theta}_1 + (r_2 \mathbf{k}_2 + r'_2 \mathbf{k}'_2) \cdot \boldsymbol{\theta}_2] \\ &\times (\mathbf{k}_{\perp 1} \cdot \mathbf{k}_{\perp 1}') (\mathbf{k}_{\perp 2} \cdot \mathbf{k}_{\perp 2}') \langle \tilde{\phi}(\mathbf{k}_1) \tilde{\phi}(\mathbf{k}'_1) \tilde{\phi}(\mathbf{k}_2) \tilde{\phi}(\mathbf{k}'_2) \rangle, \end{aligned} \quad (\text{D.9})$$

which raises the difficulty of computing the four-point correlation $\langle \tilde{\phi}(\mathbf{k}_1) \tilde{\phi}(\mathbf{k}'_1) \tilde{\phi}(\mathbf{k}_2) \tilde{\phi}(\mathbf{k}'_2) \rangle$. Assuming that ϕ is reasonably modelled by a Gaussian random field, we may neglect the tri-spectrum and apply Wick’s theorem as

$$\begin{aligned} \langle \tilde{\phi}(\mathbf{k}_1) \tilde{\phi}(\mathbf{k}'_1) \tilde{\phi}(\mathbf{k}_2) \tilde{\phi}(\mathbf{k}'_2) \rangle &= \langle \tilde{\phi}(\mathbf{k}_1) \tilde{\phi}(\mathbf{k}'_1) \rangle \langle \tilde{\phi}(\mathbf{k}_2) \tilde{\phi}(\mathbf{k}'_2) \rangle \\ &+ \langle \tilde{\phi}(\mathbf{k}_1) \tilde{\phi}(\mathbf{k}_2) \rangle \langle \tilde{\phi}(\mathbf{k}'_1) \tilde{\phi}(\mathbf{k}'_2) \rangle + \langle \tilde{\phi}(\mathbf{k}_1) \tilde{\phi}(\mathbf{k}'_2) \rangle \langle \tilde{\phi}(\mathbf{k}'_1) \tilde{\phi}(\mathbf{k}_2) \rangle. \end{aligned} \quad (\text{D.10})$$

Each of the three terms leads to a different contribution. The first one exactly compensates $\langle \delta r_{\text{geo}} \rangle^2$ in Eq. (D.8); the third one vanishes in Limber’s approximation; the second one holds the interesting correlation and we eventually get

$$\begin{aligned} \xi_{\text{geo}}(\theta) &= 16 \int_0^{r_s} dr \int_0^r dr' \left[\frac{(r_s - r)r'}{r_s} \right]^2 \frac{1}{(rr')^4} \\ &\times \int \frac{d^2 \boldsymbol{\ell}}{(2\pi)^2} \frac{d^2 \boldsymbol{\ell}'}{(2\pi)^2} (\boldsymbol{\ell} \cdot \boldsymbol{\ell}')^2 e^{i(\boldsymbol{\ell} + \boldsymbol{\ell}') \cdot (\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2)} P_\phi \left(\frac{\boldsymbol{\ell}}{r} \right) P_\phi \left(\frac{\boldsymbol{\ell}'}{r'} \right). \end{aligned} \quad (\text{D.11})$$

The power spectrum $P_{\text{geo}}(L)$ must satisfy

$$C_{\text{geo}}(\theta) = \int \frac{d^2 \boldsymbol{\ell}}{(2\pi)^2} e^{i\boldsymbol{\ell} \cdot \boldsymbol{\theta}} P_{\text{geo}}(\boldsymbol{\ell}). \quad (\text{D.12})$$

Thus, introducing the variable $\mathbf{L} \equiv \boldsymbol{\ell} + \boldsymbol{\ell}'$ and performing the change of variable $(\boldsymbol{\ell}, \boldsymbol{\ell}') \mapsto (\boldsymbol{\ell}, \mathbf{L})$ in Eq. (D.11), we immediately identify

$$\begin{aligned} P_{\text{geo}}(L) &= 16 \int_0^{r_s} dr \int_0^r dr' \left[\frac{(r_s - r)r'}{r_s} \right]^2 \frac{1}{(rr')^4} \int \frac{d^2 \boldsymbol{\ell}}{(2\pi)^2} \\ &\times [(\mathbf{L} - \boldsymbol{\ell}) \cdot \boldsymbol{\ell}]^2 P_\phi \left(\eta_0 - r, \frac{\boldsymbol{\ell}}{r} \right) P_\phi \left(\eta_0 - r', \frac{\mathbf{L} - \boldsymbol{\ell}}{r'} \right). \end{aligned} \quad (\text{D.13})$$

In the last integral, the direction of the vector \mathbf{L} does not matter, because integration over $\boldsymbol{\ell}$ makes everything isotropic. In practice, one may take \mathbf{L} to be aligned with \mathbf{e}_x . With that convention,

the two-dimensional integral over $\boldsymbol{\ell}$ becomes

$$\begin{aligned} &\int \frac{d^2 \boldsymbol{\ell}}{(2\pi)^2} [(\mathbf{L} - \boldsymbol{\ell}) \cdot \boldsymbol{\ell}]^2 P_\phi \left(\eta_0 - r, \frac{\boldsymbol{\ell}}{r} \right) P_\phi \left(\eta_0 - r', \frac{\mathbf{L} - \boldsymbol{\ell}}{r'} \right) \\ &= \int_0^\infty \frac{\ell d\ell}{2\pi} P_\phi \left(\eta_0 - r, \frac{\boldsymbol{\ell}}{r} \right) \int_0^{2\pi} \frac{d\psi}{2\pi} [\ell(L \cos \psi - \ell)]^2 \\ &\quad \times P_\phi \left(\eta_0 - r', \frac{\sqrt{L^2 + \ell^2 - 2L\ell \cos \psi}}{r'} \right). \end{aligned} \quad (\text{D.14})$$

Introducing the integration variable $k \equiv \ell/r$, and then making the change $L \rightarrow \ell$, we get the final result

$$\begin{aligned} P_{\text{geo}}(\ell) &= 16 \int_0^{r_s} dr \int_0^r dr' \left[\frac{(r_0 - r)}{r_0 r'} \right]^2 \\ &\times \int_0^\infty \frac{k^3 dk}{2\pi} \int_0^{2\pi} \frac{d\psi}{2\pi} (\ell \cos \psi - kr)^2 \\ &\times P_\phi(\eta_0 - r, k) P_\phi \left(\eta_0 - r', \frac{\sqrt{\ell^2 + (kr)^2 - 2\ell kr \cos \psi}}{r'} \right). \end{aligned} \quad (\text{D.15})$$

D.3. Angular power spectrum of $\delta r(z)$

Regarding the relative perturbations on the comoving distance at constant observed redshift, since we only account for the Doppler effect, the prediction is proportional to the velocity field. We therefore compute its angular power spectrum similarly to Eq. (D.7) but with the pre-factor $1/\mathcal{H}r$ instead of $(1 - 1/\mathcal{H}r)$.

D.4. Constrained variance

All the other quantities that we investigate in Sect. 4.2, that is $\delta r(\eta)$, $\delta\lambda(\eta)$ and $\delta r(\lambda)$, are obtained by line-of-sight integrations of the gravitational potential. As such, these are particularly impacted by the constraint at the observer (see Sect. 3.11). Here we show how to compute the variance for such quantities.

We consider a scalar quantity X that is a line-of-sight projection of the potential ϕ , with

$$X = \int_0^{r_s} dr \mathcal{K}_X(r) \phi(\eta_0 - r, r), \quad (\text{D.16})$$

where \mathcal{K}_X is the kernel associated with X . To estimate the variance given the constrained field $\phi(\mathbf{r})$ it is easier to work in configuration space and compute the variance using (see Sect. 3.9)

$$\sigma_{\text{ss}}^2 = \int d^2 \boldsymbol{\theta} d^2 \boldsymbol{\theta}' W(\boldsymbol{\theta}) W(\boldsymbol{\theta}') \langle X(\boldsymbol{\theta}) X(\boldsymbol{\theta}') \rangle. \quad (\text{D.17})$$

In Eq. (60) we used a window function for a cone-shaped geometry with circular base for simplicity. Actually, our narrow cones are pyramid-shaped (we expect the difference to be negligible compared to the circular case). To compute Eq. (D.17), we set the integration boundaries to $\varphi = [-\Delta/2, \Delta/2]$, $\vartheta = [\pi/2 + \Delta/2, \pi/2 - \Delta/2]$, where (φ, ϑ) are the angles in spherical coordinates. For the intermediate and deep cones, $\Delta = 50$ and 20 degrees respectively.

Assuming that $\langle X(\boldsymbol{\theta}) X(\boldsymbol{\theta}') \rangle = \omega_X(|\boldsymbol{\theta} - \boldsymbol{\theta}'|)$ is statistically isotropic, we find

$$\omega_X(|\boldsymbol{\theta} - \boldsymbol{\theta}'|) = \int_0^{r_s} dr \int_0^r dr' \mathcal{K}_X(r) \mathcal{K}_X(r') H_X(r\theta, r'\theta'). \quad (\text{D.18})$$

Then, to compute the relevant variance for Sect. 4.2, we used the functions shown in Table D.1.