

Controlling intrinsic-shear alignment in three-point weak lensing statistics

X. Shi^{1,2}, B. Joachimi¹, and P. Schneider¹

¹ Argelander-Institut für Astronomie (AlfA), Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
e-mail: xun@astro.uni-bonn.de

² International Max Planck Research School (IMPRS) for Astronomy and Astrophysics at the Universities of Bonn and Cologne

Received 3 February 2010 / Accepted 5 July 2010

ABSTRACT

Three-point weak lensing statistics provide cosmic information that complements two-point statistics. However, both statistics suffer from intrinsic-shear alignment, which is one of their limiting systematics. The nulling technique is a model-independent method developed to eliminate intrinsic-shear alignment at the two-point level. In this paper we demonstrate that the nulling technique can also be naturally generalized to the three-point level, thereby controlling the corresponding GGI systematics.

We show that under the assumption of exact redshift information the intrinsic-shear alignment contamination can be completely eliminated. To show how well the nulling technique performs on data with limited redshift information, we apply the nulling technique to three-point weak lensing statistics from a fictitious survey analogous to a typical future deep imaging survey, in which the three-point intrinsic-shear alignment systematics is generated from a power-law toy model.

Using 10 redshift bins, the nulling technique leads to a factor of 10 suppression of the GGI/GGG ratio and reduces the bias on cosmological parameters to less than the original statistical error. More detailed redshift information allowing for finer redshift bins leads to better bias reduction performance. The information loss during the nulling procedure doubles the statistical error on cosmological parameters. A comparison of the nulling technique with an unconditioned compression of the data suggests that part of the information loss can be retained by considering higher order nulling weights during the nulling procedure. A combined analysis of two- and three-point statistics confirms that the information contained in them is of comparable size and is complementary, both before and after nulling.

Key words. cosmology: theory – methods: data analysis – gravitational lensing: weak – large-scale structure of Universe – cosmological parameters

1. Introduction

Weak gravitational lensing refers to the mild distortion of the light from distant sources by the large-scale matter inhomogeneity between the source and the observer. One observable effect of weak gravitational lensing is the coherent shape distortion of the light sources, known as cosmic shear. Cosmic shear is sensitive to all cosmological parameters which have influence on the density perturbations and/or the geometry of the universe, including those concerning properties of dark energy, which have been a key concern after the discoveries made by observations of supernovae, the cosmic microwave background, and the large-scale structure (for a review see e.g. [Munshi et al. 2008](#)).

Since its first detection in 2000 ([Bacon et al. 2000](#); [Kaiser et al. 2000](#); [Van Waerbeke et al. 2000](#); [Wittman et al. 2000](#)), cosmic shear has been developed into a competitive cosmological probe. Its constraining power on cosmological parameters is now comparable to other probes (e.g. [Spergel et al. 2007](#); [Fu et al. 2008](#)).

With forthcoming large field multicolor imaging surveys (e.g. DES¹, KIDS², EUCLID³, etc.), photometric redshift and shape information of a huge number of galaxies will be available, rendering cosmic shear even greater statistical power. In

particular, cosmic shear is considered to be one of the most promising dark energy probes ([Albrecht et al. 2006](#); [Peacock et al. 2006](#)) when the results of these surveys become available.

Such constraining power will be further enhanced by the use of higher-order statistics. Second-order statistics measure only the Gaussian signature of a random field. Even if the primordial cosmic density field is Gaussian, non-Gaussianity will be generated due to the nonlinear nature of gravitational clustering. Such non-Gaussianity in the cosmic density field will then show up via its lensing effect, leading to non-Gaussian signals in the cosmic shear field. A common way of measuring non-Gaussianity is to use higher-order statistics. In cosmic shear studies, several authors have shown that the lowest order of them, i.e. the third-order statistics, already provide a valuable probe for cosmological parameter estimates; in particular it can break the near degeneracy between the density parameter Ω_m and the power spectrum normalization σ_8 ([Bernardeau et al. 1997](#); [Jain & Seljak 1997](#); [van Waerbeke et al. 1999](#); [Hui 1999](#)). A more recent study by [Takada & Jain \(2004\)](#) (TJ04 afterwards) showed that including third-order statistics can improve parameter constraints significantly, typically by a factor of three.

But the ultimate performance of these future surveys still largely depends on the how well the systematic errors can be controlled (e.g. [Huterer et al. 2006](#)). In this paper we focus on a particularly worrisome systematic error in cosmic shear studies: the intrinsic-shear alignment, and demonstrate a way to control it for shear three-point statistics.

¹ <http://www.darkenergysurvey.org/>

² <http://www.astro-wise.org/projects/KIDS/>

³ <http://sci.esa.int/science-e/www/area/index.cfm?fareaid=102>

In the weak lensing limit the observed ellipticity of a galaxy ϵ_{obs} can be written as the sum of the intrinsic ellipticity ϵ_1 of the galaxy, and the shear γ which is caused by gravitational lensing of the foreground matter distribution. Here ϵ_{obs} , ϵ_1 and γ are complex quantities. Intrinsic-shear alignment is defined in two-point cosmic shear statistics as the correlation between the intrinsic ellipticity of one galaxy and the shear of another galaxy (the GI term, [Hirata & Seljak 2004](#)). Three-point statistics $\langle \epsilon_{\text{obs}}^i \epsilon_{\text{obs}}^j \epsilon_{\text{obs}}^k \rangle$, a correlator of ellipticities of three galaxy images i , j and k , can also be expanded into lensing (GGG), intrinsic-shear (GGI and GII), and intrinsic (III) terms:

$$\langle \epsilon_{\text{obs}}^i \epsilon_{\text{obs}}^j \epsilon_{\text{obs}}^k \rangle = \text{GGG} + \text{GGI} + \text{GII} + \text{III}, \text{ with} \quad (1)$$

$$\text{GGG} = \langle \gamma^i \gamma^j \gamma^k \rangle, \quad (2)$$

$$\text{GGI} = \langle \epsilon_1^i \gamma^j \gamma^k \rangle + \langle \epsilon_1^j \gamma^i \gamma^k \rangle + \langle \epsilon_1^k \gamma^i \gamma^j \rangle, \quad (3)$$

$$\text{GII} = \langle \epsilon_1^i \epsilon_1^j \gamma^k \rangle + \langle \epsilon_1^j \epsilon_1^i \gamma^k \rangle + \langle \epsilon_1^k \epsilon_1^i \gamma^j \rangle, \quad (4)$$

$$\text{III} = \langle \epsilon_1^i \epsilon_1^j \epsilon_1^k \rangle. \quad (5)$$

Physically, if one assumes that galaxies are randomly oriented on the sky, only the desired GGG term remains on the right-hand side of (1). However, when these galaxies are subject to the tidal gravitational force of the same matter structure (e.g. they formed under the influence of the same massive dark matter halo), their shapes can intrinsically align and become correlated, giving rise to a nonvanishing III term. Furthermore, GGI and GII terms can be generated when a matter structure tidally influences close-by galaxies and at the same time contributes to the shear signal of background objects, leading to correlations among them.

In two-point statistics, the corresponding intrinsic (II) and intrinsic-shear (GI) terms have been subject to detailed studies both theoretically (e.g. [Catelan et al. 2001](#); [Croft & Metzler 2000](#); [Heavens et al. 2000](#); [Hui & Zhang 2002](#); [Mackey et al. 2002](#); [Jing 2002](#); [Hirata & Seljak 2004](#); [Heymans et al. 2006](#); [Bridle & Abdalla 2007](#); [Schneider & Bridle 2010](#)) and observationally ([Brown et al. 2002](#); [Heymans et al. 2004](#); [Mandelbaum et al. 2006, 2009](#); [Hirata et al. 2007](#); [Fu et al. 2008](#); [Brainerd et al. 2009](#); [Okumura et al. 2009](#); [Okumura & Jing 2009](#)). Although the results of these studies show large variations, most of them are consistent with a 10% contamination by both II and GI correlations for future surveys with photometric redshift information. Especially, neglecting these correlations can bias the dark energy equation of state parameter w_0 by as much as 50% ([Bridle & King 2007](#)) for a “shallow” survey described in [Amara & Réfrégier \(2007\)](#). For three-point shear statistics, there have been few measurements up to now ([Bernardeau et al. 2002b](#); [Pen et al. 2003](#); [Jarvis et al. 2004](#)). However the potential systematics level in these studies is found to be high. A recent numerical study by [Semboloni et al. \(2008\)](#) showed that intrinsic alignments affect three-point weak lensing statistics more strongly than at the two-point level for a given survey depth. In particular, neglecting GGI and GII systematics would lead to an underestimation of the GGG signal by 5–10% for a moderately deep survey like the CFHTLS Wide. Therefore, to match the statistical power expected for cosmic shear in the future surveys, it is essential to control these systematics.

The intrinsic alignment, II(III) in the two-(three-) point case, is relatively straightforward to eliminate, since it requires that the galaxies in consideration are physically close to each other, i.e. have very similar redshifts and angular positions

([King & Schneider 2002, 2003](#); [Heymans & Heavens 2003](#); [Takada & White 2004](#)). The control of intrinsic-shear systematics, GI for the two-point case and GGI in the three-point case (GII also requires that two of the three galaxies are physically close and thus can be eliminated in the same way as II and III), turns out to be a much greater challenge. However, as already pointed out by HS04, the characteristic dependence on galaxy redshifts is a valuable piece of information that helps to control the intrinsic-shear alignments.

Several methods for this have already been constructed in the context of two-point statistics. They can be roughly classified into three categories: modeling ([King 2005](#); [Bridle & King 2007](#)), nulling ([Joachimi & Schneider 2008](#), JS08 hereafter; [Joachimi & Schneider 2009](#)) and self-calibration ([Zhang 2008](#); [Joachimi & Bridle 2009](#)). Modeling separates cosmic shear from the intrinsic-shear alignment effect by constructing template functions for the latter. It suffers from uncertainties of the model due to the lack of knowledge of the angular scale and redshift dependence of the intrinsic-shear signal. The nulling technique employs the characteristic redshift dependence of the intrinsic-shear signal to “null it out”. It is a purely geometrical method and is model-independent, but suffers from a significant information loss. Self-calibration intends to solve the problem of information loss by using additional information from the galaxy distribution to “calibrate” the signal. The original form of self-calibration, proposed by [Zhang \(2008\)](#), is model-independent but strong assumptions have been made. [Joachimi & Bridle \(2009\)](#) then develop it into a modeling method, by treating intrinsic alignments and galaxy biasing as free functions of scale and redshift.

All these methods have the potential of being generalized to three-point statistics. In this paper we focus on the nulling technique, and establish it as a method to reduce the three-point intrinsic-shear alignments GGI and GII. Since GII can be removed by discarding close pairs of galaxies as in the case of II controlling (e.g. [Heymans & Heavens 2003](#)), we focus on the control of GGI systematics.

As known from the case of two-point statistics, the nulling technique introduces significant information loss while (in principle) completely removing the intrinsic-shear alignment from the signal. In this work we compare the nulling technique to an unconditioned linear compression of the data, distinguish different sources of such information loss, and discuss the possible ways of reducing it. We also study the combined constraints on cosmological parameters with both two- and three-point cosmic shear statistics.

In Sect. 2 we demonstrate why and how the nulling technique can be applied to three-point lensing statistics. We then apply the nulling technique to the modeled lensing bispectrum which we contaminate by intrinsic-shear alignment. The modeling details are described in Sect. 3. The method of nulling weights construction and the corresponding results are shown in Sect. 4, while the results concerning the constraints on cosmological parameters are presented in Sect. 5. We conclude in Sect. 6.

We will work in the context of a spatially flat CDM cosmology with a variable dark energy whose equation of state w is parameterized as $w = w_0 + w_a(1 - a)$, with a the cosmic scale factor. The adopted fiducial values for cosmological parameters are $\Omega_m = 0.3$, $\Omega_b = 0.045$, $\Omega_{\text{de}} = 0.7$, $w_0 = -0.95$, $w_a = 0.0$, $h = 0.7$, $n_s = 1.0$, and $\sigma_8 = 0.8$. Here, Ω_m , Ω_b and Ω_{de} are the density parameters of the matter (including cold dark matter and baryons), baryons and the dark energy at present time, n_s is the spectral index of the primordial power spectrum of scalar perturbations, h is the dimensionless Hubble parameter defined by

$H_0 = 100 h \text{ km s}^{-1}/\text{Mpc}$, and σ_8 is the rms mass fluctuation in spheres of radius $8 h^{-1} \text{ Mpc}$.

2. The nulling technique applied to three-point shear tomography

2.1. Principle of the nulling technique

The shear on the image of a distant galaxy is a result of gravitational distortion of light caused by the inhomogeneous three-dimensional matter distribution in the foreground of that galaxy. For notational simplicity, we will use the dimensionless surface mass density (the convergence) κ instead of the shear γ as a measure for the lensing signal throughout the paper, although in reality the signal is based on the measurement of the shear. This will not affect our results since κ and γ are linearly related on each redshift plane while our method is dealing with the redshift dependence of them (the same reason justifies the turning to the Fourier domain in the next subsection).

When one measures the shear γ , the direct observable is the galaxy ellipticity $\epsilon_{\text{obs}} = \epsilon_1 + \gamma$. The shear γ is a signal caused by gravitational distortion which is a deterministic process, where the intrinsic ellipticity ϵ_1 can be further written as the sum of a deterministic part ϵ_1^{det} which is caused by intrinsic alignment, and a stochastic part ϵ_1^{ran} which does not correlate with any other quantity. There is no correlation between ϵ_1^{ran} of different galaxies either.

We define κ_{obs} and κ_1 which are the correspondences of $\epsilon_1^{\text{det}} + \gamma$ and ϵ_1^{det} . We remove the stochastic part since κ is deterministic. Note that κ_{obs} and κ_1 are analogs of the dimensionless surface mass density κ but do not have any direct physical meaning as κ does. They are complex quantities in general and can lead to a B-mode signal. To better distinguish the real measurable κ from them, we denote it as κ_G in the rest of the paper since it is the physical quantity which is related to the gravitational lensing signal. Keeping the dominating linear term, the convergence κ_G can be written as (details see e.g. [Schneider 2006](#)):

$$\kappa_G(\boldsymbol{\theta}, \chi_s) = \frac{3 \Omega_m H_0^2}{2c^2} \int_0^{\chi_s} d\chi \frac{\chi(\chi_s - \chi)}{\chi_s} \frac{\delta(\chi, \boldsymbol{\theta}, \chi)}{a(\chi)}, \quad (6)$$

where δ is the three-dimensional matter density contrast, χ_s is the comoving distance of the background galaxy which is acting as a source, and $a(\chi)$ is the cosmic scale factor at the comoving distance χ of δ which is acting as a lens.

Equation (6) clearly shows that the contribution of the matter inhomogeneity δ at comoving distance χ_i to the cosmic shear signal of background galaxies can be considered as a function of the source distance χ_s , and this function is proportional to $1 - \chi_i/\chi_s$. The nulling technique takes advantage of this characteristic dependence on source distance χ_s by constructing a weight function $T(\chi_i, \chi_s)$ such that the product of $T(\chi_i, \chi_s)$ and $1 - \chi_i/\chi_s$ has an average of zero on the range between χ_i and the comoving distance to the horizon χ_{hor} :

$$\int_{\chi_i}^{\chi_{\text{hor}}} d\chi_s T(\chi_i, \chi_s) \left(1 - \frac{\chi_i}{\chi_s}\right) = 0. \quad (7)$$

One then uses this weight function as a weight for integrating over the source distance:

$$\hat{\kappa}_G(\chi_i, \boldsymbol{\theta}) := \int_{\chi_i}^{\chi_{\text{hor}}} d\chi_s T(\chi_i, \chi_s) \kappa_G(\boldsymbol{\theta}, \chi_s). \quad (8)$$

The resulting new measure of shear signal $\hat{\kappa}_G(\chi_i, \boldsymbol{\theta})$ is then free of contributions from the matter inhomogeneity at distance

χ_i . Note that although the weight function T has two arguments χ_i and χ_s here, we consider it as a function of χ_s for a particular χ_i .

Consider a correlator $\langle \kappa_{\text{obs}}^i \kappa_{\text{obs}}^j \rangle$ with comoving distances $\chi_i < \chi_j$. With a similar decomposition as (1), it is straightforward to see that the GI term in it is $\langle \kappa_1^i \kappa_G^j \rangle$. The term $\langle \kappa_G^i \kappa_1^j \rangle$ vanishes since the lensing signal at χ_i is correlated only with matter with $\chi \leq \chi_i$, whereas κ_1^j originates solely from physical processes happening at χ_j . If we integrate $\langle \kappa_1^i \kappa_G^j \rangle$ over χ_j with a weight function that eliminates the contributions to κ_G^j by the matter inhomogeneity at distance χ_i , this correlator will also vanish,

$$\int_{\chi_i}^{\chi_{\text{hor}}} d\chi_j T(\chi_i, \chi_j) \langle \kappa_1^i \kappa_G^j \rangle = 0, \quad (9)$$

since it is just the matter inhomogeneity at distance χ_i that gives rise to the correlation between κ_1^i and κ_G^j . Thus, when we integrate over $\langle \kappa_{\text{obs}}^i \kappa_{\text{obs}}^j \rangle$ with the same weight function, the GI contamination in it will be ‘‘nulled out’’. Equation (7) is the condition that the weight function T should satisfy in order to ‘‘null’’ the intrinsic-shear alignment terms, so we call it ‘‘the nulling condition’’.

The same applies to three-point statistics. Consider a correlator $\langle \kappa_{\text{obs}}^i \kappa_{\text{obs}}^j \kappa_{\text{obs}}^k \rangle$ with χ_i being the smallest comoving distance of the three. Both GII and GGI systematics contained in it also originate from the matter inhomogeneity at distance χ_i . Typically, the generation of GII systematics requires that $\chi_i \approx \chi_j < \chi_k$, while the generation of GGI requires $\chi_i < \chi_j$ and $\chi_i < \chi_k$. For both cases, the dependence of GII or GGI systematics on χ_k is also just $1 - \chi_i/\chi_k$. So new measures built as $\int_{\chi_i}^{\chi_{\text{hor}}} d\chi_k T(\chi_i, \chi_j, \chi_k) \langle \kappa_{\text{obs}}^i \kappa_{\text{obs}}^j \kappa_{\text{obs}}^k \rangle$ with T satisfying the nulling condition for three-point statistics

$$\int_{\chi_i}^{\chi_{\text{hor}}} d\chi_k T(\chi_i, \chi_j, \chi_k) \left(1 - \frac{\chi_i}{\chi_k}\right) = 0 \quad (10)$$

will be free of both GII and GGI contamination. Again, $T(\chi_i, \chi_j, \chi_k)$ here should be seen as a function of χ_k whose form depends on χ_i and χ_j .

Note that this method only depends on the characteristic redshift dependencies of the lensing signal and intrinsic-alignment signals, and is not limited to E-mode fields. This is a reassuring feature since while the κ_G field is a pure E-mode field to first order, the κ_1 field can have a B-mode component. However, if parity-invariance is assumed, any correlation function which contains an odd number of B-mode shear components vanishes ([Schneider 2003](#)), thus there should be no B-mode component in the GGI signal.

2.2. Nulling formalism for lensing bispectrum tomography

Since the nulling technique relies on the distinct redshift dependence of the intrinsic-alignment signal, redshift information is crucial for it. With the help of infrared bands, forthcoming multicolor imaging surveys can provide rather accurate photometric redshift information for the galaxies (e.g. [Abdalla et al. 2008](#); [Bordoloi et al. 2010](#)), allowing tomographic studies of cosmic shear statistics. We base our study on cosmic shear bispectrum tomography, and outline the corresponding formalism of the nulling technique in the following.

Given the galaxy redshift probability distribution of redshift bin i which we denote as $p_s^{(i)}(z) = p_s^{(i)}(\chi_s) d\chi_s/dz$, one can define the average convergence field in redshift bin i by integrating

$\kappa(\boldsymbol{\theta}, \chi_s)$ in (6) over $p_s^{(i)}(\chi_s)$. We turn to angular frequency space now and define

$$\tilde{\kappa}_G^{(i)}(\boldsymbol{\ell}) := \int_0^{\chi_{\text{hor}}} d\chi_s p_s^{(i)}(\chi_s) \tilde{\kappa}_G(\boldsymbol{\ell}, \chi_s), \quad (11)$$

where $\tilde{\kappa}_G(\boldsymbol{\ell}, \chi_s)$ is the Fourier transform of $\kappa_G(\boldsymbol{\theta}, \chi_s)$. To better show the relation between $\tilde{\kappa}_G^{(i)}(\boldsymbol{\ell})$ and three-dimensional matter inhomogeneity in Fourier space $\tilde{\delta}(k, \chi)$, one can write, basing on (6) and (11),

$$\tilde{\kappa}_G^{(i)}(\boldsymbol{\ell}) = \int_0^{\chi_{\text{hor}}} d\chi W^{(i)}(\chi) \tilde{\delta}(\boldsymbol{\ell}/\chi, \chi), \quad (12)$$

by defining a lensing weight function $W^{(i)}(\chi)$ as

$$W^{(i)}(\chi) := \frac{3 \Omega_m H_0^2 \chi}{2 a(\chi) c^2} \int_{\chi}^{\chi_{\text{hor}}} d\chi_s p_s^{(i)}(\chi_s) \frac{\chi_s - \chi}{\chi_s}. \quad (13)$$

The tomographic lensing bispectrum is defined via

$$\langle \tilde{\kappa}_G^{(i)}(\boldsymbol{\ell}_1) \tilde{\kappa}_G^{(j)}(\boldsymbol{\ell}_2) \tilde{\kappa}_G^{(k)}(\boldsymbol{\ell}_3) \rangle = (2\pi)^2 B_{\text{GGG}}^{(ijk)}(\ell_1, \ell_2, \ell_3) \delta_D(\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2 + \boldsymbol{\ell}_3), \quad (14)$$

where the Dirac delta function ensures that the bispectrum is defined only when $\boldsymbol{\ell}_1$, $\boldsymbol{\ell}_2$, and $\boldsymbol{\ell}_3$ form a triangle. This fact arises from statistical homogeneity, while that the bispectrum can be defined as a function independent of the directions of the angular frequency vectors arises from statistical isotropy.

In a survey, the convergence field $\tilde{\kappa}_{\text{obs}}$ is determined from the observed galaxy ellipticities, and the corresponding bispectrum B_{obs} suffers from intrinsic-shear alignments. As we did with the three-point correlator in Sect. 1, we separate the observed lensing bispectrum into the four terms:

$$B_{\text{obs}} = B_{\text{GGG}} + B_{\text{GGI}} + B_{\text{GGII}} + B_{\text{III}}. \quad (15)$$

Among them, B_{GGI} , B_{GGII} and B_{III} can be linked to the convergence in a similar way as (14), for example

$$\langle \tilde{\kappa}_1^{(i)}(\boldsymbol{\ell}_1) \tilde{\kappa}_G^{(j)}(\boldsymbol{\ell}_2) \tilde{\kappa}_G^{(k)}(\boldsymbol{\ell}_3) \rangle = (2\pi)^2 B_{\text{GGI}}^{(ijk)}(\ell_1, \ell_2, \ell_3) \delta_D(\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2 + \boldsymbol{\ell}_3). \quad (16)$$

Here we assume disjoint redshift bins and let i to be the redshift bin with the lowest redshift, so $\langle \tilde{\kappa}_G^{(i)}(\boldsymbol{\ell}_1) \tilde{\kappa}_1^{(j)}(\boldsymbol{\ell}_2) \tilde{\kappa}_G^{(k)}(\boldsymbol{\ell}_3) \rangle$ and $\langle \tilde{\kappa}_G^{(i)}(\boldsymbol{\ell}_1) \tilde{\kappa}_G^{(j)}(\boldsymbol{\ell}_2) \tilde{\kappa}_1^{(k)}(\boldsymbol{\ell}_3) \rangle$ both vanish due to the same reason as explained in Sect. 2.1 for the two-point statistics.

The purpose of the nulling technique is to filter B_{obs} in such a way that the GGI term is strongly suppressed in comparison with the GGG term. The GII and III terms can be removed by ignoring the signal coming from bispectrum $B_{\text{obs}}^{(ijk)}(\ell_1, \ell_2, \ell_3)$ with two or three equal redshift bins.

To fulfill this purpose, we construct our new measures as

$$Y^{(ij)}(\ell_1, \ell_2, \ell_3) := \sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k) B_{\text{obs}}^{(ijk)}(\ell_1, \ell_2, \ell_3) \chi'_k \Delta z_k, \quad (17)$$

where N_z is the total number of redshift bins, χ'_k is the derivative of comoving distance with respect to redshift, and Δz_k is the width of redshift bin k . The weight function is written now as $T^{(ij)}(\chi_k)$ since i and j indicate two redshift bins, i.e. two populations of galaxies, rather than two comoving distances as in the previous subsection. The weight $T^{(ij)}$ is required to satisfy the nulling condition (7) in its discretized form,

$$O^{(ij)} := \sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k) \left(1 - \frac{\chi_i}{\chi_k}\right) \chi'_k \Delta z_k = 0, \quad (18)$$

for all $j > i$. Here, χ_i and χ_k should be chosen such that they represent well the distance to redshift bins i and k . In this paper we choose them to be the distances corresponding to the median redshift of the bin. The summation over index k runs from $i + 1$ rather than i since we consider only bispectrum measures with $j > i$ and $k > i$ to avoid III and GII systematics. In this case $B_{\text{obs}}^{(ijk)}$ in (17) can be written as a sum of $B_{\text{GGG}}^{(ijk)}$ and $B_{\text{GGI}}^{(ijk)}$, and $Y^{(ij)}$ can be expressed as

$$Y^{(ij)}(\ell_1, \ell_2, \ell_3) = \sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k) B_{\text{GGG}}^{(ijk)}(\ell_1, \ell_2, \ell_3) \chi'_k \Delta z_k + \sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k) B_{\text{GGI}}^{(ijk)}(\ell_1, \ell_2, \ell_3) \chi'_k \Delta z_k. \quad (19)$$

Suppose one has infinitely many redshift bins, then the lensing signal in bin k caused by the matter inhomogeneity in bin i is exactly proportional to $1 - \chi_i/\chi_k$, which means $B_{\text{GGI}}^{(ijk)}(\ell_1, \ell_2, \ell_3)$ can be written as a product of $1 - \chi_i/\chi_k$ and some function of the parameters other than χ_k :

$$B_{\text{GGI}}^{(ijk)}(\ell_1, \ell_2, \ell_3) = \mathcal{F}(\chi_i, \chi_j, \ell_1, \ell_2, \ell_3) \left(1 - \frac{\chi_i}{\chi_k}\right). \quad (20)$$

Then we have

$$\sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k) B_{\text{GGI}}^{(ijk)}(\ell_1, \ell_2, \ell_3) \chi'_k \Delta z_k = \mathcal{F}(\chi_i, \chi_j, \ell_1, \ell_2, \ell_3) \sum_{k=i+1}^{N_z} T^{(ij)}(\chi_k) \left(1 - \frac{\chi_i}{\chi_k}\right) \chi'_k \Delta z_k = 0. \quad (21)$$

This suggests that only the GGG contribution is left in the nulled measure $Y^{(ij)}$, the GGI contribution has been ‘‘nulled out’’ due to the nulling condition. If only a limited number of redshift bins is available, (20) holds only approximately, leading to a residual in (21).

Since the nulling condition is the only condition that the weight $T^{(ij)}$ must satisfy in order to ‘‘null’’, there is much freedom in choosing the form of it. We would like to further specify its form such that it preserves as much Fisher information in $Y^{(ij)}$ as possible. The method we have adopted for the nulling weight construction will be detailed in Sect. 4.

Note that for each (i, j) combination, one can in principle apply more than one nulling weight to the original bispectrum, and obtain more nulled measures. If one retains the condition of maximizing the Fisher information and demands that all the weight functions built for one (i, j) combination are orthogonal to each other, one arrives at higher-order modes that have the second-most, third-most, etc., information content (higher-order weights, see JS08). The total number of such linearly independent nulled measures for a certain (i, j) equals the possible values of $k \geq i + 1$. In this schematic study we will only use the optimum, i.e. the first-order nulling weights. We will assess the information loss due to this limitation in Sect. 5.4.

3. Modeling

3.1. Survey characteristics

We set up a fictitious survey with a survey size of $A = 4000 \text{ deg}^2$ which is similar to the survey size of DES. This can be easily scaled to any survey size using the proportionality of statistical

errors to $A^{-1/2}$. We assume a galaxy intrinsic ellipticity dispersion $\sigma_\epsilon = \sigma(\epsilon_1^{\text{int}}) = 0.35$. As galaxy redshift probability distribution we adopt the frequently used parameterization (Smail et al. 1994),

$$p_s(z) \propto \left(\frac{z}{z_0}\right)^\alpha \exp\left\{-\left(\frac{z}{z_0}\right)^\beta\right\}, \quad (22)$$

and use $z_0 = 0.64$, $\alpha = 2$, $\beta = 1.5$. The distribution is cut at $z_{\text{max}} = 3$ and normalized to 1. The corresponding median redshift of this fictitious survey is $z_m = 0.9$, which is compatible to a survey like EUCLID. We adopt an average galaxy number density $\bar{n}_g = 40 \text{ arcmin}^{-2}$ which is again EUCLID-like.

Disjunct redshift bins without photo- z error are assumed, which means that the galaxy redshift probability distribution in redshift bin i takes the form $p_s^{(i)}(z) \propto p_s(z)$ if and only if the redshift that corresponds to comoving distance χ_s is within the boundaries of redshift bin i . A number of 10 redshift bins are used by default. The boundaries of the redshift bins are set such that each bin contains the same number of galaxies.

We adopt 20 angular frequency bins spaced logarithmically between $\ell_{\text{min}} = 50$ and $\ell_{\text{max}} = 3000$, and denote the characteristic angular frequency of a bin as $\bar{\ell}$. Within this range the noise properties of the cosmic shear field are still not too far in the non-Gaussian regime, allowing a more realistic theoretical estimation of the bispectrum and its covariance. Whether this number of angular frequency bins can reconstruct the angular frequency dependence of the bispectrum is tested, and 20 bins are found to be sufficient for our requirements on precision. This is also expected since the bispectrum is rather featureless as a function of angular frequency.

3.2. Bispectrum and its covariance

We show the modeling of B_{GGG} and its covariance in this section. We will only consider the tomographic bispectrum at redshift bins satisfying $z_i < z_j$ and $z_i < z_k$, which already ensures an elimination of B_{III} and B_{GII} systematics in our case.

Applying Limber's equation, it can be shown that the tomographic convergence bispectrum can be written as a projection of the three-dimensional matter bispectrum $B_\delta(k_1, k_2, k_3; \chi)$ (see e.g. TJ04):

$$B_{\text{GGG}}^{(ijk)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) = \int_0^{\chi_{\text{hor}}} d\chi \frac{W^{(i)}(\chi)W^{(j)}(\chi)W^{(k)}(\chi)}{\chi^4} B_\delta\left(\frac{\bar{\ell}_1}{\chi}, \frac{\bar{\ell}_2}{\chi}, \frac{\bar{\ell}_3}{\chi}; \chi\right). \quad (23)$$

To compute B_δ , we employ the fitting formula by Scoccimarro & Couchman (2001), which is based on hyper-extended perturbation theory (Scoccimarro & Frieman 1999). A comparison of this formula with the halo model results can be found in Takada & Jain (2003a,b).

Estimating the bispectrum covariance is often done within a flat-sky spherical harmonic formalism (Hu 2000, Hu00 hereafter), which suffers – at least formally – from drawbacks since its basis functions Y_{lm} are only defined for discrete angular frequency values and full sky coverage. In Joachimi et al. (2009) another approach exclusively based on the two-dimensional Fourier formalism was constructed, which we use here.

The bispectrum covariance is a six-point correlation function which can be expanded into its connected parts as outlined in e.g. Bernardeau et al. (2002a). As argued in TJ04, for angular scales $\ell \leq 3000$, the term which is a triple of two-point functions still

dominates. Keeping only this term, the bispectrum covariance reads

$$\text{Cov}\left(B_{\text{GGG}}^{(ijk)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3), B_{\text{GGG}}^{(lmn)}(\bar{\ell}_4, \bar{\ell}_5, \bar{\ell}_6)\right) = \frac{(2\pi)^3}{A \bar{\ell}_1 \bar{\ell}_2 \bar{\ell}_3 \Delta \bar{\ell}_1 \Delta \bar{\ell}_2 \Delta \bar{\ell}_3} \Lambda^{-1}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) \times \left(\bar{P}^{(il)}(\bar{\ell}_1) \bar{P}^{(jm)}(\bar{\ell}_2) \bar{P}^{(kn)}(\bar{\ell}_3) \delta_{\bar{\ell}_1 \bar{\ell}_4} \delta_{\bar{\ell}_2 \bar{\ell}_5} \delta_{\bar{\ell}_3 \bar{\ell}_6} + 5 \text{ perms.}\right), \quad (24)$$

in which $\Delta \bar{\ell}_i$ is the bin width of the angular frequency bin with typical value $\bar{\ell}_i$, and $\bar{P}^{(ij)}(\bar{\ell})$ is the observed power spectrum which contains the intrinsic ellipticity noise (e.g. Kaiser 1992; Hu 1999; Joachimi et al. 2008):

$$\bar{P}^{(ij)}(\bar{\ell}) = P^{(ij)}(\bar{\ell}) + \delta_{ij} \frac{\sigma_\epsilon^2}{2\bar{n}_i}, \quad (25)$$

where \bar{n}_i is the galaxy number density in redshift bin i . The term $\Lambda(\ell_1, \ell_2, \ell_3)$ is defined as

$$\Lambda(\ell_1, \ell_2, \ell_3) \equiv \begin{cases} \left\{ \frac{1}{4} \sqrt{2\ell_1^2 \ell_2^2 + 2\ell_1^2 \ell_3^2 + 2\ell_2^2 \ell_3^2 - \ell_1^4 - \ell_2^4 - \ell_3^4} \right\}^{-1} & \text{if } |\ell_1 - \ell_2| < \ell_3 < \ell_1 + \ell_2, \\ 0 & \text{else.} \end{cases} \quad (26)$$

When $|\ell_1 - \ell_2| < \ell_3 < \ell_1 + \ell_2$ is satisfied, $\Lambda^{-1}(\ell_1, \ell_2, \ell_3)$ is the area of a triangle with side lengths ℓ_1, ℓ_2 and ℓ_3 . We use the Eisenstein & Hu (1998) transfer function to evaluate the linear three-dimensional matter power spectrum, and the Smith et al. (2003) fitting function for the nonlinear power spectrum.

Note that there is no intrinsic ellipticity noise in the observed bispectrum, since the galaxy intrinsic ellipticity distribution is assumed to be skewless.

Under the assumptions of a compact survey geometry and scales much smaller than the extent of the survey area, (24) provides a bispectrum covariance that naturally incorporates the scaling with survey size, is not restricted to integer angular frequencies, and allows for any appropriate binning. In terms of Fisher information, the result given by this approach and the Hu00 one agree to high accuracy (Joachimi et al. 2009).

3.3. Toy intrinsic-shear alignment model

In this section we present a toy model for generating GGI systematics. Since the physical generation of intrinsic-shear alignments concerns nonlinear growth of structure and complex astrophysical processes which are not easy to quantify, a realistic model is not yet available. Current simulations involving baryonic matter also have some way to go before they can simulate the generation of the GGI systematics reliably.

Up to now there has not been any attempt to measure GGI and GII in galaxy surveys. Semboloni et al. (2008) studied these systematics using ray-tracing simulations. They provided fits in real space to projected GII and GGI signals, but the results are still too crude to lead to sufficient constraints on an intrinsic-shear alignment model.

This situation emphasizes the importance of a method intended to control intrinsic-shear alignment to be model-independent, especially at the three-point level. Since this is the case for the nulling technique, for this work we only require a simple model for $B_{\text{GGI}}^{(ijk)}$ which satisfies the characteristic redshift dependence and leads to a reasonable bias.

Based on the observation that the lensing bispectrum expression (23) comes directly from (12) and the definition of

the tomography bispectrum (14), we link $B_{\text{GGI}}^{(ijk)}$ also to a three-dimensional bispectrum $B_{\delta,\delta\delta}$ via

$$B_{\text{GGI}}^{(ijk)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) = \int_0^{\chi_{\text{hor}}} d\chi \frac{p_s^{(i)}(\chi) W^{(j)}(\chi) W^{(k)}(\chi)}{\chi^4} B_{\delta,\delta\delta}\left(\frac{\bar{\ell}_1}{\chi}, \frac{\bar{\ell}_2}{\chi}, \frac{\bar{\ell}_3}{\chi}; \chi\right). \quad (27)$$

Similar to $B_{\delta}(k_1, k_2, k_3)$ which is given by

$$\langle \tilde{\delta}(\mathbf{k}_1, \chi) \tilde{\delta}(\mathbf{k}_2, \chi) \tilde{\delta}(\mathbf{k}_3, \chi) \rangle = (2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B_{\delta}(k_1, k_2, k_3; \chi), \quad (28)$$

$B_{\delta,\delta\delta}$ is defined via

$$\langle \tilde{\delta}_1(\mathbf{k}_1, \chi) \tilde{\delta}(\mathbf{k}_2, \chi) \tilde{\delta}(\mathbf{k}_3, \chi) \rangle = (2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B_{\delta,\delta\delta}(k_1, k_2, k_3; \chi), \quad (29)$$

where $\tilde{\delta}_1(\mathbf{k})$ is the three-dimensional density field which is responsible for the intrinsic alignment, and it satisfies

$$\tilde{\kappa}_1^{(i)}(\boldsymbol{\ell}) = \int_0^{\chi_{\text{hor}}} d\chi p_s^{(i)}(\chi) \tilde{\delta}_1\left(\frac{\boldsymbol{\ell}}{\chi}, \chi\right), \quad (30)$$

The definition of both $\tilde{\kappa}_1^{(i)}$ and $\tilde{\delta}_1$ originates from the deterministic part of galaxy intrinsic ellipticity ϵ_1^{det} . We have assumed the existence of these underlying smooth fields. Similar quantities have been defined in [Joachimi & Bridle \(2009\)](#), see also [Hirata & Seljak \(2004\)](#) and [Schneider & Bridle \(2010\)](#). We would like to point out again that, although we introduce these quantities for the clarity of our model, we do not need them for the main purpose of this paper. What we need to model is the projected GGI bispectrum $B_{\text{GGI}}^{(ijk)}$.

Note that in (27), the weight for the lowest redshift bin i is the source distribution function $p_s^{(i)}$ which is zero outside redshift bin i , rather than the lensing weight $W^{(i)}$ which is a much broader function. Since $\tilde{\kappa}_1^{(i)}$ depends only on physical processes at redshift bin i and is inferred from ellipticity measurements in this bin, and $\tilde{\kappa}_G^{(j)}$ is linked to the three-dimensional matter density through the lensing weight $W^{(j)}$, this assignment of weight functions will ensure the correct redshift dependence of $B_{\text{GGI}}^{(ijk)}$.

When the redshift bins are not disjunct, however, the intrinsic alignment signal can no longer be associated with bin i . There will be two permutations in both the left-hand side of (16) and the right-hand side of (27), similar to the two-point case, e.g. Eq. (11) in [Hirata & Seljak \(2004\)](#).

The modeling of $B_{\delta,\delta\delta}$ is then a pure matter of choice. We build a simple three-dimensional GGI bispectrum with power-law dependence on both redshift z and spatial frequency k :

$$B_{\delta,\delta\delta}(k_1, k_2, k_3; \chi) := -\mathcal{A} B_{\delta}(k_{\text{ref}}, k_{\text{ref}}, k_{\text{ref}}; \chi(z_{\text{med}})) \left(\frac{1+z}{1+z_{\text{med}}}\right)^{r-2} \times \left\{ \left(\frac{k_1}{k_{\text{ref}}}\right)^{2(s-2)} + \left(\frac{k_2}{k_{\text{ref}}}\right)^{2(s-2)} + \left(\frac{k_3}{k_{\text{ref}}}\right)^{2(s-2)} \right\}, \quad (31)$$

where z_{med} is the median redshift of the whole survey, and \mathcal{A} , k_{ref} , r , s are free parameters. Among them the parameter k_{ref} is designed to be a characteristic wave number, whose value we set to be a weakly nonlinear scale of $10 h \text{Mpc}^{-1}$ here. The minus sign ensures that the contamination of GGI systematics leads to an underestimation of the GGG signal, as found by [Semboloni et al. \(2008\)](#).

Little is known about the redshift and angular scale dependence of $B_{\delta,\delta\delta}$. However one can roughly estimate how it compares to the $B_{\delta\delta\delta}$ signal. A linear alignment model suggests $\delta_I \propto \delta_{\text{lin}} \bar{\rho}(z) / ((1+z) D_+(z))$ (see e.g. [Hirata & Seljak 2004](#)), in which $\bar{\rho}(z)$ is the mean density of the universe, $D_+(z)$ is the growth factor, and δ_{lin} is the linear matter density contrast. Thus we have, very roughly, $\delta_I \propto (1+z)^3 \delta_{\text{lin}}$ which suggests $B_{\delta,\delta\delta} \propto (1+z)^3 B_{\delta\delta\delta}$. The linear alignment model assumes that the intrinsic alignment is linearly related to the local tidal gravitational field (e.g. [Catelan et al. 2001](#); [Hirata & Seljak 2004](#)). If this holds true, we also expect $B_{\delta,\delta\delta}$ to have a stronger angular scale dependence than $B_{\delta\delta\delta}$ since tidal gravitational interaction follows the inverse cube law rather than the inverse square law which gravity itself follows. For a Λ CDM model, in the weakly nonlinear regime where perturbation theory holds, the dependence of $B_{\delta\delta\delta}$ on $(1+z)$ has a negative power shallower than -4 , and the dependence on k has a power of around -2 . In this paper we choose $r = 0$, $s = 1$ as default. We also study the cases of $r = -2$, $r = 2$, and $s = 0$ whose results will be shown in Fig. 4 below.

As for the amplitude \mathcal{A} of the GGI signal, the only direct study up to now is [Semboloni et al. \(2008\)](#), which suggests an overall GGI/GGG ratio of 10% for a $z_m = 0.7$ survey for elliptical galaxies and few percent for a mixed sample of elliptical and spiral galaxies. In this paper we adjust \mathcal{A} such that the amplitude of the tomographic GGI bispectrum is limited to be within 10% of the amplitude of the lensing GGG signal, i.e. $\text{GGI/GGG} \lesssim 10\%$ at redshift bin combinations with $z_i \ll z_j$ and $z_i \ll z_k$ where the GGI signal is expected to be most significant. This leads to a relatively modest overall GGI/GGG ratio at percent level. We will show examples of the generated GGI and GGG signals in Fig. 3. As an order-of-magnitude estimate, one can also relate the GGI/GGG ratio to that of GI/GG by expanding three-point signals to couples of two-point signals using perturbation theory, in analogy to the [Scoccimarro & Couchman \(2001\)](#) fitting formula. For the case of $z_i \ll z_j \approx z_k$, the leading order terms would give that the GGI/GGG ratio approximates that of GI/GG evaluated at redshifts z_i and z_j . This suggests that our adopted GGI/GGG ratio is also consistent with available observational studies of the GI signal ([Mandelbaum et al. 2006, 2009](#); [Hirata et al. 2007](#); [Fu et al. 2008](#); [Okumura et al. 2009](#); [Okumura & Jing 2009](#)), although the results of these studies vary a lot according to different median redshift, color and luminosity of the selected galaxy sample.

4. Construction of nulling weights

As mentioned in Sect. 2.2, we would like to construct a single first-order weight function $T^{(ij)}(\chi)$ for each (i, j) combination which preserves the maximum of information. This can be seen as a constrained optimization problem. The constraining condition here is the nulling condition and the quantity to be optimized is the Fisher information after nulling. In JS08, several practical methods were developed to solve this optimization problem at the two-point level, and very good agreement was found among the different methods.

We adopt the simplified analytical approach as described in JS08, and reformulate it for three-point statistics here. For convenience we introduce the following notations:

$$\begin{aligned} \text{CovB}_{\ell_1, \ell_2, \ell_3}^{(ijk); lmn} &:= \\ \text{Cov}\left(B_{\text{GGG}}^{(ijk)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3), B_{\text{GGG}}^{(lmn)}(\bar{\ell}_4, \bar{\ell}_5, \bar{\ell}_6)\right); \end{aligned} \quad (32)$$

the covariance matrix $\text{Cov}Y$ of the nulled bispectra Y , whose elements are

$$\begin{aligned} \text{Cov}Y_{\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3}^{(ij); \bar{\ell}_4, \bar{\ell}_5, \bar{\ell}_6}{}^{lm} &:= \text{Cov}\left(Y^{(ij)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3), Y^{(lm)}(\bar{\ell}_4, \bar{\ell}_5, \bar{\ell}_6)\right) = \\ &\sum_{k=i+1}^{N_z} \sum_{n=i+1}^{N_z} \text{Cov}\left(B_{\text{GGG}}^{(ijk)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3), B_{\text{GGG}}^{(lmn)}(\bar{\ell}_4, \bar{\ell}_5, \bar{\ell}_6)\right) \\ &\times T^{(ij)}(\chi_k) T^{(lm)}(\chi_n) \chi'_k \chi'_n \Delta z_k \Delta z_n; \end{aligned} \quad (33)$$

a vector $\mathbf{B}_{,\mu}$ whose elements are partial derivatives of the bispectrum with respect to the cosmological parameter p_μ

$$B_{,\mu}{}^{ijk}{}_{\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3} := \frac{\partial B_{\text{GGG}}^{(ijk)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)}{\partial p_\mu}; \quad (34)$$

and a corresponding vector $\mathbf{Y}_{,\mu}$ for nulled bispectra Y , whose elements are

$$Y_{,\mu}{}^{ij}{}_{\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3} := \frac{\partial Y^{(ij)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)}{\partial p_\mu}. \quad (35)$$

Then the Fisher information matrix from the original bispectra can be written as (following TJ04)

$$F_{\mu\nu}^i = \mathbf{B}_{,\mu} \text{Cov}B^{-1} \mathbf{B}_{,\nu}, \quad (36)$$

and that from the nulled bispectra can be written as

$$F_{\mu\nu}^f = \mathbf{Y}_{,\mu} \text{Cov}Y^{-1} \mathbf{Y}_{,\nu}. \quad (37)$$

Here the matrix multiplication is a summation of all possible angular frequency combinations $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ and redshift bin combinations, (ijk) for the original bispectra and (ij) for the nulled bispectra. In (36) and (37), $\text{Cov}B^{-1}$ and $\text{Cov}Y^{-1}$ indicate the inverse of the covariance matrix. When the covariance is approximated by triples of power spectra, the covariance between two different angular frequency combinations $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) \neq (\bar{\ell}_4, \bar{\ell}_5, \bar{\ell}_6)$ is zero, see (24), which means that the covariance matrix is block diagonal. In this case the matrix inversion can be done separately for each block specified by an angular frequency combination $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$.

According to the idea of the simplified analytical approach, we consider the Fisher information on one cosmological parameter contained in bispectrum measures $B_{\text{GGG}}^{(ijk)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ with a single $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ combination and with redshift bin (i, j, k) combinations having common (i, j) indices. For every (i, j) combination we build nulling weights $T^{(ij)}$ which maximizes the nulled Fisher matrix using the method of Lagrange multipliers. Since here the nulled Fisher matrix receives contribution only from certain angular frequency and redshift combinations, we denote it as $F_o^{(ij)}$ to avoid ambiguity. $F_o^{(ij)}$ has only one component since only one cosmological parameter is taken into consideration. As only a single $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ combination is involved, we will omit the $\bar{\ell}$ -dependence in all variables in the rest of this subsection to keep a compact form.

Again for notational simplicity, we follow JS08 and introduce a vector notation as follows. For each (i, j) in consideration, let the values of the weights $T^{(ij)}(\chi_k)$ form a vector $\mathbf{T} = T_k$, and define another vector $\boldsymbol{\rho}$ and a matrix $\bar{\mathbf{C}}$ with elements

$$\rho_k := B_{,\mu}{}^{ijk} \chi'_k \Delta z_k, \quad (38)$$

$$\bar{C}_{kn} := \text{Cov}B^{(ijk, ijm)} \chi'_k \chi'_n \Delta z_k \Delta z_n. \quad (39)$$

Thus $F_o^{(ij)}$ can be expressed, according to (37), as

$$F_o^{(ij)} := \mathbf{Y}_{,\mu}{}^{(ij)} \text{Cov}Y^{-1}{}^{(ij, ij)} \mathbf{Y}_{,\mu}{}^{(ij)} = \frac{(\mathbf{T} \cdot \boldsymbol{\rho})^2}{\mathbf{T}^r \bar{\mathbf{C}} \mathbf{T}}. \quad (40)$$

We further define a vector \mathbf{f} with elements

$$f_k = \left(1 - \frac{\chi_i}{\chi_k}\right) \chi'_k \Delta z_k \quad (41)$$

to write the nulling condition (18) as

$$O^{(ij)} = \mathbf{T} \cdot \mathbf{f} = 0. \quad (42)$$

The problem of finding nulling weights \mathbf{T} which maximize $F_o^{(ij)}$ under the constraint given by the nulling condition can be solved with the method of Lagrange multipliers by defining a function

$$G := F_o^{(ij)} + \lambda O^{(ij)} = \frac{(\mathbf{T} \cdot \boldsymbol{\rho})^2}{\mathbf{T}^r \bar{\mathbf{C}} \mathbf{T}} + \lambda \mathbf{T} \cdot \mathbf{f} \quad (43)$$

with λ being the Lagrange multiplier, and setting the gradient of G with respect to \mathbf{T} to zero,

$$\nabla_{\mathbf{T}} G = 2\boldsymbol{\rho} \frac{(\mathbf{T} \cdot \boldsymbol{\rho})}{\mathbf{T}^r \bar{\mathbf{C}} \mathbf{T}} - 2 \bar{\mathbf{C}} \mathbf{T} \left(\frac{(\mathbf{T} \cdot \boldsymbol{\rho})}{\mathbf{T}^r \bar{\mathbf{C}} \mathbf{T}}\right)^2 + \lambda \mathbf{f} = 0. \quad (44)$$

The solution to this equation is (for more details see JS08)

$$\mathbf{T} = \mathcal{N} \left\{ \bar{\mathbf{C}}^{-1} \boldsymbol{\rho} - \frac{\mathbf{f}^r \bar{\mathbf{C}}^{-1} \boldsymbol{\rho}}{\mathbf{f}^r \bar{\mathbf{C}}^{-1} \mathbf{f}} \bar{\mathbf{C}}^{-1} \mathbf{f} \right\}, \quad (45)$$

with the normalization \mathcal{N} adjusted to give $|\mathbf{T}|^2 = 1$.

Apparently the thus constructed nulling weights depend on which $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ combination is considered and with respect to which cosmological parameter we optimize the information content. In this paper the default cosmological parameter to optimize is Ω_m , and we choose for each (i, j) combination the $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ combination which maximizes $F_o^{(ij)}$. However one needs to be aware that this serves only as a clear choice of a $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ combination and is not necessarily the best in terms of information preservation considering all angular frequency bins and all cosmological parameters.

To show which triangle shapes and sizes contain more information, we plot $F_o^{(ij)}$ against the $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ triangle shape and size for four typical (i, j) combinations in Fig. 1. In the left panel, the nulled information $F_o^{(ij)}$ contained in different triangles with a common shortest side length $\bar{\ell}_1 = 171$ is plotted against α , which is the angle opposite to $\bar{\ell}_1$. Due to our logarithmic binning in angular frequency, only eight $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ combinations with $\bar{\ell}_1 = 171$ can form triangles. One sees that the more elongated triangles (small α) contain much more Fisher information than the almost equilateral triangles (large α). The small separation between the 3rd and the 4th points from the left is caused by the degeneracy of different triangle shapes with respect to α , e.g. two equal and very long side lengths can result in the same value of α as two shorter side lengths with a length difference close to the length of the shortest side length. The right panel shows the distribution of the Fisher information contained in one $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ bin over the triangle size. When the redshift in consideration is higher, the peak of the information distribution moves to higher angular frequencies. The figure suggests that most information comes from high redshifts and small angular scales.

To explore the sensitivity of nulling weights on the choice of the cosmological parameter, we construct seven sets of weight

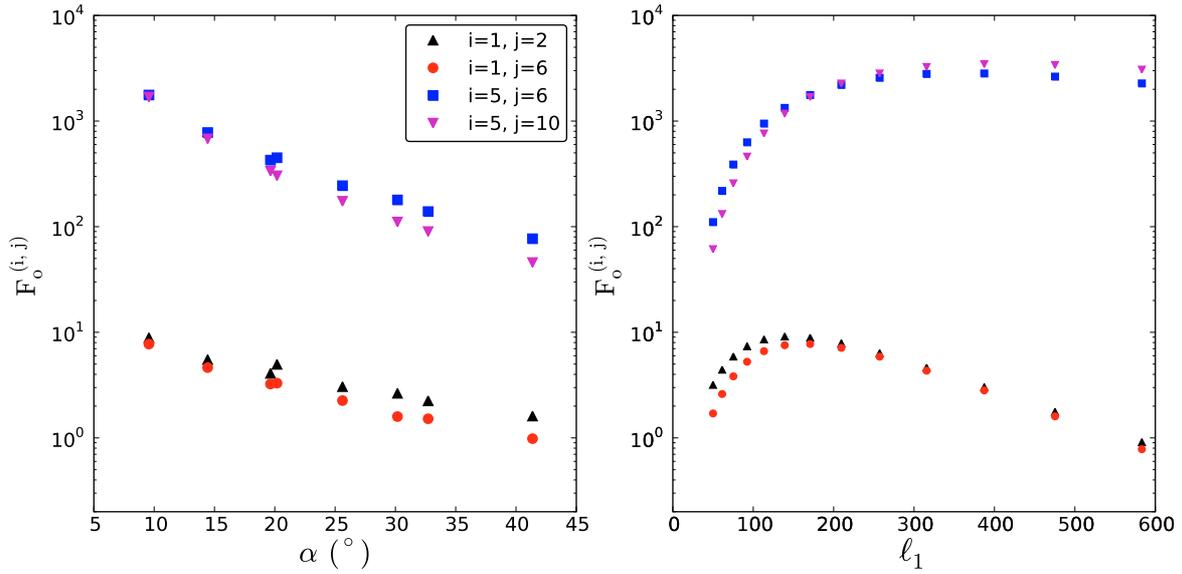


Fig. 1. Distribution of the nulled Fisher information as defined in (40) per $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ bin and per redshift bin combination among different angular frequency triangle shapes and sizes. Results for four redshift bin combinations (i, j) are presented. *Left panel:* distribution of the nulled Fisher information among different triangle configurations. We consider triangles with the common shortest side length $\bar{\ell}_1 = 171$ which corresponds to the 7th angular frequency bin. Due to our logarithmic binning and the constraint that the three side lengths must be able to form a triangle, only 8 such triangle configurations exist. Plotted is the nulled Fisher information contained in these 8 triangles against α , which is the angle opposite to the shortest side length in that triangle. Smaller α correspond to more elongated triangles, and larger α correspond to almost equilateral triangles. *Right panel:* distribution of the nulled Fisher information contributed by each $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ bin over different triangle sizes. A fixed triangle shape with $\bar{\ell}_1 : \bar{\ell}_2 : \bar{\ell}_3 = 1 : 3.64 : 4.52$ (corresponds to the leftmost points in the left panel) is chosen. The nulled Fisher information contained in one $(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3)$ bin is plotted against the shortest side length $\bar{\ell}_1$ of each triangle.

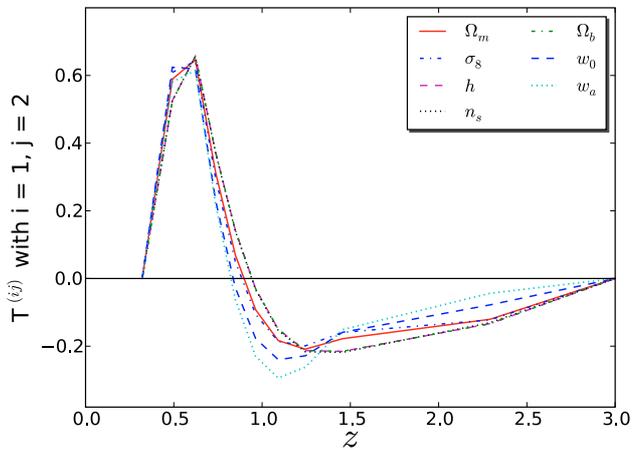


Fig. 2. Nulling weights $T^{(ij)}$ for redshift bins $i = 1, j = 2$ are plotted against the redshift value of the third redshift index k . Remarkable consistency is found between nulling weights optimized on different parameters, shown with different line styles.

functions, each optimizing the information content in terms of one parameter. For all (i, j) combinations we find that the nulling weights are not very sensitive to the choice of parameter. As an example, the weights for $(i, j) = (1, 2)$ are shown in Fig. 2. This result is rather surprising at first sight, since for different parameters the distribution of information (contained in the bispectrum) over redshift bins is quite different. However, such insensitivity suggests that the shapes of nulling weights are already strongly constrained under our construction scheme. One constraint is,

evidently, the nulling condition. Moreover, considering the fact that we optimize the nulling weights for each (i, j) combination with respect to the information content they preserve, we have already required the shapes of these first order nulling weights to be as smooth as possible.

The fact that these two conditions have already imposed strong constraints on the nulling weights also suggests that nulling weights can be robustly and efficiently constructed, i.e. it is not critical to construct the “best” nulling weights.

5. Performance of the nulling technique

5.1. GGI/GGG ratio

What the nulling technique “nulls” is the GGI signal B_{GGI} , so the GGI/GGG ratio is the most direct quantification of its performance. We plot the modeled GGI and GGG bispectra before and after nulling in Fig. 3. The original GGI signal is shown in the left panels by dashed lines. For comparison the GGG signals are shown as solid curves. The results are shown for equilateral triangle configurations for the convenience of presenting. One sees that when the redshift bin number j and/or k increase, the changes in GGG and GGI signals are different, which shows the expected different redshift dependence. For all redshift bin combinations the GGI signal is modeled to be subdominant to the GGG signal. In the nulled measures shown in the right panels, the GGI/GGG ratio is suppressed by a factor of 10 over all angular scales, which reflects the success of the nulling technique.

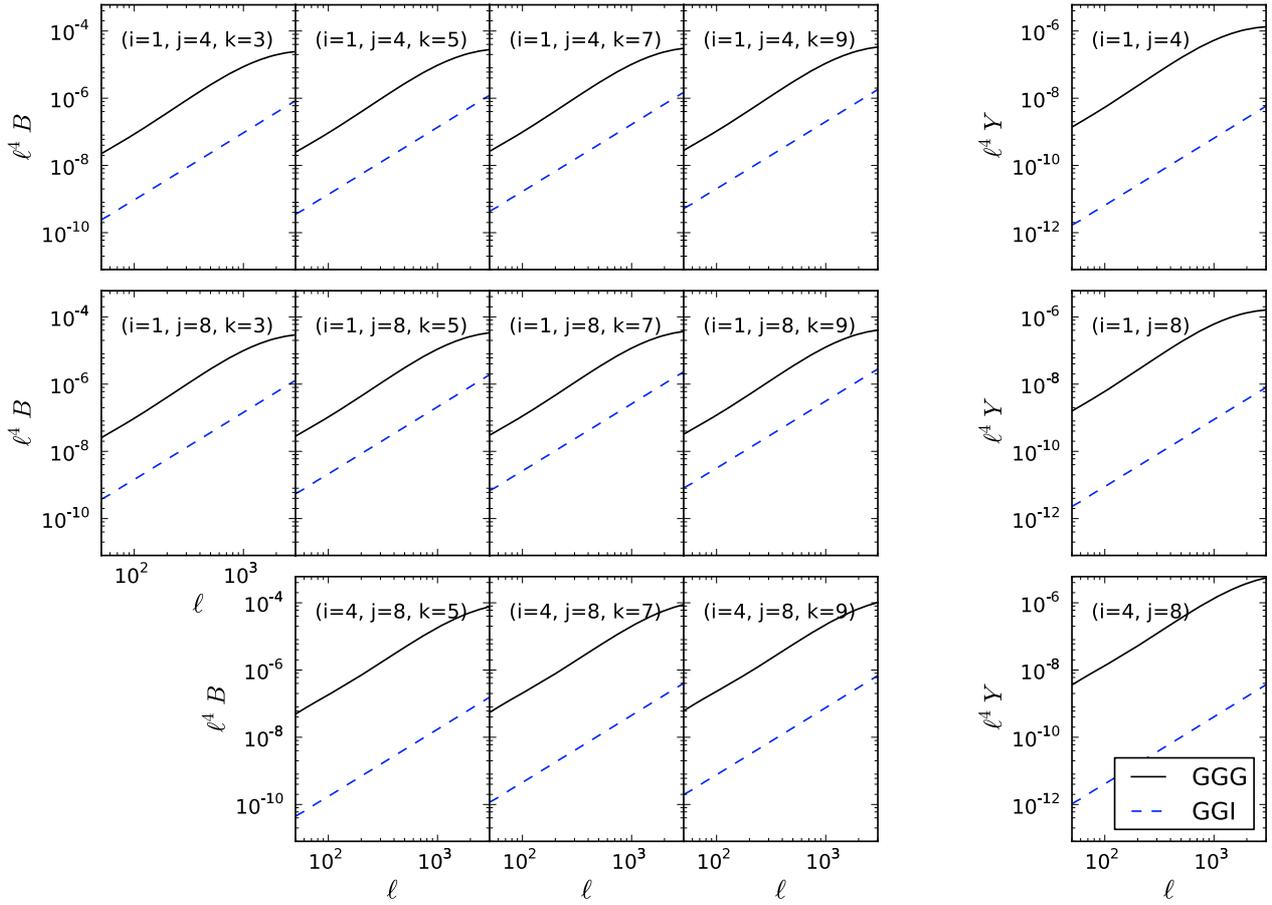


Fig. 3. Tomographic convergence bispectrum (GGG, solid curves) and intrinsic-shear alignment (GGI, dashed curves) for equilateral triangles are plotted against triangle side length. Measures both before (*left panel*) and after (*right panel*) applying the nulling technique (*B* and *Y* respectively) are shown for three typical redshift bin (*i, j*) combinations in the three rows.

5.2. Information loss and downweighting of systematics

We further evaluate the performance of the nulling technique by looking at the constraining power of cosmic shear bispectrum tomography on cosmological parameters, as well as the biases caused by the GGI systematics before and after nulling.

The full characterization of the bispectrum involves three angular frequency vectors which form a triangle. In some works concerning three-point statistics, only equilateral triangle configurations i.e. $\ell_1 = \ell_2 = \ell_3 = \ell$ are used for simplicity reasons (e.g. Pires et al. 2009). But as several authors have pointed out (e.g. Kilbinger & Schneider 2005; Bergé et al. 2010), only a low percentage of information is contained in equilateral triangles. Thus, to calculate the full information content, we use general triangle configurations but limit our calculation to triangles with three different side lengths, again for reasons of simplicity (for details see Appendix A).

We will use the figure of merit (FoM, Albrecht et al. 2006) to quantify the goodness of parameter constraints. Here the FoM for constraints in the parameter plane $p_\alpha - p_\beta$ is defined to be proportional to the inverse of the area of the parameter constraint ellipses:

$$\text{FoM}(p_\alpha, p_\beta) \equiv \left((\mathbf{F}^{-1})_{\alpha\alpha} (\mathbf{F}^{-1})_{\beta\beta} - (\mathbf{F}^{-1})_{\alpha\beta}^2 \right)^{-\frac{1}{2}}. \quad (46)$$

To compute biases, we adopt a method based on a simple extension of the Fisher matrix formalism (e.g. Huterer et al. 2006; Amara & Réfrégier 2008). Then one needs to define a bias vector \mathbb{B}^{GGI} which in our case reads:

$$\mathbb{B}_{v,i}^{\text{GGI}} = \mathbf{B}_{\text{GGI}} \text{CovB}^{-1} \mathbf{B}_{,v}, \quad (47)$$

$$\mathbb{B}_{v,f}^{\text{GGI}} = \mathbf{Y}_{\text{GGI}} \text{CovY}^{-1} \mathbf{Y}_{,v}, \quad (48)$$

with

$$\mathbf{B}_{\text{GGI}}^{(ijk)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) := B_{\text{GGI}}^{(ijk)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3), \quad (49)$$

$$\mathbf{Y}_{\text{GGI}}^{(ij)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3) := Y_{\text{GGI}}^{(ij)}(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3). \quad (50)$$

The bias of the parameter estimator \hat{p}_μ is given by the difference between its ensemble average and the fiducial value of the parameter p_μ^{fid} :

$$b_\mu = \langle \hat{p}_\mu \rangle - p_\mu^{\text{fid}} = \sum_v (\mathbf{F}^{-1})_{\mu\nu} \mathbb{B}_v^{\text{GGI}}. \quad (51)$$

The information content before and after nulling can be seen in Fig. 6. On the cost of increasing the error on each parameter to about twice its original value, GGI systematics are reduced to be within the original statistical error. The relative information loss

in terms of FoM can be found in Table 1. The constraints shown in Fig. 6 do not represent the best constraints obtainable from a cosmic shear bispectrum analysis since we consider only the triangles with angular scale $\bar{\ell}_1 \neq \bar{\ell}_2 \neq \bar{\ell}_3$. Also note that the nulling technique can in principle remove the GGI systematics completely. But as shown in Fig. 6, the systematics still cause some residual biases on cosmological parameters after nulling, due to the finite number of redshift bins. The GGI systematics will be reduced to a lower level when more redshift bins are available. We will discuss this further in the following subsection.

5.3. How many redshift bins are needed?

Analyzing the cosmic shear signal in a tomographic way was originally meant to maximize the information. For this purpose alone, a crude redshift binning will suffice (Hu 1999). However, to control intrinsic-shear alignment, which is a redshift-dependent effect, much more detailed redshift information is required (e.g. King & Schneider 2002; Bridle & King 2007; Joachimi & Schneider 2008). Thus, for a method intended to eliminate intrinsic-shear alignment, it is necessary to show its requirement on the redshift precision. In the case of nulling, detailed redshift information is not only needed for the method to be able to eliminate the bias, but also for the preservation of a reasonably large amount of information through the nulling process. JS08 examined the number of redshift bins required for the nulling technique in the two-point case, and showed that 10 redshift bins already ensure that parameters are still well-constrained after nulling.

To re-assess this problem at the three-point level, we consider two different situations to address the requirements coming from control of the intrinsic-shear alignment and preservation of the information content separately. In both cases we split the redshift range between $z = 0$ and $z = 3$ into 5, 10, 15, and 20 (only in the first situation) redshift bins, with the redshift bins split in a way that there is an equal number of galaxies in each bin.

First we consider a single cosmological parameter, Ω_m , to be free and study the biases introduced by the GGI signal on Ω_m both before and after nulling. We use only equilateral triangle configurations to reduce the amount of calculation. The results are shown in Fig. 4. Within the range of consideration, the ratio of the nulled and the original biases drops quickly with the increase of the number of redshift bins for all GGI models. For most of the models, 5 redshift bins seem to be not sufficient for the nulling technique to control the bias induced by GGI down to a percent level. Going from 5 redshift bins to 10 redshift bins is very rewarding in terms of bias reduction. However, we note that a decrease $|b_f/b_i|$ doesn't necessarily indicate a better performance of the nulling method, or generally speaking, of any method intended to control the intrinsic-shear alignment. One can see the reason for this by noticing that, it is the original unbinned GGI/GGG signal that is directly controlled by any of these methods. Between $|b_f/b_i|$ and the original unbinned GGI/GGG signal lies the binning process as well as the summation over angular frequency bins and redshift bins. Since the signs of the biases contributed by different angular frequencies and redshifts can be different, there can be bias cancellation during these processes. In another word, $|b_f/b_i|$ can depend on binning choices.

We then vary two cosmological parameters (Ω_m and σ_8) and investigate how the original and the nulled parameter constraints change with respect to the number of redshift bins available. For this case we use all triangle shapes to enable a comparison with results for two-point statistics.

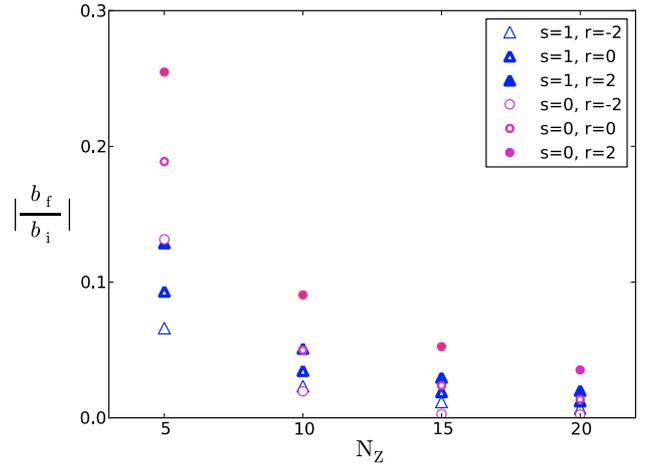


Fig. 4. Ratio of the nulled and the original biases for cosmological parameter Ω_m as a function of number of redshift bins N_z . Results for different GGI models are shown. Parameter s and r are the slopes of angular frequency and redshift dependence of our power-law model (31).

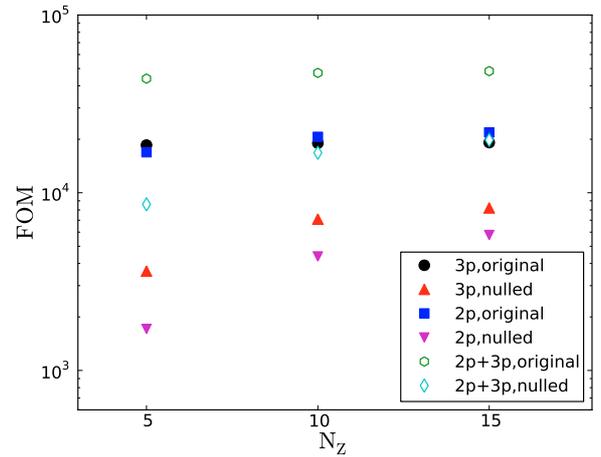


Fig. 5. Figure of merit (FoM) as defined in (46) in the $\Omega_m - \sigma_8$ plane as a function of number of redshift bins N_z . FoM from two-point (2p) measures, three-point measures (3p) and combined (2p+3p) are shown both before nulling (original) and after nulling (nulled).

Our result (Fig. 5) shows that a further increase of the number of redshift bins beyond 10 is not very rewarding in terms of information preservation as characterized by the FoM, in either 2p, 3p, or 2p+3p cases. This suggests, when the possibility of more redshift bins exists, the choice of redshift bin number should be based mainly on the requirement of bias reduction level in case of negligible photometric errors. When there are non-negligible photometric errors, however, the information loss will probably be more severe, as found by Joachimi & Schneider (2009) for the two-point case.

5.4. The nulling technique as a conditioned compression of data

The necessity of carrying out data compression in cosmology has long been recognized (e.g. Tegmark et al. 1997) and has been ever increasing due to the increasing size of the data sets.

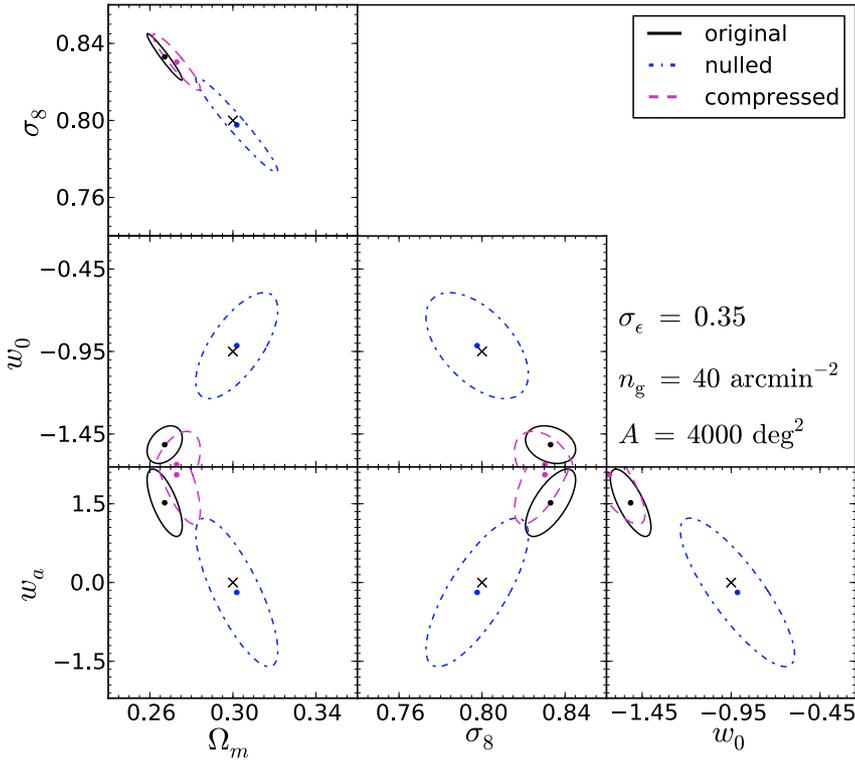


Fig. 6. Projected 1-sigma (68% CL) parameter constraints from cosmic shear bispectrum tomography. Hidden parameters are marginalized over. The black solid and blue dash-dotted ellipses correspond to the original constraints and those after nulling, respectively. The black cross in the center of each panel represents the fiducial values adopted for the parameters, and the distance from the center of one ellipse to the black cross reflects the bias caused by intrinsic-alignment GGI systematics on the corresponding parameter. As nulling can be seen as a linear data compression under the constraint of the nulling condition, we also plot the constraints and biases after an unconditioned linear data compression as magenta dashed ellipses for comparison (see Sect. 5.4).

In cosmic shear studies the survey area of next generation multi-color imaging surveys will be an order of magnitude larger than the current ones. The study of three-point statistics also implies a huge increase in the amount of data directly entering the Fisher-matrix/likelihood analysis, compared to the two-point case.

The basic principle of data compression is to reduce the amount of data while preserving most of the information. This is already naturally encoded in the nulling technique. If one keeps only the first-order weights for nulling, as we do in this paper, the nulling procedure reduces the number of data entries in each angular frequency bin from the number of redshift bin (i, j, k) combinations, to the number of (i, j) combinations, which means roughly from N_z^3 to N_z^2 . The nulling transformation is linear since the resulting nulled entry is a linear combination of k original entries weighted by the nulling weight (17). In the sense that an “optimum” set of nulling weights is constructed, the nulling technique also intends to preserve as much information as possible. But there is yet another additional constraining condition in the nulling procedure: the nulling condition (7), which largely confines the shape of the nulling weights by requiring the existence of at least one zero-crossing (see Fig. 2). In short, the nulling technique can be seen as a conditioned linear compression of data.

It is then interesting to know how much of the information loss during the nulling process actually comes from the nulling condition, and how much just comes from the fact that a data compression process is naturally involved in nulling. To explore this, we perform an unconditioned linear data compression, by simply ignoring the nulling condition in the whole nulling procedure i.e. dropping the Lagrange multiplier term in (43), but otherwise keeping the simplifications inherent to the analytical approach. The results are shown in Fig. 6. A summary of the FoM from the original and the nulled bispectrum measures as well as the compressed measures is shown in Table 1.

Table 1. Change of cosmic shear bispectrum statistical power after nulling (null) and linear data compression (compress).

	i	Null	Null/ i	Compress	Compress/ i
$\Omega_m - \sigma_8$	21455	4609	21.5%	12242	57.1%
$\Omega_m - w_0$	637	123	19.3%	428	67.2%
$\Omega_m - w_a$	145	33	23.0%	110	75.9%
$\sigma_8 - w_0$	434	87	20.0%	299	68.9%
$\sigma_8 - w_a$	101	26	25.4%	72	71.3%
$w_0 - w_a$	11.4	2.3	20.2%	8.0	70.2%

Notes. Presented are FoM on two-dimensional parameter planes between cosmological parameters Ω_m , σ_8 , w_0 and w_a . The cosmological parameters h , Ω_b and n_s are marginalized over. The second column is the FoM from the original bispectrum; the third and fifth columns are FoM from the nulled and the compressed measures, respectively; the fourth (sixth) column shows the percentage of the third (fifth) column compared to the first column, which reflects the relative information loss through the nulling (the unconditioned compression) procedure.

In contrast to nulling, an unconditioned linear compression does not eliminate the parameter bias, but increases or reduces some of them marginally. Regarding the parameter constraints, although the increase in the size of the ellipses is much less than in the case of nulling, around one third of the information in terms of FoM is lost through compression, which means that the amount of degradation in parameter constraints after compression is not negligible. This suggests that keeping only the first-order terms contributes to non-negligible information loss. To regain part of this information, one could add higher-order weights to the nulling procedure. But the difference between the nulled and the compressed FoM serves as an indication for the inevitable information loss through the nulling process, which is imposed by the nulling condition.

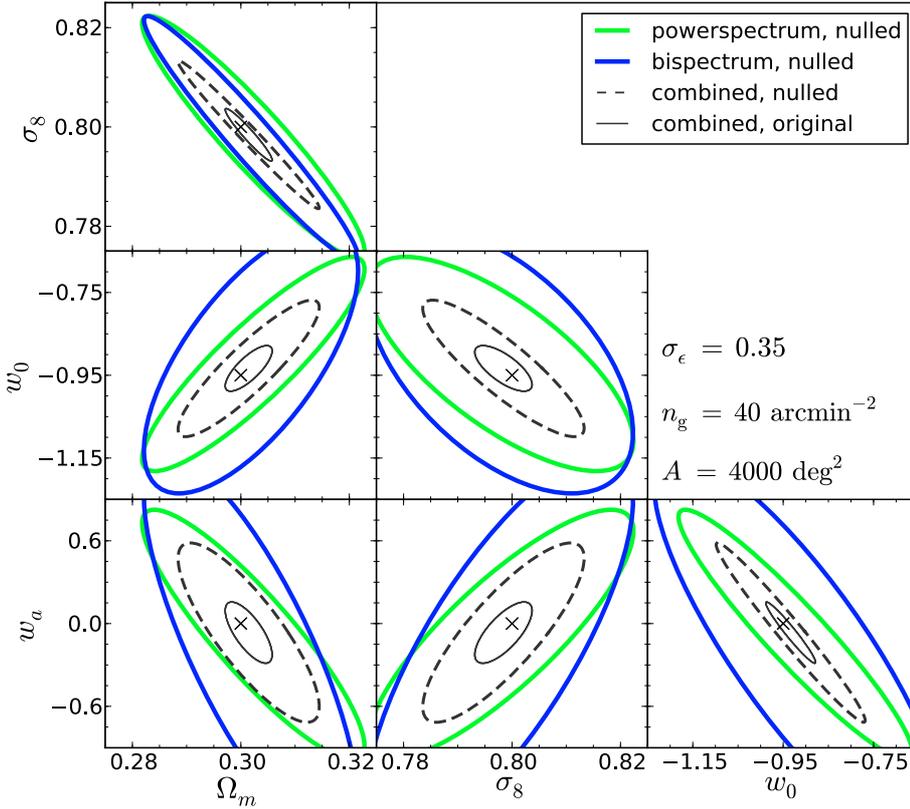


Fig. 7. The thick green (gray) solid, thick blue (black) solid and thin black dashed ellipses indicate 1-sigma (68% CL) parameter constraints from the nulled power spectrum measures, bispectrum measures, and combined. Hidden parameters are marginalized over. The distance from the center of an ellipse to the black cross reflects the nulled bias on the corresponding parameter. The original biases from bispectrum measures can be seen in Fig. 6. The thin black solid ellipses over-plotted on to the centers of the nulled combined constraint ellipses indicate the statistical power (68% CL) of combined constraints before nulling. Note the different ranges of parameters compared to Fig. 6.

Table 2. FoM before (“i”) and after (“f”) nulling and their ratio, using the cosmic shear power spectrum (2pt), bispectrum (3pt), and combined (2pt+3pt) analysis.

	2pt, i	3pt, i	2pt+3pt, i	2pt, f	2pt, i/f	3pt, f	3pt, i/f	2pt+3pt, f	2pt+3pt, i/f
$\Omega_m - \sigma_8$	21774	21455	86851	3297	15.1%	4609	21.5%	18555	21.4%
$\Omega_m - w_0$	1590	637	3806	236	14.8%	123	19.3%	600	15.8%
$\Omega_m - w_a$	517	145	872	69	13.3%	33	23.0%	121	13.9%
$\sigma_8 - w_0$	864	434	3832	132	15.2%	87	20.0%	488	17.2%
$\sigma_8 - w_a$	326	101	709	47	14.4%	26	25.4%	107	15.1%
$w_0 - w_a$	45	11	184	7.4	16.4%	2.3	20.2%	27	14.5%

5.5. Two-point and three-point constraints combined

Besides constraining cosmological parameters using three-point cosmic shear alone, we investigate the combined constraints from both two-point and three-point cosmic shear measures. The performance of the nulling technique on cosmic shear power spectrum tomography alone and the resulting constraints on cosmological parameters were presented in JS08. For consistency, we use the same setting for the cosmic shear power spectrum as described for the bispectrum in Sect. 3. In particular, we neglect photometric redshift errors, use only a limited range and number of ℓ -bins, and adopt a power-law intrinsic-shear alignment model with a form described by (36) in JS08 and a slope of 0.4. We have confirmed the consistency between our power spectrum and bispectrum codes with those used in Bergé et al. (2010). Our power spectrum code agrees also with iCosmo (Réfrégier et al. 2008).

Figure 7 shows the resulting constraint ellipses after nulling from the cosmic shear power spectrum analysis, the bispectrum analysis, and the two combined. To show how much information is lost during the nulling process, we overplot the original two- and three-point combined constraints on top of the nulled

constraint ellipses in Fig. 7, but center them on the corresponding nulled constraints by subtracting the bias difference before and after nulling. The information content in terms of FoM for each parameter pair is presented in Table 2.

One sees that the amount of information contained in bispectrum measures and power spectrum measures are indeed comparable. With bispectrum information added, typically three times better constraints in terms of FoM are achieved, both before and after nulling. This factor is smaller than the result in TJ04, although the same angular frequency range and the same set of 7 cosmological parameters are chosen for both studies. However a direct comparison is prohibited by different fiducial values adopted and different survey specification.

Through the nulling procedure, around 15% of the original information in terms of FoM is preserved in the two-point case, and around 20% in the three-point case. It is a bit higher in the three-point case, in accordance to the fact that a roughly $N_z^3 \rightarrow N_z^2$ compression is involved in the three-point case and a $N_z^2 \rightarrow N_z^1$ one in the two-point case, while this fact is due to the summation over one redshift bin index during the nulling procedure (the same trend is evident in Fig. 5). The information loss

is considerable, but it is a price to pay for a model-independent method. As we have discussed in the previous subsection, the difference between the information loss through the nulling and the unconditioned compression procedures represents the inevitable loss of information through nulling. However, this difference is less than 50% in the considered three-point case. The other information loss is due to the simplifications we adopted in this study, including using only the first-order weights, and discarding the measures with two or three equal redshift bins. A further detailed consideration of these aspects can regain part of the lost information. Another simplification we have made in the three-point case is to use only triangles with three different angular frequencies. This reduces both the original and the nulled information contained in the three-point measures. However, this simplification can be easily removed with a careful distinction of all cases.

Also notice that, the dependence of number of possible bispectrum modes, i.e. triangles, on the maximum angular frequency ℓ_{\max} is roughly ℓ_{\max}^3 , while that of power spectrum modes is roughly ℓ_{\max}^2 . For this study $\ell_{\max} = 3000$ is chosen. If reliable information on smaller angular scales can be obtained, the three-point statistics will possibly give us more information than the two-point statistics.

6. Conclusion

In this study we developed a method to control the intrinsic-shear alignment in three-point cosmic shear statistics by generalizing the nulling technique. We showed that the generalization of the nulling technique to three-point statistics is quite natural, providing a model-independent method to reduce the intrinsic-shear alignment signals (GGI and GII) in comparison to the lensing GGG signal.

To test the performance of the nulling technique, we assumed a fictitious survey with a setup typical of future multi-color imaging surveys, and applied the nulling technique to the modeled bispectra with intrinsic-shear alignment contamination. The lensing bispectra (GGG) was computed based on perturbation theory, while the GGI signal was modeled by a simple power-law toy model. We focused on the reduction of the GGI contaminant, since GII can be removed simply by not considering tomographic bispectra with two or three equal redshift bins.

The reduction of the intrinsic-shear alignment contamination at the three-point level by the nulling technique was demonstrated both in terms of the GGI/GGG ratio, and in terms of biases on cosmological parameters in the context of an extended Fisher matrix study. In terms of the GGI/GGG ratio, a factor of 10 suppression is achieved after nulling over all angular scales. Correspondingly, the biases on cosmological parameters are reduced to be less than or comparable to the original statistical errors. We studied the performance of the nulling technique when 5, 10, 15, or 20 redshift bins are available, and found that the performance on bias reduction, rather than how much information is preserved during the nulling procedure, depends more significantly on the number of redshift bins. In case one requires better control of intrinsic-shear alignment, more detailed redshift information allowing more redshift bins is the most direct way to go.

When dealing with real data, there is one further source of complication which we did not consider in this paper, that is the photometric redshift uncertainty. The photometric redshift uncertainty can be characterized by a redshift-dependent photometric redshift scatter and catastrophic outliers. Joachimi & Schneider (2009) studied the influence of photometric redshift

uncertainty on the performance of the nulling technique at the two-point level. They found that the photometric scatter places strong bounds on the remaining power to constrain cosmological parameters after nulling. The existence of catastrophic outliers, on the other hand, can lead to an incomplete removal of the intrinsic (II, III) alignments as well as the intrinsic-shear alignments (GI, GII, GGI). However, methods to control the photometric redshift uncertainty have been proposed. For example, recent studies concerning the problem of catastrophic outliers point to the solutions of either limiting the lensing analysis to $z < 2.5$ or by conducting an additional small-scale spectroscopic survey (Sun et al. 2009; Bernstein & Huterer 2009; Bordoloi et al. 2010).

As already demonstrated by JS08 in the two-point case, some information loss is inherent to the nulling procedure. For the setup of this paper we found that, in terms of FoM about 20% of the original information is preserved through the nulling procedure in the three-point case, and 15% in the two-point case. We further studied the source of such information loss by comparing the nulling technique to an unconditioned linear compression of the data, since the nulling procedure can be seen as a linear compression of data under the constraint of the nulling condition (7). We found that around one third of the original information is lost through an unconditioned compression of the data, suggesting that this situation can be improved by considering higher-order terms in the nulling and compression processes.

Results on parameter constraints from the two- and three-point cosmic shear statistics combined are also presented. The amount of nulled information contained in bispectrum measures and power spectrum measures are comparable. With bispectrum information added, typically three-times better constraints are achieved both before and after nulling, in terms of FoM.

Again, due to the large amount of information existing in the three-point cosmic shear field, one would certainly like to exploit it in the future. The nulling method we developed in this work solves a potentially severe problem hampering the use of three-point information, namely the intrinsic-shear alignment systematic. Our method works at the cost of a large information loss, which can hopefully be avoided by a future method of removing the intrinsic-shear alignment contaminants. But as the only completely model-independent method so far, the nulling technique can serve as a working method now and can provide a valuable cross check even with the availability of better methods.

Acknowledgements. We thank Joel Bergé and Masahiro Takada for the help in comparing power spectrum and bispectrum codes. We are also grateful to the anonymous referee for the helpful comments and suggestions. BJ acknowledges support by the Deutsche Telekom Stiftung and the Bonn-Cologne Graduate School of Physics and Astronomy. This work was supported by the RTN-Network ‘DUEL’ of the European Commission, and the Deutsche Forschungsgemeinschaft under the Transregional Research Center TR33 ‘The Dark Universe’.

Appendix A: Counting of triangles

A triangle is specified by six indices, i.e. three redshift bin indices $\{i, j, k\}$ and three angular frequency bin indices ℓ_1, ℓ_2, ℓ_3 . To ensure that we count each triangle configuration only once, we set the condition that $\ell_1 \leq \ell_2 \leq \ell_3$. Moreover, we would like the first index among $\{i, j, k\}$ in (36) to have the lowest redshift, i.e. $z_i < z_j$ and $z_i < z_k$, for the convenience of performing the nulling technique. The possible $\{i, j, k\}$ combinations under these constraints in the case of $N_z = 4$ are listed in Fig. A.1.

However, setting both conditions is problematic. Inspecting the definition of the tomographic bispectrum (14), one sees that

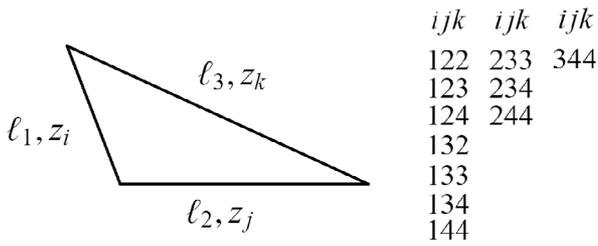


Fig. A.1. List of possible triangles (redshift bin combinations) with condition $z_i < z_j$ and $z_i < z_k$ when 4 redshift bins are available. An angular frequency combination satisfying $\ell_1 \leq \ell_2 \leq \ell_3$ is chosen. Note that the redshift indices and the angular frequencies are linked in pairs due to the definition of the tomographic bispectrum (14). In this paper a default of 10 redshift bins is assumed.

the redshift indices and the angular frequencies are linked in pairs, e.g. convergence κ in redshift bin i has angular frequency ℓ_1 , which is not desirable since the smallest angular scale does not necessarily correspond to the lowest redshift. To solve this problem, we perform nulling three times for each general angular frequency combination with $\ell_1 < \ell_2 < \ell_3$, swapping the redshift-angular scale correspondence in-between, thus allowing each redshift to be able to correspond to any angular frequency.

Note that the situation complicates a bit when two of the angular frequencies are equal, since then the swapping may lead to exactly the same configuration. To avoid this, we will restrict ourselves to three different angular frequencies. This can exclude a high percentage of possible configurations. In our case, i.e. 20 logarithmically spaced bins between $\ell_{\min} = 50$ and $\ell_{\max} = 3000$, 37% of the angular frequency combinations which can form a triangle have been excluded. However, this is only a technical complication which can be solved with a careful distinction of all cases. Since this paper is intended to be a proof of applicability of the nulling technique to three-point statistics, we defer the intricacies of accounting for all triangle configurations to future work.

References

Abdalla, F. B., Amara, A., Capak, P., et al. 2008, MNRAS, 387, 969
 Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, unpublished [arXiv:astro-ph/0609591]
 Amara, A., & Réfrégier, A. 2007, MNRAS, 381, 1018
 Amara, A., & Réfrégier, A. 2008, MNRAS, 391, 228
 Bacon, D. J., Refregier, A. R., & Ellis, R. S. 2000, MNRAS, 318, 625
 Bergé, J., Amara, A., & Réfrégier, A. 2010, ApJ, 712, 992
 Bernardeau, F., van Waerbeke, L., & Mellier, Y. 1997, A&A, 322, 1
 Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002a, Phys. Rep., 367, 1
 Bernardeau, F., Mellier, Y., & van Waerbeke, L. 2002b, A&A, 389, L28
 Bernstein, G., & Huterer, D. 2009, MNRAS, 1648
 Bordoloi, R., Lilly, S. J., & Amara, A. 2010, MNRAS [arXiv:0910.5735]
 Brainerd, T. G., Agustsson, I., Madsen, C. A., & Edmonds, J. A. 2009 [arXiv:0904.3095], ApJ, submitted
 Bridle, S., & Abdalla, F. B. 2007, ApJ, 655, L1
 Bridle, S., & King, L. 2007, New J. Phys., 9, 444
 Brown, M. L., Taylor, A. N., Hambly, N. C., & Dye, S. 2002, MNRAS, 333, 501
 Catelan, P., Kamionkowski, M., & Blandford, R. D. 2001, MNRAS, 320, L7

Croft, R. A. C., & Metzler, C. A. 2000, ApJ, 545, 561
 Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
 Fu, L., Semboloni, E., Hoekstra, H., et al. 2008, A&A, 479, 9
 Heavens, A., Refregier, A., & Heymans, C. 2000, MNRAS, 319, 649
 Heymans, C., & Heavens, A. 2003, MNRAS, 339, 711
 Heymans, C., Brown, M., Heavens, A., et al. 2004, MNRAS, 347, 895
 Heymans, C., White, M., Heavens, A., Vale, C., & van Waerbeke, L. 2006, MNRAS, 371, 750
 Hirata, C. M., & Seljak, U. 2004, Phys. Rev. D, 70, 063526
 Hirata, C. M., Mandelbaum, R., Ishak, M., et al. 2007, MNRAS, 381, 1197
 Hu, W. 1999, ApJ, 522, L21
 Hu, W. 2000, Phys. Rev. D, 62, 043007
 Hui, L. 1999, ApJ, 519, L9
 Hui, L., & Zhang, J. 2002, unpublished [arXiv:astro-ph/0205512]
 Huterer, D., Takada, M., Bernstein, G., & Jain, B. 2006, MNRAS, 366, 101
 Jain, B., & Seljak, U. 1997, ApJ, 484, 560
 Jarvis, M., Wilson, G., & Jain, B. 2004, MNRAS, 352, 338
 Jing, Y. P. 2002, MNRAS, 335, L89
 Joachimi, B., & Bridle, S. L. 2009 [arXiv:0911.2454], A&A, submitted
 Joachimi, B., & Schneider, P. 2008, A&A, 488, 829
 Joachimi, B., & Schneider, P. 2009, A&A, 507, 105
 Joachimi, B., Schneider, P., & Eifler, T. 2008, A&A, 477, 43
 Joachimi, B., Shi, X., & Schneider, P. 2009, A&A, 508, 1193
 Kaiser, N. 1992, ApJ, 388, 272
 Kaiser, N., Wilson, G., & Luppino, G. A. 2000, unpublished [arXiv:astro-ph/0003338]
 Kilbinger, M., & Schneider, P. 2005, A&A, 442, 69
 King, L., & Schneider, P. 2002, A&A, 396, 411
 King, L. J. 2005, A&A, 441, 47
 King, L. J., & Schneider, P. 2003, A&A, 398, 23
 Mackey, J., White, M., & Kamionkowski, M. 2002, MNRAS, 332, 788
 Mandelbaum, R., Hirata, C. M., Ishak, M., Seljak, U., & Brinkmann, J. 2006, MNRAS, 367, 611
 Mandelbaum, R., Blake, C., Bridle, S., et al. 2009 [arXiv:0911.5347], MNRAS, submitted
 Munshi, D., Valageas, P., van Waerbeke, L., & Heavens, A. 2008, Phys. Rep., 462, 67
 Okumura, T., & Jing, Y. P. 2009, ApJ, 694, L83
 Okumura, T., Jing, Y. P., & Li, C. 2009, ApJ, 694, 214
 Peacock, J. A., Schneider, P., Efstathiou, G., et al. 2006, ESA-ESO Working Group on Fundamental Cosmology, Tech. rep.
 Pen, U.-L., Zhang, T., van Waerbeke, L., et al. 2003, ApJ, 592, 664
 Pires, S., Starck, J., Amara, A., Réfrégier, A., & Teyssier, R. 2009, A&A, 505, 969
 Réfrégier, A., Amara, A., Kitching, T., & Rassat, A. 2008, A&A, submitted [arXiv:0810.1285]
 Schneider, M. D., & Bridle, S. 2010, MNRAS, 402, 2127
 Schneider, P. 2003, A&A, 408, 829
 Schneider, P. 2006, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro, ed. G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, & J. Wambsganss, 269
 Scoccimarro, R., & Couchman, H. M. P. 2001, MNRAS, 325, 1312
 Scoccimarro, R., & Frieman, J. A. 1999, ApJ, 520, 35
 Semboloni, E., Heymans, C., van Waerbeke, L., & Schneider, P. 2008, MNRAS, 388, 991
 Smail, I., Ellis, R. S., & Fitchett, M. J. 1994, MNRAS, 270, 245
 Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311
 Spergel, D. N., Bean, R., Doré, O., et al. 2007, ApJS, 170, 377
 Sun, L., Fan, Z., Tao, C., et al. 2009, ApJ, 699, 958
 Takada, M., & Jain, B. 2003a, ApJ, 583, L49
 Takada, M., & Jain, B. 2003b, MNRAS, 344, 857
 Takada, M., & Jain, B. 2004, MNRAS, 348, 897
 Takada, M., & White, M. 2004, ApJ, 601, L1
 Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, ApJ, 480, 22
 van Waerbeke, L., Bernardeau, F., & Mellier, Y. 1999, A&A, 342, 15
 Van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, A&A, 358, 30
 Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. 2000, Nature, 405, 143
 Zhang, P. 2008, unpublished [arXiv:0811.0613]