

Galaxy and mass assembly: Redshift space distortions from the clipped galaxy field

F. Simpson,^{1,2,*} C. Blake,³ J. A. Peacock,² I. K. Baldry,⁴ J. Bland-Hawthorn,⁵ A. F. Heavens,⁶ C. Heymans,² J. Loveday,⁷ and P. Norberg⁸

¹*ICC, University of Barcelona (UB-IEEC), Marti i Franques 1, 08028 Barcelona, Spain*

²*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, United Kingdom*

³*Centre for Astrophysics & Supercomputing, Swinburne University of Technology, P.O. Box 218, Hawthorn, Victoria 3122, Australia*

⁴*Astrophysics Research Institute, Liverpool John Moores University, IC2, Liverpool Science Park, 146 Brownlow Hill, Liverpool, L3 5RF United Kingdom*

⁵*Sydney Institute for Astronomy, School of Physics, University of Sydney, New South Wales 2006, Australia*

⁶*Imperial Centre for Inference and Cosmology, Imperial College, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, United Kingdom*

⁷*Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, United Kingdom*

⁸*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, United Kingdom*

(Received 19 May 2015; published 27 January 2016)

We present the first cosmological measurement derived from a galaxy density field subject to a “clipping” transformation. By enforcing an upper bound on the galaxy number density field in the galaxy and mass assembly survey (GAMA), contributions from the nonlinear processes of virialization and galaxy bias are greatly reduced. This leads to a galaxy power spectrum which is easier to model, without calibration from numerical simulations. We develop a theoretical model for the power spectrum of a clipped field in redshift space, which is exact for the case of anisotropic Gaussian fields. Clipping is found to extend the applicability of the conventional Kaiser prescription by more than a factor of 3 in wave numbers, or a factor of 30 in terms of the number of Fourier modes. By modeling the galaxy power spectrum on scales $k < 0.3 \text{ hMpc}^{-1}$ and density fluctuations $\delta_g < 4$ we measure the normalized growth rate $f\sigma_8(z = 0.18) = 0.29 \pm 0.10$.

DOI: [10.1103/PhysRevD.93.023525](https://doi.org/10.1103/PhysRevD.93.023525)

I. INTRODUCTION

The spatial distribution of galaxies encodes a wealth of information relating to the composition and evolution of the Universe. The apparent positions of galaxies in redshift space offers a glimpse into both the density and velocity perturbations associated with dark matter. These in turn are influenced by a number of phenomena in fundamental physics, such as the mass of the neutrino and the nature of gravity. However two key factors have thus far restricted our view: (a) the advanced stages of gravitational collapse are highly unpredictable and (b) the uncertainty associated with galaxy bias, defined as the manner in which the galaxy distribution reflects the dark matter distribution. In Fourier space, conventional analyses impose a maximum wave number beyond which the data are considered unpredictable and are simply discarded. For example, despite utilizing numerical simulations to calibrate the nonlinear power spectrum, recent studies of redshift space distortions typically truncate the power spectrum at $k < 0.2 \text{ hMpc}^{-1}$ [1] or exclude galaxy pairs closer than $25 \text{ h}^{-1} \text{ Mpc}$ [2]. While most

of the nonlinear behavior is successfully disposed of, so too is much of the cosmological information.

A number of different methods have been proposed to allow smaller clustering scales to be exploited. A phenomenological model has been developed by Kwan *et al.* [3], and expanded by Linder and Samsing [4], to model the power spectrum based on fits to numerical simulations. Reid *et al.* [5] extracted information from small-scale clustering in Baryon Oscillation Spectroscopic Survey (BOSS) using a model for the halo occupation distribution. While achieving a significant increase in precision, this technique also relies heavily upon calibration from numerical simulations. Alternatively, various local transformations have been explored as a means of reducing the influence of nonlinearities, such as Gaussianization [6,7] and the logarithmic transformation [8]. In [9] it was shown that simply enforcing a maximum density could greatly increase the number of Fourier modes that could be modeled with the standard set of theoretical tools. Recently it has also been found to enhance the observational signature in models of modified gravity which invoke a screening mechanism [10]. This clipping technique serves as the focus of the present work.

*fergus2@icc.ub.edu

Our inability to model the small scale power spectrum is not the only source of information loss. More fundamentally, the onset of nonlinear gravitational collapse degrades the amount of information held by the power spectrum [11]; this loss occurs for any non-Gaussian field. To extract some of this missing information, we can either perform additional measurements, such as higher-order statistics, or manipulate the field prior to evaluating the power spectrum. Previous attempts to extract cosmological information from the spatial distribution of galaxies beyond the conventional two point statistics include measuring three-point statistics [12–14], Minkowski functionals [15], and the shapes of voids [16]. Local transformations also appear promising in restoring this information to the power spectrum (see for example [8, 17–19]). Carron and Szapudi [19] present a transformation that is optimized for extracting information from the power spectrum of a Poisson-sampled field. Our approach differs slightly: instead of optimizing the extraction of all information from the observable galaxy density field ρ_g , we seek to selectively extract only the *predictable* information from the field. At the highest values of ρ_g both the nonlinear structure of the dark matter and the environmental impacts on galaxy formation present formidable obstacles in interpreting its value. Conversely, regions closer to the mean density are expected to behave in a more predictable and robust manner. The action of clipping preserves the location of galaxy clusters as useful information while discarding information relating to the precise value of their density contrast.

Clipping is already known to be a highly effective technique for improving the theoretical modeling of the galaxy bispectrum [20] and power spectrum [9], when applied to fields in real space. Before we can apply this technique to data from real surveys, it must first be verified in redshift space. It would also be desirable to develop a deeper understanding of why it has been successful. These are two of the goals of this paper. The third is to apply clipping to the GAMA survey in order to obtain a low-redshift measurement of the normalized growth rate $f\sigma_8$.

In Sec. II we review the theoretical background associated with the two point statistics of clipped fields and consider its extension to anisotropic fields. This theoretical framework is placed into a cosmological context in Sec. III, where we develop a model for the form of the clipped galaxy power spectrum. In Sec. IV we apply clipping to simulated dark matter and galaxy fields in redshift space, with the results illustrated in Sec. V. The GAMA data set is introduced in Sec. VI, while the main results of this work are presented in Sec. VII before our concluding remarks in Sec. VIII.

II. STATISTICAL PROPERTIES OF CLIPPED FIELDS

Clipping is a local transformation characterized by the application of a saturation value δ_0 to a scalar field $\delta(\mathbf{x})$ such that

$$\begin{aligned}\delta_c(\mathbf{x}) &= \delta_0 & (\delta(\mathbf{x}) > \delta_0) \\ \delta_c(\mathbf{x}) &= \delta(\mathbf{x}) & (\delta(\mathbf{x}) \leq \delta_0),\end{aligned}\quad (1)$$

yielding the clipped field $\delta_c(\mathbf{x})$. In this section we explore the statistical properties of $\delta_c(\mathbf{x})$, with particular attention paid to its autocorrelation

$$\xi_c(\mathbf{r}) \equiv \langle [\delta_c(\mathbf{x}) - \bar{\delta}_c][\delta_c(\mathbf{x} + \mathbf{r}) - \bar{\delta}_c] \rangle, \quad (2)$$

and corresponding power spectrum $P_c(\vec{k})$. We note that the clipping transformation induces a nonzero mean in $\delta_c(x)$, which is why must be careful to specify the more general form of the autocorrelation function, as defined by (2). In the analysis of cosmological fields the subtraction of the mean is conventionally omitted from this definition, since the mean is usually zero by construction.

We begin by reviewing the special case where $\delta(\mathbf{x})$ is an isotropic Gaussian field before generalizing to anisotropic and higher-order fields.

A. Isotropic Gaussian fields

For the case of Gaussian fields we may invoke Price's theorem [21, 22] to evaluate the two-point statistics associated with the clipped field $\delta_c(\mathbf{x})$ in terms of the original correlation function $\xi(r)$:

$$\frac{\partial \xi_c(r)}{\partial \xi(r)} = \iint \frac{\partial g[\delta_1]}{\partial \delta_1} \frac{\partial g[\delta_2]}{\partial \delta_2} p(\delta_1, \delta_2, r) d\delta_1 d\delta_2, \quad (3)$$

where $g[\delta]$ is the local transformation defined by (1), $\xi_c(r)$ is the correlation function of the transformed field, and $p(\delta_1, \delta_2, r)$ is the joint probability distribution for a Gaussian process

$$p(\delta_1, \delta_2, r) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho(r)^2}} \exp\left[\frac{2\rho(r)\delta_1\delta_2 - \delta_1^2 - \delta_2^2}{2\sigma^2[1-\rho(r)^2]}\right]. \quad (4)$$

The functional derivative of the clipping transformation is unity below the threshold and zero above the threshold. This simple behavior permits an analytic solution of the clipped correlation function $\xi_c(r)$, which is given by [9]

$$\xi_c(r) = f_c^2 \xi(r) + \sigma^2 \sum_{n=1}^{\infty} \left[\frac{\xi(r)}{\sigma^2} \right]^{n+1} C_n(u_0), \quad (5)$$

where σ^2 is the variance of the field prior to clipping, u_0 is the normalized threshold value $u_0 \equiv \delta_0/\sqrt{2}\sigma$, f_c is the fraction of the field which lies below the threshold

$$f_c = \frac{1}{2} [1 + \text{erf}(u_0)], \quad (6)$$

and $C_n(x)$ is the distortion coefficient

$$C_n(x) = \frac{H_{n-1}^2(x)}{\pi 2^n (n+1)!} e^{-2x^2}, \quad (7)$$

where $H_n(x)$ is the Hermite polynomial of order n . Beyond a scale-independent reduction in amplitude, clipping induces a distortion in the shape of the correlation function. However provided the slope of the spectral power is not too steep ($|n| \lesssim 3$), and the clipping remains weak ($u_0 \gtrsim 1$), the leading order correction makes only a small contribution to the resultant power. Furthermore, terms at higher values of n decay rapidly.

B. Anisotropic Gaussian fields

For the more general case of anisotropic Gaussian fields we may express the two-point correlation function $\xi(r, \theta)$ in terms of both the pair separation r and the orientation vector $\vec{\theta}$. In deriving the expression given by (5) we made use of the joint probability distribution for a Gaussian process as given by (4), which is not directly applicable to anisotropic fields. However we can proceed by applying this single-parameter transformation separately at each fixed value of $\vec{\theta}$, which in itself constitutes a single-parameter Gaussian process. Since the variance of each subspace is the same, $\sigma^2 = \xi(r=0, \vec{\theta})$ for all orientations, the transformation maintains the same functional form for all values of $\vec{\theta}$. Therefore the expression (5) can be generalized to anisotropic fields.

Transforming to Fourier space leaves us with the expression for the clipped power spectrum

$$P_c(\vec{k}) = f_c^2 P(\vec{k}) + \sigma^2 \sum_{n=1}^{\infty} C_n(u_0) \left[\frac{P(\vec{k})}{\sigma^2} \right]^{*(n+1)}, \quad (8)$$

where the notation $*n$ represents a self-convolution of order n . In practice it is computationally more straightforward to evaluate the higher order terms using powers of the correlation function, rather than performing multiple convolutions of the power spectrum. Further details of this calculation, as applied to mock cosmological density fields, can be found in Appendix B.

In line with the definition of the correlation function (2), our definition of the power spectrum in (8) is specified in terms of the mean subtracted field. For a clipped Gaussian field the mean is given by

$$\bar{\delta}_c = \delta_0(1 - f_c) - \frac{\sigma}{\sqrt{2\pi}} \exp(-\delta_0^2/2\sigma^2). \quad (9)$$

In practice, it is not critical to account for this constant offset since it only contributes to the power spectrum at $k=0$.

C. Second order anisotropic fields

In order to gain insight into how higher order terms respond to clipping, we repeat the procedure above using the square of a Gaussian random field, $e(x) \equiv \delta^2(x) - \sigma^2$. The correlation function of this second order field, clipped at $e = \delta_0$, is well approximated by [9]

$$\xi_c(r, e) = a_2 \xi(r, e) + \mathcal{O}\left(\frac{\xi^2(r, e)}{\sigma^2}\right), \quad (10)$$

$$a_2 = \frac{1}{\pi} [\sqrt{\pi} \text{erf}(u'_0) - 2u'_0 e^{-u_0'^2}]^2, \quad (11)$$

$$u'_0 \equiv \frac{\sqrt{\delta_0 + \sigma^2}}{\sqrt{2}\sigma}, \quad (12)$$

where σ refers to the standard deviation associated with the original Gaussian field, and a_2 quantifies the amplitude of the power spectrum relative to the original field. Following the same line of reasoning given in the previous subsection, we can generalize this result to the anisotropic case:

$$P_c(\vec{k}, e) \simeq a_2 P(\vec{k}, e). \quad (13)$$

As before, the clipped two-point statistics of a second order field maintain the same shape as the unclipped case, at least for weak transformations. Note that when clipping at equivalent thresholds, $u_0 = u'_0$, the higher order field is subject to a significantly stronger suppression of its two point statistics than for the Gaussian case, and this trend strengthens with yet higher order fields.

D. Hybrid fields

Practical applications of clipping will inevitably involve the superposition of a Gaussian field with other components that contaminate the desired signal, particularly where the amplitude of the field is large. In this scenario, clipping can assist in extracting the power spectrum associated with the original Gaussian field.

Consider a hybrid field $h(x)$ that is a linear combination of a Gaussian field $\delta(x)$, a higher order field $\delta^2(x)$, and a nuisance field δ^X , which characterises some unknown departure from the model:

$$h(x) = \alpha_1 \delta(x) + \alpha_2 (\delta^2(x) - \sigma^2) + \delta^X(x). \quad (14)$$

Upon clipping at a given threshold δ_0 , and provided the nuisance field is constant where $h(x) < \delta_0$, we can subtract the mean to remove any residual contribution from δ^X

$$h_c(x) = \alpha_1 \delta_c + \alpha_2 (\delta_c^2 - \sigma^2), \quad (15)$$

where the component fields δ_c and δ_c^2 are now also clipped fields. This result is important as it shows that we can

cleanly remove any trace of our nuisance field δ^X . In most practical applications the recovery will be imperfect, as the nuisance field is likely to vary outside of the clipped region. However this is a much better state of affairs than the conventional approach—be it perturbation theory or a model of galaxy bias—where we assume that the extra terms missing from our model (as specified by δ^X) vanish everywhere. With clipping, we can now make the much more reasonable approximation that δ^X only vanishes where $h(x)$ is small.

The two component fields δ_c and δ_c^2 each experience their own distinct thresholds, which may be found by solving (15) with the condition $h_c(x) = \delta_0$. The power spectrum of $h_c(x)$ is given by

$$P_c(k) = \alpha_1^2 P_{cL}(k) + 2\alpha_1\alpha_2 P_{c12}(k) + \alpha_2^2 P_{c22}(k), \quad (16)$$

where $P_{cL}(k)$ and $P_{c22}(k)$ denote the power spectra associated with the clipped $\delta(x)$ and $\delta^2(x)$ fields respectively. The cross spectrum $P_{c12}(k)$ vanishes in the limit of a high threshold, but becomes increasingly prominent as the threshold is lowered. In order to estimate $P_{c12}(k)$, we may decompose it as

$$\langle \delta_c^{(1)} \delta_c^{(2)} \rangle = \langle r_1 r_2 \rangle - \langle r_2 \delta^{(1)} \rangle - \langle r_1 \delta^{(2)} \rangle, \quad (17)$$

where the residual fields are defined as $r_i \equiv \delta^{(i)} - \delta_c^{(i)}$. Provided the clipping is weak, $\delta_0 > \sigma$, the residual fields are closely related $r_2 \simeq 2t_0 \frac{\alpha_2}{\alpha_1} r_1$, where t_0 is the threshold experienced by the Gaussian field. This may be reexpressed in the form

$$P_{c12}(k) \simeq - \left[2t_0(f_c^{-1} - 1)P_L(k) + \frac{\alpha_1 - \alpha_2}{2t_0\alpha_2} P_{22}(k) \right]. \quad (18)$$

Since $P_{c12}(k)$ may be expressed as a linear combination of the first and second order power spectra, we can now rewrite (16) in the form

$$P_c(k) \simeq a_L P_L(k) + a_2 P_{22}(k), \quad (19)$$

where a_L and a_2 are the apparent amplitudes of the original linear and second order spectra. This result helps explain why the simple model used in Simpson *et al.* [9] was particularly successful at reproducing the clipped dark matter power spectrum, without explicitly accounting for the cross-spectrum $P_{c12}(k)$.

Adopting this higher order model, rather than relying on the linear solution from Sec. II B, holds two advantages. First of all weaker clipping thresholds can be used, allowing the power spectrum to maintain a high amplitude. In addition, the inclusion of a higher order term potentially allows the degeneracy between linear bias and σ_8 to be lifted. The disadvantage of this approach is the difficulty in

estimating a_2 , which could either be calibrated from simulations, or simply treated as an additional free parameter.

III. GALAXY DENSITY FIELDS

Galaxy redshift surveys continue to develop an intricate mosaic of the low redshift Universe. By convolving the pointlike positions of galaxies with a suitable kernel, a continuous density field can be generated. These galaxy fields are heavily influenced by both nonlinear structure and galaxy bias, which have proved highly challenging to model. In each case, it is the highest density regions which are particularly troublesome, and this motivates the application of clipping. The three dimensional nature of the data ensures that clipping can be applied very efficiently. Selecting a threshold that affects only $\sim 1\%$ of the field's volume typically leads to a reduction in large scale power by a factor of 2. Maps which are two-dimensional projections, such as those derived from a photometric survey, could also be subject to clipping but a greater proportion of the area would need to be clipped in order to achieve the same degree of suppression.

In this section we explore the consequences of clipping a galaxy density field in redshift space. We shall work in the distant-observer approximation such that all line-of-sight displacements may be considered parallel.

A. Galaxy bias

If we model the fractional overdensity of galaxies $\delta_g \equiv \rho_g/\bar{\rho}_g - 1$ as an arbitrary function of the local dark matter density [23],

$$\delta_g = \sum_{k=0}^{\infty} b_k \frac{\delta_m^k}{k!}, \quad (20)$$

then applying the clipping transformation suppresses higher order b_k terms in the same way that higher order terms in perturbation theory are suppressed. The simplest extension to the linear bias model would be the introduction of b_2 , which for the case of Gaussian dark matter fluctuations leaves us with a hybrid field as defined in (14). With a sufficiently low clipping threshold, the linear bias parameter b_1 dominates such that the clipped galaxy power spectrum is highly insensitive to the initial value of b_2 . This linearization process was demonstrated explicitly in Fig. 4 of [9]. Even in the context of more complex models of galaxy bias, such as those induced by tidal fields [24], we expect a similar behavior. Nonlinear contributions to the galaxy bias still predominantly arise in regions where δ is large, and these are the regions suppressed by the clipping transformation.

There is however a fundamental limit on how much we can shield ourselves from the influence of the highest density regions, and this stems from the estimation of the

mean number density. In defining the fractional overdensity, $\delta_g \equiv \rho_g/\bar{\rho}_g - 1$, the quantity $\bar{\rho}_g$ necessarily incorporates the abundance of galaxies across the whole volume, prior to clipping. Unlike dark matter perturbations where the total particle number is conserved, no such restriction applies to galaxy bias. For example if baryonic effects reduce the abundance of galaxies in clusters such that the total galaxy count across the survey volume is lowered by a small fraction y , then the inferred amplitude of fluctuations across the rest of the volume are overestimated by

$$\hat{\delta}_g \simeq \delta_g^t(1 + y) + y, \quad (21)$$

where δ_g^t is the true fractional perturbation.

B. Redshift space distortions: Linear model

Whilst the true spatial distribution of galaxies is expected to be statistically isotropic, their redshift-inferred distances receive an additional displacement due to their peculiar velocities, generating a statistically anisotropic configuration. This permits a measurement of $f\sigma_8$, where the logarithmic linear growth rate is given by $f \equiv d \ln \delta / d \ln a$, and σ_8 defines the amplitude of linear perturbations.

An additional source of anisotropic clustering arises from inaccuracies in the assumed geometry of the Universe, which is required when converting the observed values of angles and redshifts into a Euclidean framework. This can potentially generate false measurements of the growth rate [25]. In this work we shall consider the background expansion to be fixed to a flat Λ CDM model with $\Omega_m = 0.27$ unless specified otherwise.

Clipping in redshift space carries additional complications. The small scale velocity dispersion associated with the ‘‘fingers of God’’ effect will tend to move galaxies out of the high density peaks and potentially into a surrounding region that lies below the clipping threshold. We should therefore expect the removal of nonlinear effects to be less efficient in redshift space. The velocity dispersion also causes the power spectrum to steepen at larger wave numbers along the line of sight. This strong spectral slope enhances the relative amplitude of the higher order terms in (8), so these should not be neglected.

On all but the largest scales, the real space cosmological density field at low redshifts is not well described by a Gaussian field. It may instead be considered as a superposition of a Gaussian component $\delta^{r(1)}$ and an extra field δ^X representing the conglomeration of all nonlinear corrections.

$$\delta_m^r = \delta_m^{r(1)} + \delta_m^X. \quad (22)$$

The δ^X field is largest where the linear approximation is most strongly violated—both from the truncation of higher order terms in perturbation theory and more fundamentally

from the assumption of a single-valued and curl-free velocity field. The matter density field is traced by the galaxy number density field, which again may be decomposed into a Gaussian component, with a linear bias factor b , and a residual term such that

$$\delta_g^r = b\delta_m^{r(1)} + \delta_g^X. \quad (23)$$

Now moving to redshift space, the Gaussian component is described by the Kaiser model [26], which relates the real space linear density perturbations with those in redshift space, which we couple with a Lorentzian model of the velocity dispersion:

$$\delta_g^{s(1)}(k, \mu) = \frac{b + f\mu^2}{\sqrt{1 + k^2\sigma_v^2\mu^2/2}} \delta_m^{r(1)}(k), \quad (24)$$

where $\mu \equiv k_{\parallel}/k$ is defined as the relative fraction of the wave vector that extends along the line of sight. The pairwise velocity dispersion σ_v effectively smooths the field along the line of sight. Theoretically this quantity is given by

$$\sigma_v^2 = \frac{f^2 H_0^2}{3H^2(z)\pi^2} \int P_{\theta\theta}(k) dk, \quad (25)$$

where $P_{\theta\theta}(k)$ is the velocity power spectrum [27]. In practice we shall treat σ_v as a free parameter, due to the uncertain behavior of nonlinear motions.

The clipped galaxy field δ_{cg}^s can be represented as the sum of a clipped Gaussian field and a residual term δ^X . Therefore the resulting power spectrum may be expressed as the sum of three terms, the two autocorrelations and the cross-spectrum

$$P_c^s(k) = P_{cL}^s(k) + 2P_{c1x}^s(k) + P_{xx}^s(k). \quad (26)$$

Given that the nonlinear terms encapsulated by δ^X dominate the clustering statistics at larger values of δ , it experiences a much greater loss of power than the linear component. Therefore we should expect that after clipping the first term remains the dominant contribution to the total power over a broader range of scales. Our simplest model for the clipped galaxy power spectrum is therefore encapsulated by

$$\begin{aligned} P_c^s(k, \mu) &\simeq P_{cL}^s(k), \\ &= \mathcal{C}\{P_g^s(k, \mu); u_0\}, \end{aligned} \quad (27)$$

where \mathcal{C} denotes the transformation defined by (8), u_0 is the normalized clipping threshold experienced by the Gaussian field as given by (6), and

$$P_g^s(k, \mu) = \frac{b^2(1 + \beta\mu^2)^2}{1 + (k\sigma_v\mu)^2/2} P_L(k), \quad (28)$$

where $P_g^s(k, \mu)$ is the Gaussian contribution to the galaxy power spectrum in redshift space, $P_L(k)$ is the linear matter power spectrum in real space, and the anisotropy parameter $\beta \equiv f/b$ quantifies the level of anisotropy in the galaxy power spectrum induced by linear velocity perturbations.

For the case of a hybrid field such as the one defined in (14), then as clipping is applied, and nonlinear contaminations are suppressed relative to the linear contributions, we should expect the recovered value of β to be closer to the theoretical value for a given k_{\max} ; alternatively, we should be able to achieve the same level of systematic error in β at a higher k_{\max} . The actual level of error and/or smallest scale to probe must be determined empirically using simulations, as described in the following section.

C. Redshift space distortions: Nonlinear model

Our base model is defined by the set of four parameters $\{\beta, b\sigma_8, \sigma_v, a_L\}$, and relies upon the recovery of the linear matter power spectrum. However the linear power spectrum decays rapidly towards higher wave numbers, and by $k \sim 0.5 \text{ hMpc}^{-1}$ is typically an order of magnitude lower than the contribution from the one-loop correction to the power spectrum. Therefore despite the suppression of higher order terms, the inclusion of a suppressed one-loop term substantially improves the model for the real space power spectra of matter and galaxies [9]. Further motivated by the results of Sec. II D, we introduce an extended model with the additional parameter ϵ , which accounts for a higher order contribution to the power spectrum

$$P_c^s(k, \mu) = a_L \frac{b^2(1 + \beta\mu^2)^2}{1 + (k\sigma_v\mu)^2/2} [P_L(k) + \epsilon P_{1\text{loop}}(k)]. \quad (29)$$

The parameter a_L is the coefficient of the linear power which contributes to the clipped power spectrum, as defined in the model of (19). In general the value of a_L cannot be determined *a priori*. It may be evaluated empirically by considering the fractional change in amplitude of the large scale clustering of the field,

$$a_L = \lim_{r \rightarrow \infty} \frac{\xi_c(r)}{\xi(r)}. \quad (30)$$

IV. SIMULATIONS

In order to test our theoretical models, we construct mock density fields for both dark matter and galaxies. These are derived from numerical simulations, and transformed into redshift space using the distant observer approximation. In this section we summarize our methods for generating and modeling the power spectra associated

with clipped cosmological fields, and for the estimation of their covariance matrices.

A. Number density fields in redshift space

For our mock dark matter field we utilize the $z = 0$ snapshot from the Horizon Run 2 simulation [28], which consists of 6000^3 particles within a periodic box of size $7200 \text{ h}^{-1}\text{Mpc}$. The amplitude of linear perturbations is $\sigma_8 = 1/1.26$, with a matter density $\Omega_m = 0.26$. The redshift-space density field is defined by considering a 1920^3 grid using a nearest grid point (NGP) scheme, where the particles are displaced along one axis in accordance with their peculiar velocity. This leaves us with a grid size of $3.75 \text{ h}^{-1}\text{Mpc}$ corresponding to a Nyquist frequency of $k_N = 0.84 \text{ hMpc}^{-1}$.

The mock galaxy catalogues are taken from [29], which applies a semianalytic model to the halo merger trees of the Millennium-I simulation [30]. Following [20] we use the two snapshots at $z = 0$ and $z = 0.687$ in order to explore different amplitudes of matter perturbations and growth rates. With $\sigma_8 = 0.9$, the $z = 0$ snapshot possesses a slightly higher amplitude of fluctuations than the Horizon simulation. Applying a stellar mass cut of $\log_{10}(M_*/M_\odot h) \geq 9$ leaves us with a distribution resembling that found in our GAMA sample [31]. Number density fields are formed on both 128^3 and 256^3 grids across the $500 \text{ h}^{-1}\text{Mpc}$ box, with Nyquist frequencies of $k_N = 0.8 \text{ hMpc}^{-1}$ and 1.6 hMpc^{-1} respectively. The standard deviations are 3.1 for the coarse grid and 6.0 for the high resolution field.

In order to evaluate the true value of β we need the linear bias parameter. The ratio of the real space galaxy power spectrum with that of the Millennium simulation's dark matter field gives $b = 1.125 \pm 0.01$ at $z = 0$, on the largest available scales. At $z = 0.687$ this increases to $b = 1.29 \pm 0.02$, while the growth rate is well approximated by $f(z) \approx \Omega_m^{0.55}(z)$.

B. Threshold selection

After constructing each number density field, we apply the clipping transformation defined by (1). In order to select a suitable threshold value δ_0 we require an appropriate metric for defining the strength of clipping. For the case of a Gaussian field, the normalized threshold $u_0 \equiv \delta_0/\sqrt{2}\sigma$ provides a natural measure for this in relation to the standard deviation of the field. However when working with fields that are highly non-Gaussian, we do not know *a priori* what impact a given threshold will have. Some degree of iteration is therefore required in order to reach the desired reduction in power, characterized by a_L .

When working with the simulations, rather than quoting absolute values of the threshold δ_0 , we choose to work in terms of the fraction of mass (or galaxies) removed. This way fields with larger fluctuations naturally adopt higher thresholds δ_0 .

The disadvantage of a stronger (lower) threshold is a larger drop in the amplitude of the power spectrum, which in turn reduces the maximum wave number available before the shot noise contributions dominate. Very strong thresholds also induce a large contribution from the cross-power $P_{12}(k)$, as given by (18), further reducing the power. The optimal choice of threshold is therefore one that adequately suppresses contributions to the power from nonlinear structure and nonlinear bias, while maintaining a relatively high amplitude of the linear power spectrum. As shown in Figs. 6, 7, and 8 of [20], the decay of the higher order term rapidly outpaces the contribution from the linear power spectrum, such that a factor of 2 reduction in linear power is sufficient to eliminate approximately 80% of the nonlinear power.

For each field we explore a range of threshold values, since it is important to verify that different thresholds generate consistent parameter constraints. Performing a likelihood analysis that combines the power spectra derived from different thresholds may further improve parameter constraints. However an estimation of the covariance between the different spectra is beyond the scope of this work, here we shall only consider the analysis of each threshold separately.

C. Methods

The theoretical model for the clipped power spectrum is evaluated with the following procedure

- (i) Apply the redshift space model defined in (28) to determine $\hat{P}^s(k, \mu)$ from the linear power spectrum evaluated by camb [32] combined with the chosen parameter values of b , f , and σ_v .
- (ii) Interpolate the power spectrum onto a 3D grid matching the specifications of the NGP lattice generated from the simulations.
- (iii) Transform to the 3D correlation function $\xi(\sigma, \pi)$.
- (iv) Evaluate the clipped 3D correlation function $\xi_c(\sigma, \pi)$ using (5). We find it sufficient to truncate the series expansion at $n_{\max} = 4$ in order to establish sub-percent accuracy. This step is discussed further in Appendix A.
- (v) Apply the inverse Fourier transform to determine the expected 3D power spectrum as a function of wave number magnitude k and orientation μ , $P_c^s(k, \mu)$. The power spectrum is assigned to linearly spaced bins, with widths of $\Delta k = 0.025$, $\Delta \mu = 0.2$. We maintain the same binning scheme throughout this work. Note that the mean wave number contributing to the bin is in general at higher values of k than the bin centre, due to the abundance of modes within a spherical shell dk scaling as k^2 .

A four-dimensional likelihood grid is constructed, $\mathcal{L}(P_c(k, \mu) | a_L, \sigma_v, \beta, b\sigma_8)$, where the parameter set relates to the model given by (27) and (28).

With the clipped power spectrum alone as the only source of information, we would have little knowledge of how strong the applied clipping has been, and therefore a_L is poorly constrained. However the drop in power relative to the original (unclipped) power spectrum provides us with some extra information that can assist in constraining the range of a_L values. We therefore make use of the fractional loss of power experienced at $k < 0.1 h\text{Mpc}^{-1}$ and $\mu < 0.6$, to provide some external information for the value of a_L . At higher values of k and μ the larger contributions from nonlinear structure mean that they experience a greater drop in power.

We assign flat priors of $0 < a_L < 1$, $0 < \beta < 2$, $0 < b\sigma_8 < 2$, and for the velocity dispersion we use a broad Gaussian prior $\sigma_v/\sqrt{2} = 300 \pm 300 \text{ km s}^{-1}$, as defined in (28). Our results are largely insensitive to the particular choice of priors.

D. Covariance estimation

The large volume of the dark matter field in the Horizon Run 2 simulation, $(7200 h^{-1}\text{Mpc})^3$, permits a direct estimation of the 70×70 covariance matrix associated with the $P(k, \mu)$ bins by considering the covariance of power spectra evaluated from 350 subvolumes, each of size $(960 h^{-1}\text{Mpc})^3$. To evaluate the covariance matrix we employ the Ledoit-Wolf shrinkage estimator [33], following the prescription

$$\Sigma = \delta^* F + (1 - \delta^*) S, \quad (31)$$

where Σ is our estimated covariance matrix, and δ^* is the shrinkage constant. S is the sample covariance matrix, defined as the ensemble average over $N = 350$ subvolumes

$$S_{ij} = \langle \Delta P_i \Delta P_j \rangle, \quad (32)$$

where ΔP_i represents the deviation in power in the i th $P(k, \mu)$ bin from the sample mean. The shrinkage target F is defined in terms of the sample covariance

$$F_{ij} = \begin{cases} S_{ij}, & (i = j) \\ \bar{r} \sqrt{S_{ii} S_{jj}}, & (i \neq j) \end{cases} \quad (33)$$

$$\bar{r} = \frac{2}{(N-1)N} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{S_{ij}}{\sqrt{S_{ii} S_{jj}}}, \quad (34)$$

where $N = 350$. The shrinkage constant δ^* is estimated using publicly available code [34]. We emphasize that this approach yields a considerably improved estimate of the covariance matrix than simply using the sample covariance S alone.

For the galaxy sample associated with the Millennium simulation, the volume is considerably smaller and therefore a direct estimate of the covariance matrix would be

prohibitively noisy. Instead we only estimate the diagonal terms of the covariance matrix, again using the variance of subvolumes. The estimation of off diagonal terms is determined by using the dark matter covariance as a template, such that

$$C_{ij} = \text{corr}(i, j) \sqrt{C_{ii} C_{jj}}, \quad (35)$$

where $\text{corr}(i, j)$ is the correlation matrix from the dark matter power spectrum.

E. Shot noise estimation

For the conventional unclipped power spectrum, the discrete nature of sources leads to an additional contribution of power, which if we assume to be Poissonian in nature is given by $P_s = \bar{n}^{-1}$. As clipping smooths the field above the threshold value, the shot noise contribution is *reduced*. Approximating the noise field as Gaussian allows us to utilize (6) to estimate $P_s \approx f_V^2 \bar{n}^{-1}$, where f_V is the fraction of the volume of the field lying below the clipping threshold. In this work the typical volume fraction is of the order $\sim 1\%$, and therefore the correction to the shot noise is negligible.

The power spectra of clipped fields are highly robust to changes in the number density of sampled points [9]. The only noticeable consequence appears to be that fields with higher shot noise possess noisier power spectra and can therefore not utilize as wide a range of wave numbers.

V. RESULTS FROM SIMULATIONS

First we present the power spectra from the dark matter and galaxy fields, at different clipping thresholds, before reviewing the results of the likelihood analysis.

A. Clipped power spectra

The panels in the left-hand column of Fig. 1 compare the dark matter power spectrum from the Horizon Run 2 simulation with the linear theory prediction in each of five angular bins. Within each panel the uppermost set of points represent the power from the original unclipped field, while the central set is generated after applying a logarithmic transformation $\ln(1 + \delta)$. The lowest set of point corresponds to a field subject to a clipping transformations, (1), with a threshold value δ_0 chosen such that 20% of the mass is removed. Each dashed line corresponds to the model based on the linear power spectrum given by (28) where the amplitude is rescaled to fit the transformed spectra. The form of the real space linear power spectrum and the anisotropy parameter β are assigned values according to the simulation parameters. The value of $\beta \equiv f/b$ is derived using the linear growth rate $f \approx \Omega_m^{0.55}$ and the linear bias $b = 1$ since we are working directly with the dark matter. No error bars are displayed in these panels because the statistical error is considerably smaller than the marker size.

The bins in k and μ were chosen to match the power spectra derived from the GAMA survey.

For such an evolved field the linear theory prescription given by (28) typically holds only on very large scales. Even at $k \sim 0.1 \text{ hMpc}^{-1}$ the model overestimates the power in the highest μ bin by almost 10%, consistent with the findings of Jennings *et al.* [35]. However once either transformation is applied, the linear formalism of (28) provides a significantly improved description. Agreement with the clipped spectrum is better than 5% within the range ($k < 0.5 \text{ hMpc}^{-1}; \mu < 0.8$). This improvement in the modeling occurs due to the strong suppression of higher order terms in perturbation theory [9]. The leading cause of tension with the model appears to be within the highest μ bin, which is perhaps unsurprising since these modes receive contributions from very small physical scales, due to the velocity dispersion of galaxies. The central set of points illustrate the response of the power spectrum to another local transformation, the logarithm of the number density, $\ln(1 + \delta)$. [8] demonstrated that this can help linearize the power spectrum of the real space dark matter field. We find that considerable linearization also occurs when applying the log transform to the dark matter field in redshift space. The shape of the linear theory power spectrum defined by (28) is recovered to better than 10% for $k < 0.4 \text{ hMpc}^{-1}$.

The central column of Fig. 1 is in the same format as the left column, but illustrates the galaxy power spectrum from the Millennium simulation at $z = 0$ before and after applying transformations to the number density field. As before, the uppermost set of points in each panel represent the original unclipped field. The lowest set of points are generated by clipping 20% of the galaxies, which brings the data closer to the shape of the linear model (28), as given by the dashed line. To obtain the value of β for the model requires a combination of the linear bias, which is estimated from the amplitudes of the largest Fourier modes in the simulation to be $b = 1.125$, and the growth rate $f(z = 0) = 0.47$. As quantified in Sec. IV A, the true value of the linear bias is only an estimate; however, it remains a subdominant source of uncertainty. The error bars in the central and right hand columns are significantly larger than those in the left-hand column, reflecting the considerably smaller box of the Millennium-I simulation (500 Mpc/h) compared to that of the Horizon Run (7200 Mpc/h).

Deviations between the clipped spectrum and the linear model remain lower than 10% for all data at $k < 0.3 \text{ hMpc}^{-1}$. Unlike the case of dark matter, it is the lowest μ bin that causes the greatest tension with the model. This may be due to the local motions of galaxies causing them to be displaced from their high density regions, which would make the clipping process less efficient, leaving behind a considerable proportion of the nonlinear contributions to the power spectrum. The middle set of points in Fig. 1 correspond to the logarithmic transformation, but

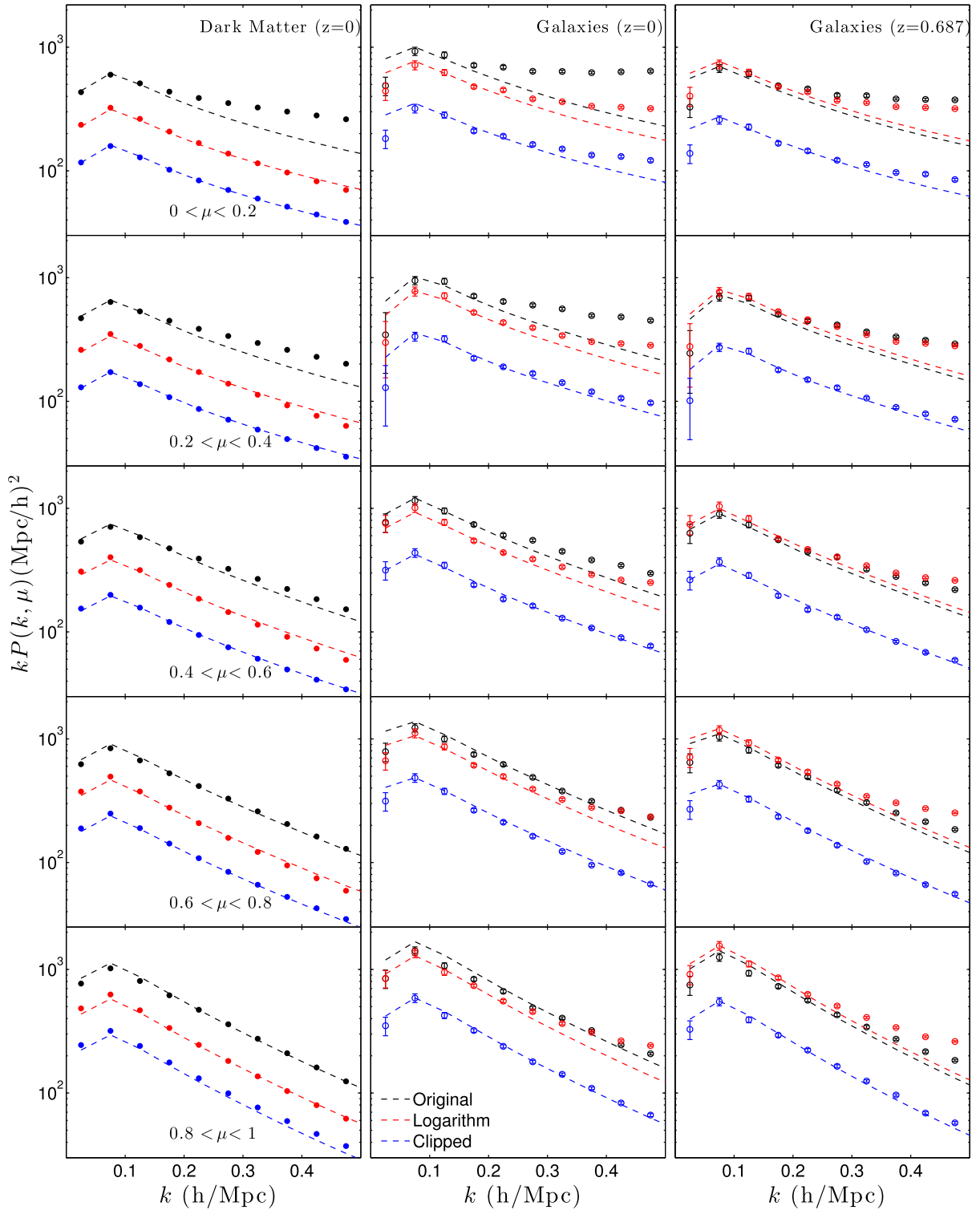


FIG. 1. The redshift space power spectra $P(k, \mu)$ from the simulations described in Sec. IV, divided into five angular bins. The three columns are dark matter from the Horizon simulation (left); galaxies at $z = 0$ from the Millennium simulation (middle); and galaxies at $z = 0.687$ from the Millennium simulation (right). Within each panel the original power spectrum is shown as black set of data while the red and blue sets correspond to the power spectra generated after the number density field has been subject to logarithmic and clipping transformations respectively. The clipping threshold is chosen such that it lies below 20% of the mass. The dashed lines represent the linear model specified by (28). For the case of the transformed spectra, the amplitude of the model is rescaled to fit the data.

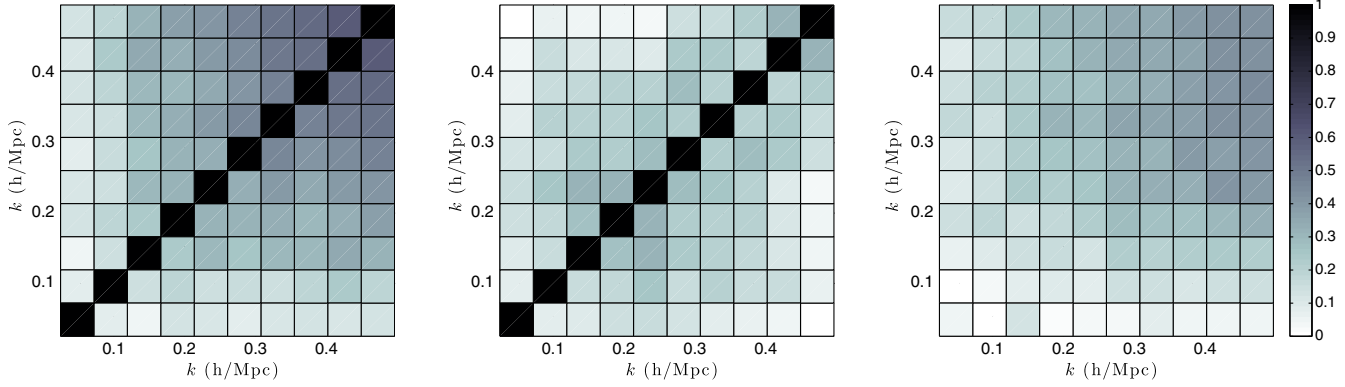


FIG. 2. Three subsections of the 70 parameter correlation matrix from the power spectrum $P(k, \mu)$ of the unclipped dark matter field in the Horizon simulation. The left-hand panel corresponds to the lowest bin in wave vector orientation, $0 < \mu < 0.2$, while the center panels relates to the highest bin, $0.8 < \mu < 1$. Each section consists of ten bins in wave number k . The right-hand panel is subsection of the correlation matrix associated with the cross-correlation between those two μ bins. The growing off diagonal correlations towards smaller scales, particularly apparent in the lowest μ bin, is an indicator of the nonlinear structure that generates mode coupling. The correlation matrix was estimated by following the prescription described in Sec. IV D.

now only a modest degree of linearization is observed. This reduced effectiveness can be attributed to the sampling noise from the galaxy fields with $\bar{n} = 0.01 \text{ Mpc}^{-3}$. Due to this sensitivity to the level of shot noise, we shall focus on the clipping transform for the remainder of this work, which by contrast is largely insensitive to shot noise.

The right-hand column of Fig. 1 explores the power spectra for a different sample of galaxies at a different redshift, $z = 0.687$. Here both the linear bias $b = 1.29$ and growth rate $f(z = 0.687) = 0.77$ have changed from those of the central column, yet the outcome is similar, in that the application of clipping significantly improves the performance of the Kaiser model defined in (28). As with the low redshift galaxy sample, the departure from linearity is less than 10% for $k < 0.3 \text{ hMpc}^{-1}$.

Figure 2 displays three subsections of the 70×70 correlation matrix (14 bins in k , 5 bins in μ) associated with the unclipped dark matter field. The coupling of modes becomes particularly apparent towards higher wave numbers, $k > 0.25 \text{ hMpc}^{-1}$. It is interesting to note that these off diagonal terms fall by approximately 20%–30% after clipping has been applied. This decorrelation of neighboring bins was previously observed in the real space power spectrum [9].

The shrinkage constant also reduces considerably after the application of clipping. From the unclipped field we find a shrinkage constant of $\delta^* = 0.25$. Thresholds selected to remove 5% and 10% of the dark matter yields shrinkage constants of 0.18 and 0.14 respectively, reflecting the increasingly Gaussian nature of these fields.

In Fig. 3 we explore the efficacy of fitting the power spectra of clipped fields with the nonlinear model, specified by (29), which is based on the model of the real space power spectrum presented in [9]. The simulations in question are the same as those used in Fig. 1. The data

used to fit the model parameters are $k < 0.4 \text{ hMpc}^{-1}$ and $\mu < 0.8$. The upper set of data corresponds to the original field, with the dashed line now making use of the one-loop power spectrum. The middle and lower sets of data relate to clipping 10% and 20% of the mass respectively. Now that an extra contribution from the one-loop power spectrum is included, the data in the lower μ bins are much better accounted for, compared with the linear model in Fig. 1. However the highest μ bin appears significantly underestimated.

We conducted further investigations by evaluating the power spectra of dark matter haloes in the Millennium simulation, before and after clipping, at the same two snapshots as the aforementioned galaxy catalogues. A very similar trend is found, whereby the clipped spectra are readily described by the model for $k < 0.4 \text{ hMpc}^{-1}$. Since these power spectra appear very similar to those displayed in Fig. 3, they are not shown here.

B. Growth of structure

First we attempt to recover the correct cosmological parameters from the clipped power spectra, using only the linear model given by (28):

$$P_c^s(k, \mu) = \mathcal{C} \left\{ \frac{b^2(1 + \beta\mu^2)^2}{1 + (k\sigma_v\mu)^2/2} P_L(k); u_0 \right\}. \quad (36)$$

Our basic set of parameters is $\{\beta, b\sigma_8, \sigma_v, a_L\}$. The normalized growth rate $f\sigma_8$ may be derived from these via the following relation:

$$f\sigma_8 = \beta b\sigma_8. \quad (37)$$

Figure 4 shows the constraints on $f\sigma_8$ when applying the linear model to the three different fields from the

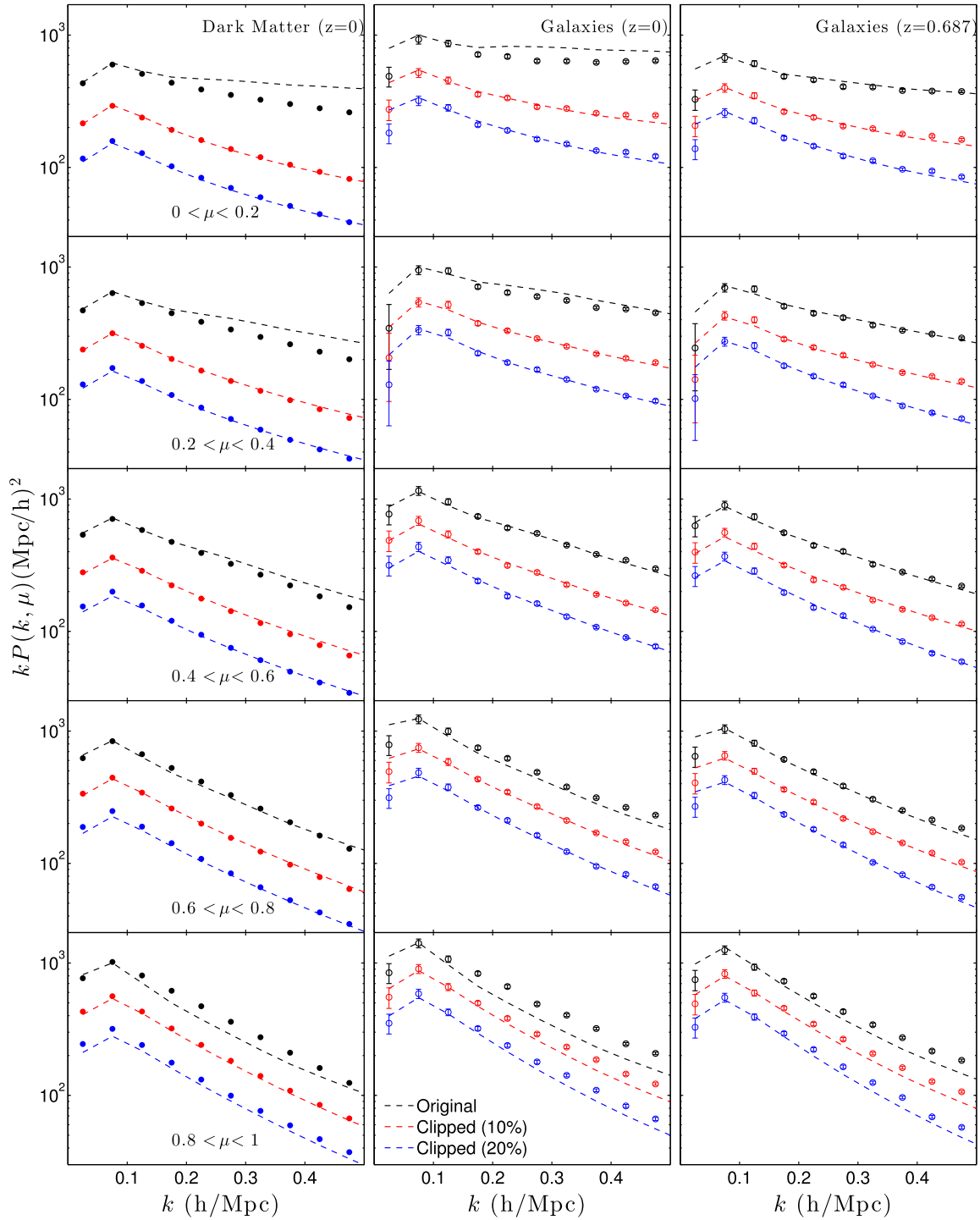


FIG. 3. A similar format to Fig. 1, except here the dashed lines represent the best-fit nonlinear model, as given by (29). The central and lower sets of data points correspond to different clipping thresholds, removing 10% and 20% of the mass respectively.

simulations, as a function of the maximum wave number k_{\max} . In each case the true value of $f\sigma_8$ is illustrated by a horizontal dashed line. The points have small horizontal offsets for clarity and appear in order of increasing clipping strength. The squares, triangles and inverted triangles correspond to clipping thresholds below 5%, 10% and

15% of the field respectively. The shaded regions represent their 68% confidence limits. For reference, the maximum likelihood points from the original unclipped field are shown as black circles.

Results from the dark matter field are shown in the left-hand panel. Removing only 5% of the mass is found to be

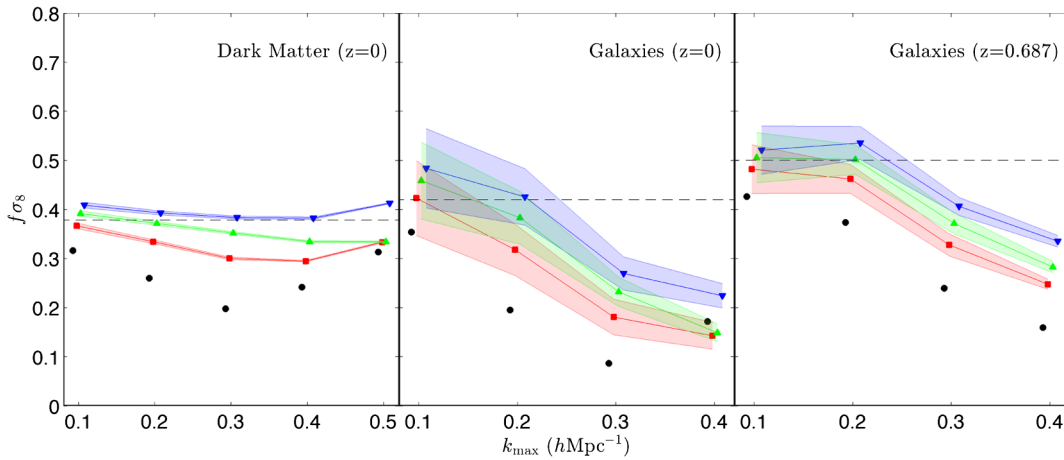


FIG. 4. Marginalized constraints on the normalized growth rate $f\sigma_8$ from the simulations, when using only the linear power spectrum (27) to model the data. The three shaded regions arise from the power spectra after clipping 5% (squares), 10% (triangles), and 15% (inverted triangles) of the mass from each field. The horizontal dashed lines indicates the true values of $f\sigma_8$ within each snapshot.

sufficient to correct for much of the nonlinear behavior on scales $k \lesssim 0.3 \text{ hMpc}^{-1}$. Similar behavior is seen in the central panel of Fig. 4, which uses the $z = 0$ galaxy sample, with a true value of $\sigma_8 = 0.9$. In the right-hand panel we find that the galaxy sample at $z = 0.687$ also significantly improves the recovery of the underlying cosmology when using the linear model on scales $k \lesssim 0.3 \text{ hMpc}^{-1}$. The shaded regions are significantly broader in the central and right-hand panels, reflecting the considerably smaller box of the Millennium-I simulation ($500 \text{ h}^{-1}\text{Mpc}$) compared to that of the Horizon Run ($7200 \text{ h}^{-1}\text{Mpc}$).

In the context of the more general model (29), which invokes additional contribution controlled by the ϵ parameter, we know that for very weak clipping ϵ will be positive and when the clipping is very strong ϵ becomes negative as the contributions from cross spectra such as $P_{12}(k)$ dominate. Therefore there is inevitably a threshold at which

ϵ vanishes and the linear model offers a good fit to the data. This is represented by the inverted triangles in Fig. 4.

C. Nonlinear model

We repeat our analysis using the extended model defined by (29), which introduces an extra parameter ϵ to control the amplitude of the clipped one-loop power spectrum, which is otherwise fully specified in terms of the linear power spectrum. Figure 5 shows the constraints derived on $f\sigma_8$ from the three simulated fields clipped with the same thresholds as Fig. 4. The extra freedom in the real space matter power spectrum leads to significantly improved measurements at weaker clipping thresholds. In the case of dark matter we find that the maximum likelihood is within 10% of the true value for each $k_{\text{max}} \leq 0.4 \text{ hMpc}^{-1}$ and for each threshold. Similarly in the galaxy field both clipping thresholds return more consistent constraints, and with only

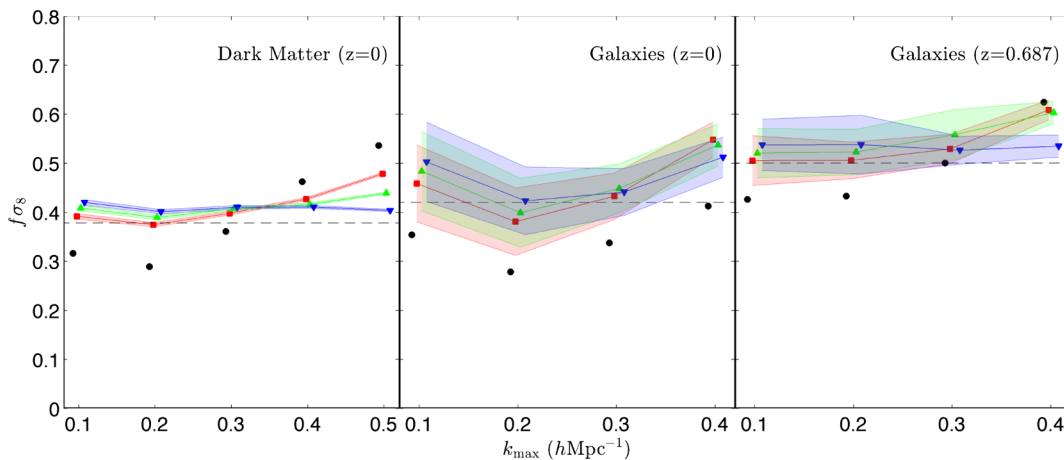


FIG. 5. The same format as Fig. 4 except we now fit the data using the nonlinear model defined by (29), with a uniform prior $-1 < \epsilon < 1$.

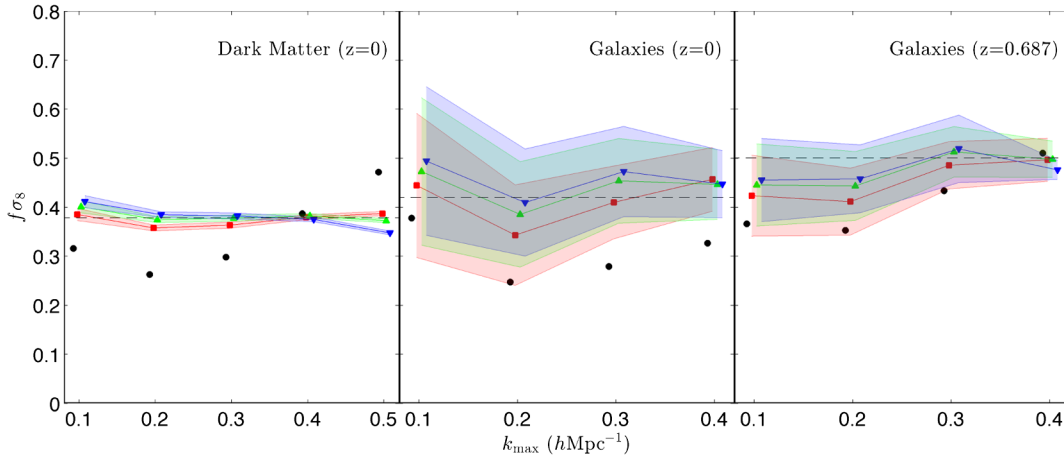


FIG. 6. The same format as Fig. 5 except here we only make use of intermediate wave vector angles ($0.2 < \mu < 0.8$).

a modest loss of precision compared to the simpler linear model. However we find that the extreme values of μ are responsible for the bulk of the tension between the model and the data. It may be the case that more complex models such as those outlined by Taruya *et al.* [36] may provide a better description of the anisotropies in the clipped power spectrum. An exploration of these models in the context of clipped fields is beyond the scope of this work. Restricting ourselves to $0.2 < \mu < 0.8$ improves agreement between the model and data to better than 2% across all wave numbers $k < 0.5 h\text{Mpc}^{-1}$. This is reflected in Fig. 6 where the tendency to overpredict $f\sigma_8$ at the highest values of k_{max} is resolved.

VI. DATA

The galaxy redshift surveys that have the greatest potential to benefit from clipping are those with a high number density of galaxies. A densely sampled field ensures that shot noise is low out to high wave numbers, even after the drop in the amplitude of the power spectrum due to clipping. The GAMA survey provides an excellent basis for the first application of clipping to a real galaxy field. In this section we present details of the survey, and how the power spectra were generated.

A. The GAMA survey

The galaxy and mass assembly (GAMA) project [37–40] is a multiwavelength photometric and spectroscopic survey. The redshift survey, which has been carried out with the Anglo-Australian telescope (AAT), has provided a dense, highly complete sampling of large-scale structure up to redshift $z \sim 0.5$. The primary target selection is $r < 19.8$ (where r is an extinction-corrected Sloan Digital Sky Survey (SDSS) Petrosian magnitude), using TILINGCATV41.

Following [41], we analyzed a highly complete subsample of the survey data set known as the GAMA II

equatorial fields. This subsample covers three 12×5 deg regions centered at 09h, 12h and 14h30m, which we refer to as G09, G12 and G15, respectively. Galaxy redshifts were obtained from the AAT spectra using a fully automatic cross-correlation code that can robustly measure absorption and emission line redshifts [42]. We restricted the redshift catalogue to galaxies with “good” redshifts ($\text{NQ} \geq 3$). In order to obtain high-resolution measurements of the density field we restricted our analysis to the redshift range $0.002 \leq z \leq 0.25$, where the galaxy number density exceeds $10^{-2} h^3\text{Mpc}^{-3}$. In the (G09, G12, G15) regions we utilized (32076, 37382, 36538) galaxies in our analysis. The comoving volume of each region is approximately $6.4 \times 10^6 (h^{-1}\text{Mpc})^3$ [41].

The survey selection function at each point, used in the calculation of the galaxy overdensity, was determined by combining the angular completeness map of the survey (which has a mean value of 97% across the three regions) with an empirical fit to the galaxy redshift distribution, performed after stacking together the data in the three regions to reduce fluctuations due to cosmic variance. Full details of the method are described in Sec. 3.2 of [41].

B. Estimating the clipped power spectrum

The clipped power spectra for each GAMA region were determined for a given overdensity threshold δ_{clip} as follows:

- (1) The galaxy distribution was binned on a common 3D grid to the selection function, with a resolution of $2 h^{-1}\text{Mpc}$. We denote the gridded distributions from the data and random samples as $D(\vec{x})$ and $R(\vec{x})$, respectively. The random catalogues are sampled from the selection function constructed for the GAMA survey data, which combines the angular completeness in each survey region with an empirical smooth redshift distribution fit to a combination of the three regions.

- (2) The distributions were smoothed using a Gaussian kernel $G(\vec{x}) = e^{-(\vec{x}\cdot\vec{x})/2\lambda^2}$. We take $\lambda = 2h^{-1}$ Mpc for our analysis. We denote the smoothed fields by $\text{sm}(D)$ and $\text{sm}(R)$.
- (3) The overdensity field for each region was estimated as $\delta(\vec{x}) = \text{sm}(D)/\text{sm}(R) - 1$, where the normalization of R was fixed such that $\langle\delta\rangle = 0$.
- (4) The mean overdensity of each region in the redshift range $z < 0.25$, relative to the average of all three regions, was estimated using the measured number of galaxies as $\delta_{\text{reg}} = (-0.0922, 0.0580, 0.0341)$ for (G09, G12, G15). The effective clipping threshold applied to the locally defined fluctuations within each region was then adjusted to $\delta_{\text{clip,eff}}$ to accommodate these mean density fluctuations, where $1 + \delta_{\text{clip,eff}} = (1 + \delta_{\text{clip}})/(1 + \delta_{\text{reg}})$.
- (5) For any grid cell with $\delta > \delta_{\text{clip,eff}}$, the unsmoothed gridded data value was lowered to $D = R(1 + \delta_{\text{clip,eff}})$.
- (6) The power spectrum of the clipped gridded data field D was measured using fast fourier transform (FFT) techniques following Sec. 3.3 of [41]. The optimal-weighting estimation scheme of Feldman, *et al.* (1994) was applied, assuming a characteristic power spectrum amplitude $P(k) = 5000 h^{-3}\text{Mpc}^3$. We binned the power spectrum by k and μ , where μ is the cosine of the angle of the wave vector \vec{k} with respect to the line-of-sight, using bin widths $\Delta k = 0.05 h\text{Mpc}^{-1}$ and $\Delta\mu = 0.2$. The integral constraint correction to the power spectrum was included in the estimation process (using the Fourier transform of the window function).
- (7) The amplitude of the measured power spectra was corrected for the misestimate of the mean density of the region, through multiplication by a factor $(1 + \delta_{\text{reg}})^2$.
- (8) The convolution matrix, which is used to project a model power spectrum $P(k, \mu)$ to form a comparison with the data given the survey window function, was determined using the method outlined in Sec. 3.3 of [41], in which the full FFT convolution is applied to a series of unit model vectors, and an equivalent matrix is constructed row-by-row.
- (9) The covariance matrix of the power spectrum measurement in (k, μ) bins was estimated by evaluating the sums described in Sec. 3.4 of [41]. Initially the measured power spectrum in each bin was used to specify the cosmic variance component. This produces an error estimate that is correlated with the data. To resolve this we modified the computation using an iterative procedure in which the best-fitting (convolved) theoretical model was determined and the covariance was reestimated using that model. Two iterations were used to ensure convergence.

We repeated the above analysis for three different clipping thresholds $\delta_0 = \{8, 5, 4\}$. These values were selected on the basis of generating a suppression of linear power between 30% and 60%. This provides an appropriate balance between the elimination of nonlinear structure and maintaining a high degree of signal to noise. The three thresholds affect approximately 0.7%, 2.1%, and 3.2% of the field in terms of volume, and approximately 6.4%, 15%, and 20% in terms of galaxies. The latter quantity is defined as the fractional reduction in the value of $\sum D(\vec{x})$ due to clipping.

Unlike nonlocal transforms such as those used for reconstructing the baryon acoustic oscillations, clipping commutes with the window function. This facilitates our interpretation of clipped power spectra, since the clipping transformation associated with the full (nonwindowed) universe can be evaluated first, before compensating for the impact associated with the window function of the survey.

VII. RESULTS FROM GAMA

In this section we perform a likelihood analysis to estimate the normalized growth rate $f\sigma_8$ at $z = 0.18$. The effective redshift of power spectrum measurements in the GAMA regions was determined by [41]. The methodology from Sec. IV is applied to each of the clipped galaxy power spectra from each of the three fields of the GAMA survey.

Figure 7 illustrates the effect of the clipping transformation. The solid bars represent the probability density function of the galaxy density field within the G09 region, defined in terms of 40 bins which are spaced equally in $\log(1 + \delta)$. The hollow bars show the resulting distribution

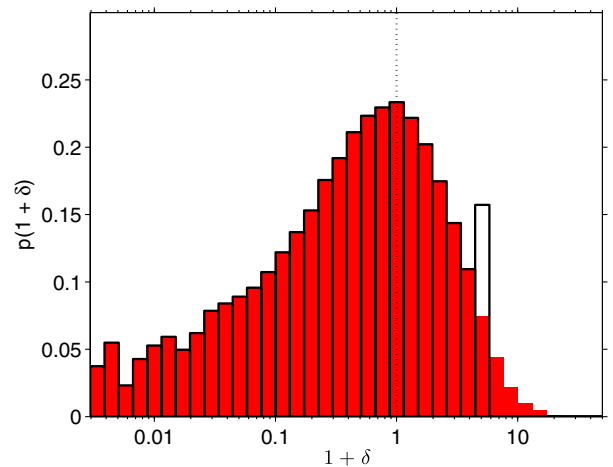


FIG. 7. The solid bars represent the probability density function for the galaxy density field within the G09 GAMA region. The hollow bars demonstrate the effect that applying a clipping threshold of $\delta_0 = 4$ would have on this field. A vertical dotted line demarcates the underdense and overdense regimes.

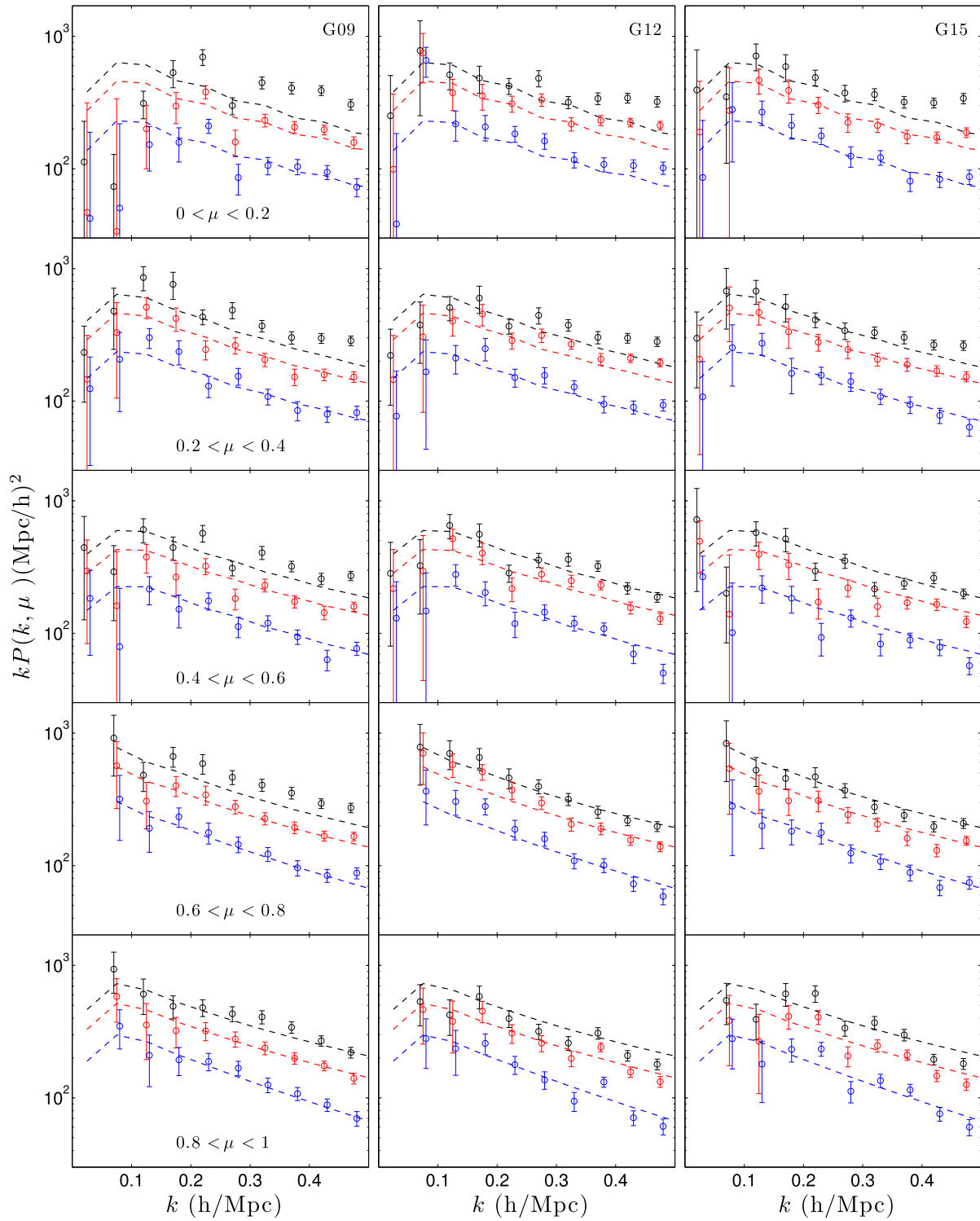


FIG. 8. The galaxy power spectra $P_g(k, \mu)$ measured within the three GAMA regions, each divided into five angular bins. Within each panel the original power spectrum is shown as the upper set of data while the middle and lower sets correspond to the power spectra generated after the number density field has been clipped with thresholds of $\delta_0 = 8$ and $\delta_0 = 4$ respectively. These thresholds generate a drop in large scale power of approximately 30% and 60%. The upper dashed line is generated from the conventional model given by (28). The lower dashed line is the fiducial model for the clipped linear power (27). At each clipping strength β and σ_v are set to their maximum likelihood values. The theoretical lines differ slightly between each region since each has been convolved with the window function of their respective region. For visual clarity a small horizontal offset is applied to the black and blue data.

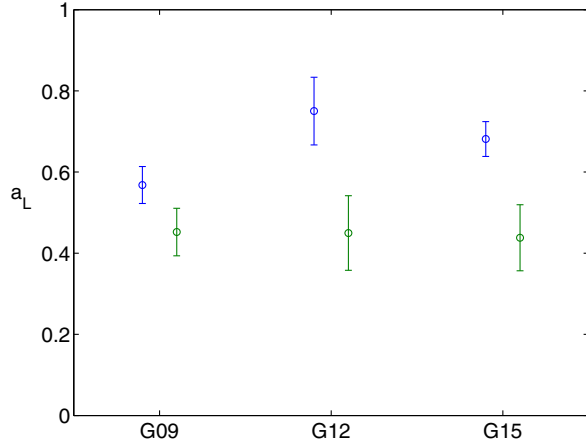


FIG. 9. Estimates of the fractional drop in linear power in each of the three GAMA regions as a result of applying clipping thresholds of $\delta_0 = 8$ (upper points) and $\delta_0 = 4$ (lower points).

function if we then apply a clipping transformation with a threshold of $\delta_0 = 4$. The distribution function remains unaltered below the threshold value δ_0 , while all contributions from greater overdensities are compressed into the bin associated with the threshold value.

A. Clipped power spectra

The panels in Fig. 8 illustrate the anisotropic power spectra derived from the three fields (G09, G12, and G15). As with the simulations, the power spectrum is divided into five equal bins in μ , spanning $0 < \mu < 1$, while the wave number bin width is taken to be $\Delta k = 0.05$. Within each field the three sets of points correspond to the power in the field before (black) and after the application of clipping thresholds $\delta_0 = 8$ (red) and $\delta_0 = 4$ (blue). At each clipping

strength, the dashed line reflects the linear model, with the maximum likelihood values of β , σ_v , and $b\sigma_8$. Estimates of the a_L parameter are shown in Fig. 9. These are determined by the fractional drop in $P(k, \mu)$ for $k < 0.1 \text{ hMpc}^{-1}$ and $\mu < 0.6$, after clipping is applied.

B. Linear model

Following the procedure outlined in Sec. VB, we use the clipped power spectra from the three GAMA regions to measure the normalized growth rate $f\sigma_8(z = 0.18)$. Here we shall present results which combine the likelihoods of the three regions. Individual results from the three separate regions can be found in Appendix C.

First we employ the linear model defined by (28) and (27), while fixing the shape of the linear power spectrum to the fiducial model. The left-hand panel of Fig. 10 show the maximum likelihood values and 68% error bars associated with $f\sigma_8$, under a range of different clipping thresholds and k_{max} values. As before, the squares, triangles and inverted triangles correspond to clipping thresholds of $\delta_0 = 8, 5$ and 4 respectively. The shaded regions represent their 68% confidence limits. For reference, the maximum likelihood points from the original unclipped field are shown as black circles. Their confidence limits are suppressed for clarity, as they do not provide an acceptable fit to the data.

With the unclipped data the constraint on $f\sigma_8$ is highly sensitive to variations in k_{max} , which is consistent with the behavior found in the simulations. Since the model inevitably underestimates the amount of real space power towards larger k , this leads to an underestimation of σ_v which in turn biases the estimate of β to be low. Applying a high clipping threshold ($\delta_0 = 8$) show a modest improvement in terms of consistency and goodness of fit. Stronger thresholds of $\delta_0 = 5$ and $\delta_0 = 4$ provide a much improved

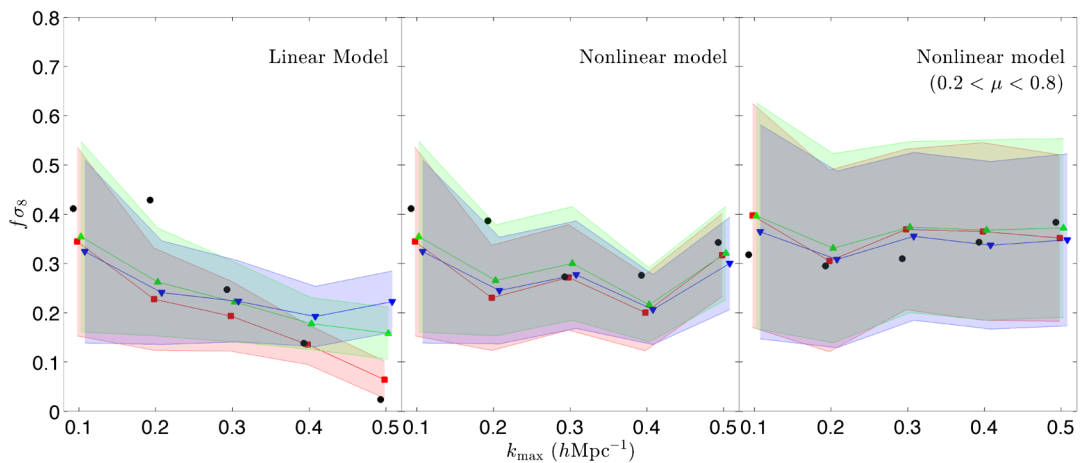


FIG. 10. Constraints on $f\sigma_8$ from the power spectra of the galaxy field from GAMA. Results from the linear model are presented in the left-hand panel, while those from the nonlinear model are illustrated in the central panel. The right-hand panel uses the same model as the central panel but only includes intermediate wave vector angles ($0.2 < \mu < 0.8$). Within each panel the squares, triangles and inverted triangles correspond to clipping thresholds δ_0 of 8, 5, and 4 respectively. The circles represent the maximum likelihood points when using the original field without clipping.

TABLE I. The maximum likelihood $f\sigma_8$ values and their associated 68% confidence intervals, when fitting the linear model to the GAMA data. Each column represents a different clipping threshold while each row represents a different k_{\max} condition. Table elements associated with a value of $\chi^2_{\text{red}} < 1.2$ are highlighted in bold. There is a strong covariance between nearby table elements, so we do not attempt to combine constraints from different clipping strengths in this work.

| k_{\max} | Unclipped | $\delta_0 = 8$ | $\delta_0 = 5$ | $\delta_0 = 4$ |
|------------|-----------------|-----------------------------------|-----------------------------------|-----------------------------------|
| 0.1 | 0.41 ± 0.07 | 0.34 ± 0.19 | 0.35 ± 0.19 | 0.32 ± 0.19 |
| 0.2 | 0.43 ± 0.08 | 0.23 ± 0.10 | 0.26 ± 0.10 | 0.24 ± 0.11 |
| 0.3 | 0.25 ± 0.05 | 0.19 ± 0.07 | 0.22 ± 0.08 | 0.22 ± 0.08 |
| 0.4 | 0.14 ± 0.02 | 0.14 ± 0.04 | 0.18 ± 0.05 | 0.19 ± 0.06 |
| 0.5 | 0.02 ± 0.01 | 0.06 ± 0.04 | 0.16 ± 0.05 | 0.22 ± 0.06 |

agreement with the model, and more consistent results towards higher wave numbers. The maximum likelihood values are displayed in Table I, and the reduced χ^2 values can be found in Table II. The power spectra associated with the clipped fields are found to adhere to the linear theory prediction more closely than the original field. From the simulations we expect a significant systematic error to arise at $k_{\max} \gtrsim 0.3 \text{ hMpc}^{-1}$, so for the linear model we use $k < 0.2 \text{ hMpc}^{-1}$ to find $f\sigma_8(z = 0.18) = 0.26 \pm 0.10$. Constraints from the clipped fields show a more consistent result across the range of wave numbers than the original field, and also have a much improved goodness of fit.

Extracting robust constraints from higher wave numbers requires a higher order model, since at these scales the amplitude of the linear power spectrum falls far below the nonlinear contributions.

C. Nonlinear model

The central panel of Fig. 10 shows the constraints on $f\sigma_8$ when using the extended model defined by (29). Again we find consistent behavior between different clipping strengths, and across a variety different maximum wave numbers. The extra degree of freedom does not appear to significantly weaken the constraints. Guided both by the performance of simulations, and the goodness of fit between the data and the model, we adopt our benchmark measurement to be $f\sigma_8(z = 0.18) = 0.29 \pm 0.10$, using

TABLE II. The reduced χ^2 values for a variety of clipping thresholds and maximum wave numbers, when fitting the linear model to the GAMA data.

| k_{\max} | Unclipped | $\delta_0 = 8$ | $\delta_0 = 5$ | $\delta_0 = 4$ |
|------------|-----------|----------------|----------------|----------------|
| 0.1 | 2.77 | 1.50 | 1.31 | 1.26 |
| 0.2 | 3.68 | 1.00 | 0.88 | 0.73 |
| 0.3 | 2.81 | 1.26 | 1.05 | 1.08 |
| 0.4 | 2.66 | 1.35 | 1.01 | 1.00 |
| 0.5 | 3.06 | 1.87 | 1.31 | 1.21 |

TABLE III. The same format as Table I but now using the nonlinear model defined by (29).

| k_{\max} | Unclipped | $\delta_0 = 8$ | $\delta_0 = 5$ | $\delta_0 = 4$ |
|------------|-----------------|-----------------------------------|-----------------------------------|-----------------------------------|
| 0.1 | 0.41 ± 0.07 | 0.34 ± 0.19 | 0.35 ± 0.19 | 0.32 ± 0.19 |
| 0.2 | 0.39 ± 0.08 | 0.23 ± 0.11 | 0.27 ± 0.11 | 0.25 ± 0.11 |
| 0.3 | 0.27 ± 0.07 | 0.27 ± 0.11 | 0.30 ± 0.11 | 0.29 ± 0.10 |
| 0.4 | 0.28 ± 0.05 | 0.20 ± 0.08 | 0.22 ± 0.08 | 0.21 ± 0.07 |
| 0.5 | 0.34 ± 0.04 | 0.32 ± 0.08 | 0.32 ± 0.10 | 0.30 ± 0.09 |

TABLE IV. The reduced χ^2 values associated with each of the maximum likelihood values presented in Table III.

| k_{\max} | Unclipped | $\delta_0 = 8$ | $\delta_0 = 5$ | $\delta_0 = 4$ |
|------------|-----------|----------------|----------------|----------------|
| 0.1 | 2.77 | 1.50 | 1.31 | 1.26 |
| 0.2 | 3.12 | 0.96 | 0.88 | 0.76 |
| 0.3 | 2.04 | 1.36 | 1.12 | 1.12 |
| 0.4 | 2.34 | 1.39 | 1.02 | 1.02 |
| 0.5 | 2.55 | 1.86 | 1.42 | 1.27 |

$k < 0.3 \text{ hMpc}^{-1}$ and $\delta_0 = 4$. This measurement is consistent with that derived from the linear model, and serves as the central result of this work. Our result is consistent with the findings of Blake *et al.* [41], who used the same (unclipped) galaxy field to determine $f\sigma_8(z = 0.18) = 0.36 \pm 0.09$. The full set of constraints on $f\sigma_8$ is presented in Table III, while the reduced χ^2 values are displayed in Table IV.

A more conservative approach is to restrict our analysis to intermediate wave vectors, $0.2 < \mu < 0.8$, and the results

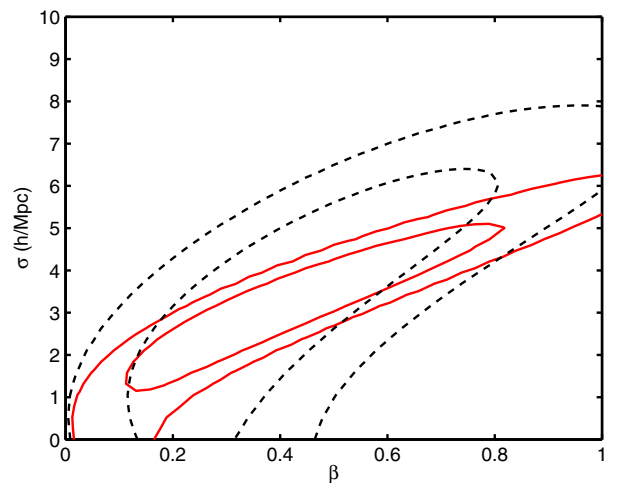


FIG. 11. The joint likelihood contours at 68% and 95% when evaluating the GAMA power spectra with $\delta_0 = 4$, using the linear model. The degeneracy with the velocity dispersion σ_v limits the gain in measuring the anisotropy parameter β when utilizing higher wave numbers. The dashed contours show the degeneracy when fitting $k_{\max} = 0.3 \text{ hMpc}^{-1}$, while the solid contours represent $k_{\max} = 0.5 \text{ hMpc}^{-1}$.

are shown in the right-hand panel of Fig. 10. While the susceptibility to systematic errors has been reduced, there is also a substantial loss of precision. Therefore even when making use of the full range of wave numbers, $k < 0.5 \text{ hMpc}^{-1}$, the resulting confidence interval is found to be $f\sigma_8 = 0.35 \pm 0.17$.

While a significant improvement in precision is achieved by increasing the k_{max} value up to 0.3 hMpc^{-1} , thereafter the gain is not as great as one might expect from the increased abundance of Fourier modes. One of the key limitations remains the degeneracy between the anisotropy parameter β and the velocity dispersion σ_v . Their joint likelihood is illustrated in Fig. 11, for two different values of k_{max} . Clearly if additional information were available to measure or predict the value of σ_v , substantial improvements in the measurement of $f\sigma_8$ could be made. Another factor which limits the gains available from smaller scales is the shot noise. Its fractional importance is amplified by the reduction in the amplitude of the power, which becomes particularly apparent at the lowest threshold.

VIII. DISCUSSION

We have developed the clipping procedure proposed in Simpson *et al.* [20] to enable its application to anisotropic fields, and applied this new analysis technique to the $z < 0.25$ sample from the GAMA survey. A simple model based on the linear power spectrum is used to measure the normalized growth rate $f\sigma_8(z = 0.18) = 0.26 \pm 0.10$ at $k_{\text{max}} = 0.2 \text{ hMpc}^{-1}$. Employing a higher order model from perturbation theory allows the use of higher wave numbers, while still not requiring numerical simulations for calibration. For this case we find $f\sigma_8(z = 0.18) = 0.29 \pm 0.10$ when using $k_{\text{max}} = 0.3 \text{ hMpc}^{-1}$ and density fluctuations $\delta_g < 4$. These results alone are not in significant tension with expectations from the Planck data [43] within the context of a standard Λ CDM model. However, they do add to a growing body of evidence that appears to prefer a lower amplitude of density perturbations at low redshifts. Such evidence includes weak gravitational lensing [44], galaxy clusters [45], and a number of other measurements of redshift space distortions [46]. This trend is also visibly apparent in Fig. 12, but there are several possible explanations for this behavior. One interpretation of this is the reduction in the quadrupole generated by nonlinear motions, relative to the Kaiser prediction, as illustrated in Fig. 2 of [35]. However our result would be largely insensitive to this effect. Another interpretation is the presence of nonlinear galaxy bias. An additional isotropic contribution to the power dilutes the strength of the anisotropic clustering signal. This effect can be seen in Fig. 5 where, before clipping is applied, the inferred value of $f\sigma_8$ is significantly lower than the correct value even when using $k \lesssim 0.1 \text{ hMpc}^{-1}$. The simplest forms of nonlinear bias are strongly suppressed by clipping, but others

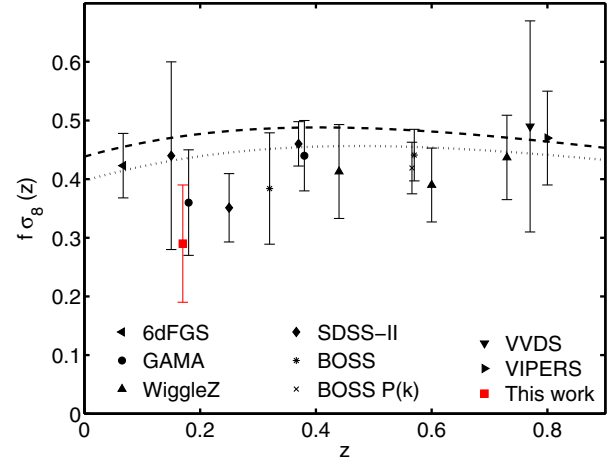


FIG. 12. Comparison of measurements of $f\sigma_8$ as measured from anisotropic clustering in galaxy redshift data. These include 6dGFS [49]; WiggleZ [50]; BOSS [51,52]; BOSS P(k) [53]; VVDS [54]; SDSS-II [55,56]; VIPERS [57]; GAMA [41]. The dotted and dashed lines represent theoretical predictions for a flat LCDM cosmology, as given by $[\Omega_m, \sigma_8]$ pairs from the maximum likelihood values from Planck [43] using only polarization [0.286; 0.796], and the full data set [0.3156; 0.831].

such as stochastic bias are likely to remain, and therefore merit further investigation. It is also important to note that most studies of redshift space distortions rely upon *a priori* on the range of possible background geometries based on results from WMAP. The portion of the error budget associated with the Alcock-Paczynski effect in each survey will therefore be highly correlated [25,47,48].

As was found to be the case in real space [9], the preferred range of clipping thresholds are typically those that reduce the linear power by around 25%–50%. For our galaxy field this corresponded to thresholds in the range $4 < \delta_0 < 8$. Higher thresholds lead to weaker clipping, which is less effective at removing the problematic contributions from nonlinear structure. Meanwhile stronger clipping from lower thresholds leads to reduced signal-to-noise and may also induce a significant cross-correlation between the Gaussian and residual fields.

At present the precision of our measurement of the growth of structure is limited by the degeneracy with the velocity dispersion σ_v . By applying a group finding algorithm to the galaxy catalogue it may be possible to reduce the influence of the fingers of God. This would also improve the efficiency with which clipping removes peaks in the density field.

Previous measurements of redshift space distortions either rely heavily on calibration from numerical simulations or on more complex approaches to perturbation theory. Each of these rely on certain model-dependent assumptions, such as a linear bias model. Clipping is a complementary approach as it can identify whether the galaxy bias is showing signs of scale-dependence. The

degree of covariance between parameter constraints obtained from the clipped analysis and a conventional nonlinear analysis has yet to be quantified; however, potentially these two approaches could be combined to yield considerable additional information within the same survey volume. It may also be beneficial to perform a combined analysis of power spectra from multiple clipping thresholds.

Clipping may also be applicable to a number of other cosmological fields, which we shall consider in turn.

1. Cosmic microwave background

In the early Universe, cosmological perturbations are known to be highly Gaussian. A recent analysis of the cosmic microwave background (CMB) from Planck severely limits the amplitude of local departures from Gaussianity $f_{\text{NL}} = 2.5 \pm 5.7$ [58]. Applying a clipping transformation to the CMB anisotropies is unlikely to be beneficial, since the Gaussian component is already highly dominant. However it may be of interest to identify whether features such as the lack of power on large angular scales, and the power asymmetry on the sky, remain intact after clipping, or are exacerbated.

2. Lyman- α forest

At later times, the cosmological perturbations are again detectable in the absorption lines of quasars. This technique has been used to detect the baryon acoustic oscillations in BOSS [59]. In this case the observed tracer already experiences a transformation similar to clipping, in that the highest density regions form damped Ly α systems. The results of Sec. II may be generalized to other local transformations. For example the transformation relating the local density to the observed flux F is often approximated as

$$g(\delta) = \exp[-A(1 + \delta)^\alpha], \quad (38)$$

which can be used in conjunction with (3) to directly compute the flux correlation function in terms of the linear power spectrum.

3. Weak gravitational lensing

The shapes of high redshift galaxies are coherently distorted by the intervening matter perturbations. Cosmic shear offers the most direct insight into the dark matter distribution at lower redshifts, yet uncertainties in the small scale power spectrum limit the amount of cosmological information that may be extracted. Gaussianization of the convergence field have been proposed by several authors, such as Seo *et al.*, Joachimi *et al.*, and Yu *et al.* [60–62]. However as highlighted in [61], in the presence of shape noise the benefits of the transformation are minimal. This is due to the substantial reduction in the amplitude

of the resulting power spectrum. Applying clipping here with a suitably high threshold may be advantageous as it can suppress the strongest sources of nonlinearity while still preserving a high level of signal to noise. But the observable field appears in a projected two-dimensional form, due to the broad lensing kernel, so the identification and suppression of peaks is a less efficient procedure compared with the full three-dimensional data that can be acquired from the distribution of galaxies.

4. Galaxy clustering

Cosmological information from the galaxy power spectrum can be split into three categories: geometric information from the baryon acoustic oscillations, primordial information from the broader shape of the power spectrum, and gravitational information from the degree of anisotropic clustering.

Local density transformations such as clipping ameliorate nonlinearities associated with high density regions. One form of nonlinearity for which this is not the case is that associated with the smoothing of the baryon acoustic peak in the galaxy correlation function. Nonlocal transformations are more appropriate for this form of peak reconstruction, as demonstrated in [63–65], where the signal can be largely restored by reversing the inferred large scale displacements.

The large-scale shape of the galaxy power spectrum is sensitive to a variety of cosmological parameters such as the matter density, the spectral index, and the neutrino mass. The precision of these parameter measurements is limited by the uncertain nature of galaxy bias, which is expected to be linear on very large scales $k \lesssim 0.1 h\text{Mpc}^{-1}$ but not on smaller scales where the vast majority of the information resides. Clipping can greatly assist in linearizing the galaxy bias, thereby ensuring the clipped galaxy power spectrum bears a close resemblance to the clipped dark matter power spectrum. Interpreting the shape is less straightforward since stronger clipping leads to a change in the shape of the power spectrum, but this is fully specified by the transformation defined in (5). Overall, then, we see considerable scope for further applications of the method presented here.

ACKNOWLEDGMENTS

F. S. acknowledges support by the European Research Council under the European Community's Seventh Framework Programme FP7-IDEAS-Phys.LSS 240117. C. B. acknowledges the support of the Australian Research Council through a Future Fellowship award. This research was supported by the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through Project No. CE110001020. P. N.

acknowledges the support of the Royal Society through the award of a University Research Fellowship and the European Research Council, through receipt of a Starting Grant (Grant No. DE GAS-259586). C. H. acknowledges support from the European Research Council under the EC FP7 Grant No. 240185. GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programs including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT, and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is <http://www.gama-survey.org>.

APPENDIX A: SPECTRAL DISTORTION

First we view how rapidly the series expansion given by (8) converges on the numerical solution. Taking the $z = 127$ snapshot as our fiducial Gaussian random field (GRF) we apply clipping and evaluate the true clipped power spectrum, $P_c(k)$. We then generate an estimate of the clipped power, $\hat{P}_c(k)$, by truncating the series of (8) at $n = 0, 2, 4$. Figure 13 illustrates the ratio of the true and estimated power spectra in each case. In order to establish subpercent precision in the estimated power spectrum, it is sufficient to stop at $n = 2$ provided the parameter $a_L > 0.5$. Throughout this work we evaluate terms at $n \leq 4$.

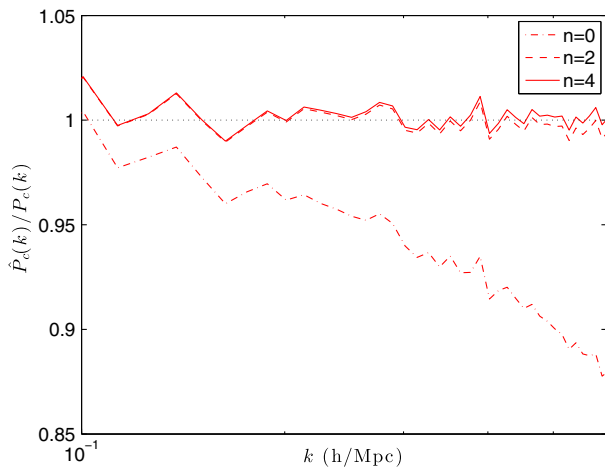


FIG. 13. The convergence of series given in (8) to the numerical solution, as the number of terms in the summation is increased from $n = 0, 2, 4$. For this case the sample GRF was clipped at the 0.7σ level, resulting in an approximately 40% loss of power on large scales.

APPENDIX B: CLIPPED ANISOTROPIC FIELDS

In this section we explore the consequences of (8) by performing numerical tests on anisotropic fields. We take the $z = 127$ Millennium-I density field in real space and impose a distortion along one axis consistent with the prescription of [26]:

$$\delta^s(k, \mu) = \delta^r(k)(1 + \beta\mu^2) \frac{1}{\sqrt{1 + k^2\mu^2\sigma_v^2/2}}, \quad (\text{B1})$$

where μ is the cosine of the angle between the wave vector and the line of sight, and the longitudinal amplification factor $\beta \equiv f/b$. This leaves us with an anisotropic GRF whose power spectrum recovers the standard form

$$P(k, \mu) = P^r(k)(1 + \beta\mu^2)^2 \frac{1}{1 + k^2\mu^2\sigma_v^2/2}. \quad (\text{B2})$$

In Fig. 14 we can see that the fractional change in the angle-averaged power spectrum induced by clipping is slightly reduced when the velocity dispersion σ_v is introduced. The larger gradient in $P(k)$ leads to a stronger contribution from higher order terms in (8).

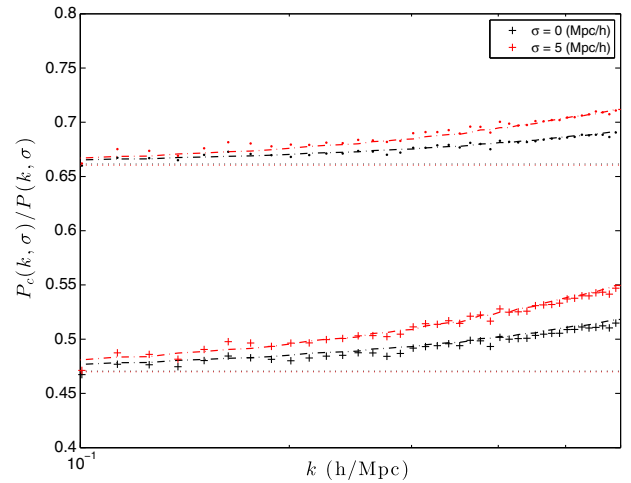


FIG. 14. The fractional change in the angle-averaged power spectrum of the field described by (B2), where we have set $\beta = 0.5$ and we explore $\sigma = 0$ and $\sigma = 5$, each evaluated at two different clipping thresholds (removing 5% and 10% of the mass). The real-space density field is taken from the $z = 127$ snapshot of the Millennium-I simulation. Larger values of σ can be seen to slightly reduce the effects of clipping on small scales. This arises from the contribution from terms in (8), which involve the self-convolution of the power spectrum. The horizontal dotted lines are the estimated clipped power when only using the first term from (8), while adding in the second leads to the dashed lines, which are in good agreement with the numerical results.

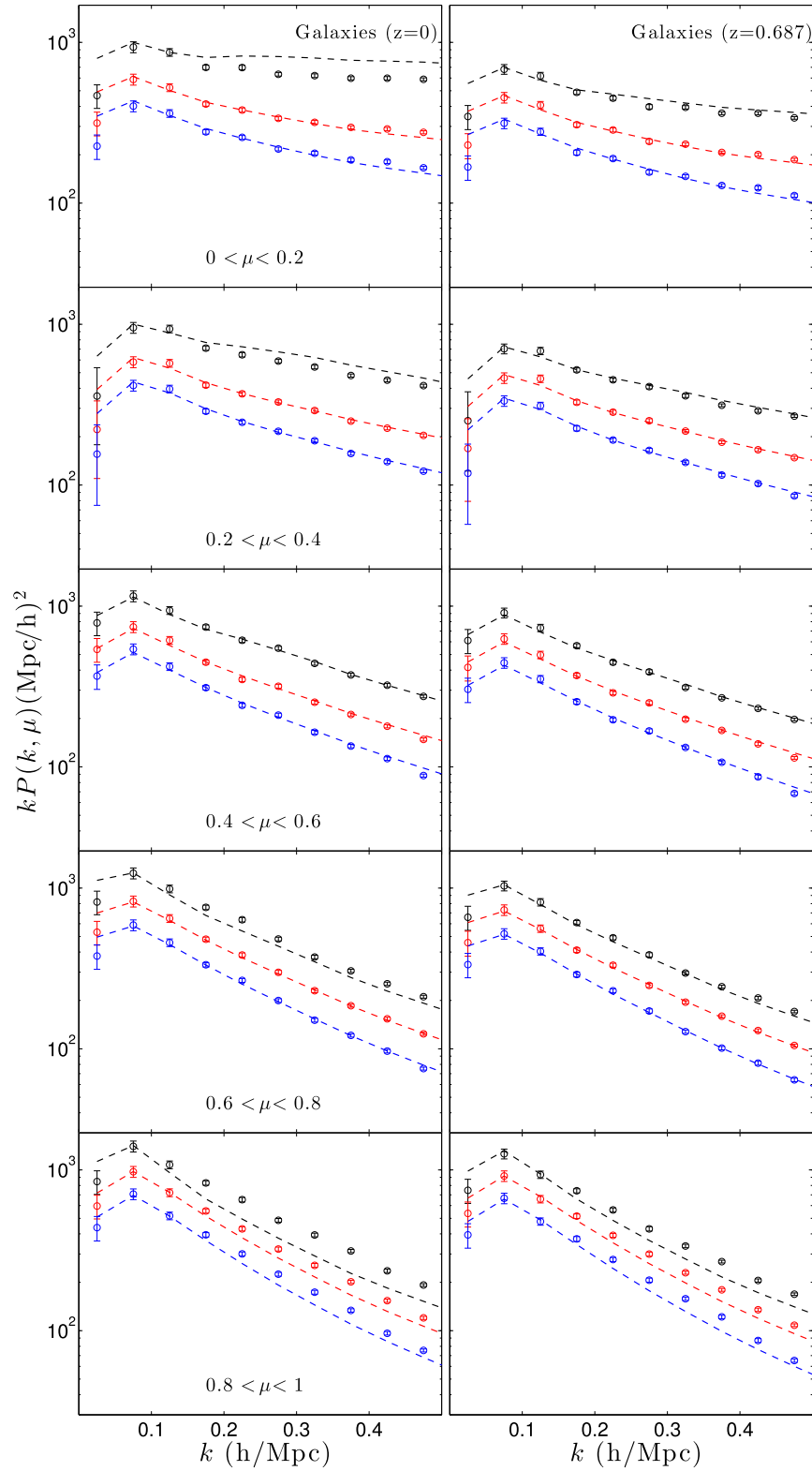


FIG. 15. The same format as Fig. 3 but we now use a finer grid to define the number density field, at 1.95 Mpc/h instead of 3.9 Mpc/h. Since the total variance of the field is greater, slightly less power is lost after removing the same fraction of galaxies as before.

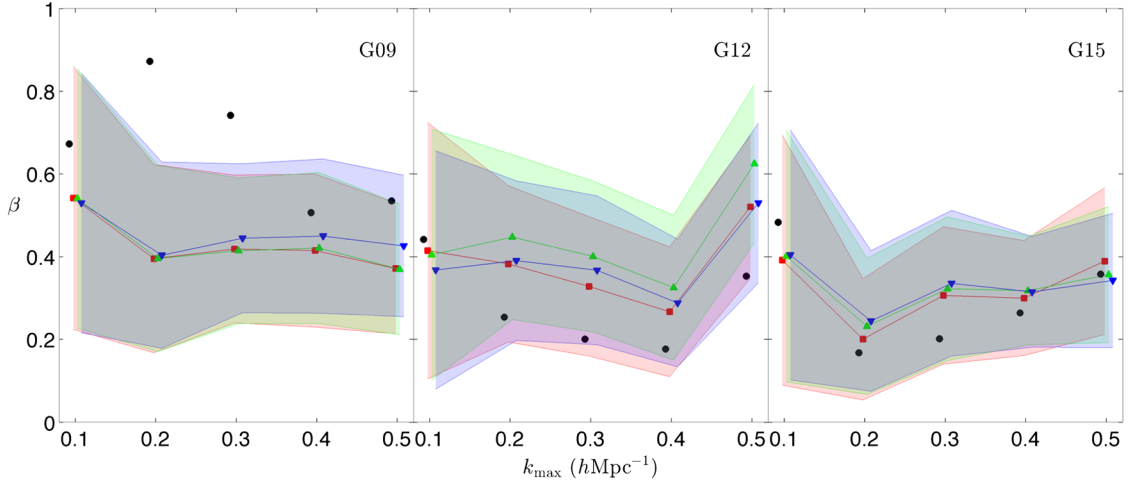


FIG. 16. Constraints on the anisotropy parameter β from the three individual GAMA regions, from left to right: G09, G12, G15, for the case of the unclipped galaxy field and three clipping thresholds δ_0 of 8, 5, and 4.

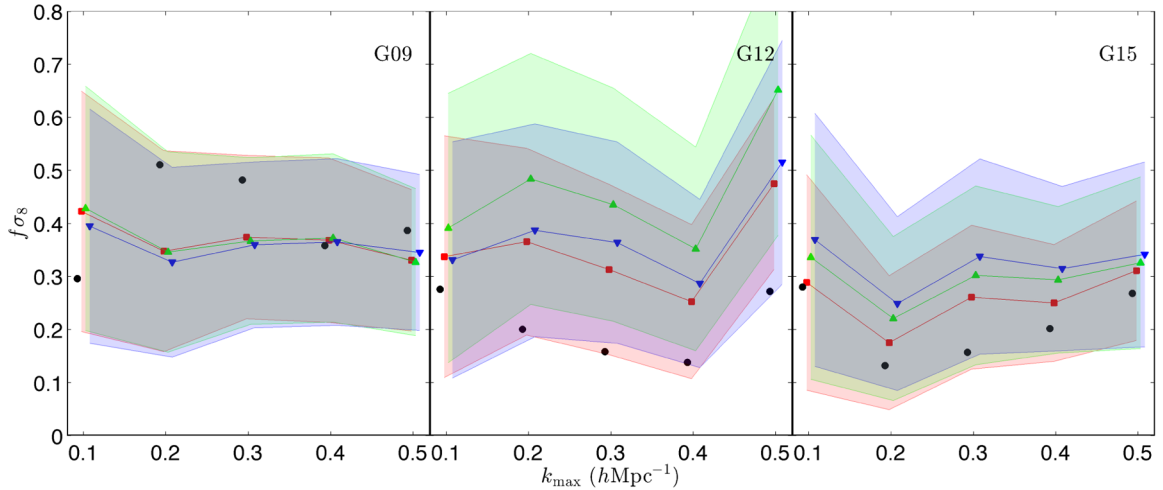


FIG. 17. Constraints on $f\sigma_8$ from the three individual GAMA regions, from left to right: G09, G12, G15, for the case of the unclipped galaxy field and three clipping thresholds δ_0 of 8, 5, and 4.

However in all cases the dashed line of the model successfully reproduces the behavior of the data.

Next we evaluate the power spectrum from the mock galaxy fields at lower redshifts. Figure 15 repeats the analysis of Fig. 3 but now with a smoothing length of 1.95 Mpc/h. The sets of parameter constraints derived from these two different smoothing lengths are found to be highly consistent with each other.

APPENDIX C: THE GAMA REGIONS

Figures 16 and 17 present constraints on β and $f\sigma_8$ from the three separate GAMA regions. In each case the results appear consistent between the three regions, and across the three thresholds within each region.

The clipping statistics for each field are presented in Table V.

TABLE V. The fraction of the field volume f_V and the effective fraction of the galaxies f_m subject to clipping from each of the GAMA regions, for each of the three threshold values.

| Region | δ_0 | f_V | f_m |
|--------|------------|-------|-------|
| 9 | 4 | 0.025 | 0.169 |
| 12 | 4 | 0.037 | 0.210 |
| 15 | 4 | 0.035 | 0.217 |
| 9 | 5 | 0.016 | 0.119 |
| 12 | 5 | 0.024 | 0.157 |
| 15 | 5 | 0.024 | 0.165 |
| 9 | 8 | 0.005 | 0.047 |
| 12 | 8 | 0.008 | 0.065 |
| 15 | 8 | 0.008 | 0.081 |

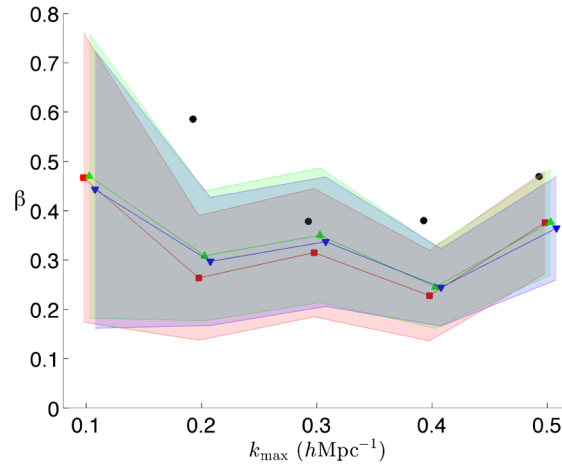


FIG. 18. Constraints on the anisotropy parameter β from the combination of the three GAMA regions for a variety of k_{\max} values, using the nonlinear model specified by (29). The squares, triangles and inverted triangles correspond to clipping thresholds δ_0 of 8, 5, and 4 respectively. The circles represent the maximum likelihood points when using the original field without clipping.

APPENDIX D: ANISOTROPIC CLUSTERING

Here we present constraints on the anisotropy parameter β from the clipped galaxy power spectrum. To do so we marginalize over the three model parameters (a_L , σ_v , $b\sigma_8$) while fixing the shape of the linear power spectrum to the fiducial model.

Figure 18 illustrates the confidence intervals derived from the combination of the three fields. The maximum likelihood values for β are plotted as a function of the maximum wave number used to compare to the model. The behavior closely reflects that of $f\sigma_8$, as seen in the central panel of Fig. 10.

-
- [1] C. Blake *et al.*, *Mon. Not. R. Astron. Soc.* **415**, 2876 (2011).
- [2] B. A. Reid *et al.*, *Mon. Not. R. Astron. Soc.* **426**, 2719 (2012).
- [3] J. Kwan, G. F. Lewis, and E. V. Linder, *Astrophys. J.* **748**, 78 (2012).
- [4] E. V. Linder and J. Samsing, *J. Cosmol. Astropart. Phys.* **2** (2013) 025.
- [5] B. A. Reid, H.-J. Seo, A. Leauthaud, J. L. Tinker, and M. White, *Mon. Not. R. Astron. Soc.* **444**, 476 (2014).
- [6] D. H. Weinberg, *Mon. Not. R. Astron. Soc.* **254**, 315 (1992).
- [7] N. McCullagh, M. Neyrinck, P. Norberg, and S. Cole, arXiv:1511.02034.
- [8] M. C. Neyrinck, I. Szapudi, and A. S. Szalay, *Astrophys. J. Lett.* **698**, L90 (2009).
- [9] F. Simpson, A. F. Heavens, and C. Heymans, *Phys. Rev. D* **88**, 083510 (2013).
- [10] L. Lombriser, F. Simpson, and A. Mead, *Phys. Rev. Lett.* **114**, 251101 (2015).
- [11] C. D. Rimes and A. J. S. Hamilton, *Mon. Not. R. Astron. Soc.* **360**, L82 (2005).
- [12] L. Verde *et al.*, *Mon. Not. R. Astron. Soc.* **335**, 432 (2002).
- [13] F. A. Marín *et al.*, *Mon. Not. R. Astron. Soc.* **432**, 2654 (2013).
- [14] H. Gil-Marín, J. Noreña, L. Verde, W. J. Percival, C. Wagner, M. Manera, and D. P. Schneider, *Mon. Not. R. Astron. Soc.* **451**, 539 (2015).
- [15] C. Blake, J. B. James, and G. B. Poole, *Mon. Not. R. Astron. Soc.* **437**, 2488 (2014).
- [16] P. M. Sutter, A. Pisani, B. D. Wandelt, and D. H. Weinberg, *Mon. Not. R. Astron. Soc.* **443**, 2983 (2014).
- [17] X. Wang, M. Neyrinck, I. Szapudi, A. Szalay, X. Chen, J. Lesgourgues, A. Riotto, and M. Sloth, *Astrophys. J.* **735**, 32 (2011).
- [18] J. Carron and I. Szapudi, *Mon. Not. R. Astron. Soc.* **434**, 2961 (2013).
- [19] J. Carron and I. Szapudi, *Mon. Not. R. Astron. Soc.* **439**, L11 (2014).
- [20] F. Simpson, J. B. James, A. F. Heavens, and C. Heymans, *Phys. Rev. Lett.* **107**, 271301 (2011).
- [21] R. Price, *IRE Transactions on Information Theory* **4**, 69 (1958).
- [22] R. Gross and D. Veeneman, in *Conference Record, "Serving Humanity Through Communications" ICC'94, SUPER-COMM/ICC'94, IEEE International Conference on Communications, New Orleans, LA, 1994* (IEEE, New York, 1994), p. 843.
- [23] J. N. Fry and E. Gaztanaga, *Astrophys. J.* **413**, 447 (1993).

- [24] T. Baldauf, U. Seljak, V. Desjacques, and P. McDonald, *Phys. Rev. D* **86**, 083540 (2012).
- [25] F. Simpson and J. A. Peacock, *Phys. Rev. D* **81**, 043512 (2010).
- [26] N. Kaiser, *Mon. Not. R. Astron. Soc.* **227**, 1 (1987).
- [27] K. B. Fisher, *Astrophys. J.* **448**, 494 (1995).
- [28] J. Kim, C. Park, G. Rossi, S. M. Lee, and J. R. Gott, III, *J. Korean Astron. Soc.* **44**, 217 (2011).
- [29] Q. Guo, S. White, M. Boylan-Kolchin, G. De Lucia, G. Kauffmann, G. Lemson, C. Li, V. Springel, and S. Weinmann, *Mon. Not. R. Astron. Soc.* **413**, 101 (2011).
- [30] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce, *Nature (London)* **435**, 629 (2005).
- [31] E. N. Taylor *et al.*, *Mon. Not. R. Astron. Soc.* **418**, 1587 (2011).
- [32] A. Lewis, A. Challinor, and A. Lasenby, *Astrophys. J.* **538**, 473 (2000).
- [33] O. Ledoit and M. Wolf, *J. Multivariate Anal.* **88**, 365 (2004).
- [34] <http://www.econ.uzh.ch/faculty/wolf/publications/covCor.m.zip>
- [35] E. Jennings, C. M. Baugh, and S. Pascoli, *Mon. Not. R. Astron. Soc.* **410**, 2081 (2011).
- [36] A. Taruya, T. Nishimichi, and F. Bernardeau, *Phys. Rev. D* **87**, 083509 (2013).
- [37] I. K. Baldry *et al.*, *Mon. Not. R. Astron. Soc.* **404**, 86 (2010).
- [38] A. Robotham *et al.*, *Pub. Astron. Soc. Aust.* **27**, 76 (2010).
- [39] S. P. Driver *et al.*, *Mon. Not. R. Astron. Soc.* **413**, 971 (2011).
- [40] J. Liske *et al.*, *Mon. Not. R. Astron. Soc.* **452**, 2087 (2015).
- [41] C. Blake, I. K. Baldry, J. Bland-Hawthorn, L. Christodoulou, M. Colless, C. Conselice, S. P. Driver, A. M. Hopkins, J. Liske, J. Loveday, P. Norberg, J. A. Peacock, G. B. Poole, and A. S. G. Robotham, *Mon. Not. R. Astron. Soc.* **436**, 3089 (2013).
- [42] I. K. Baldry, M. Alpaslan, A. E. Bauer, J. Bland-Hawthorn, S. Brough, M. E. Cluver, S. M. Croom, L. J. M. Davies, S. P. Driver, M. L. P. Gunawardhana, B. W. Holwerda, A. M. Hopkins, L. S. Kelvin, J. Liske, Á. R. López-Sánchez, J. Loveday, P. Norberg, J. Peacock, A. S. G. Robotham, and E. N. Taylor, *Mon. Not. R. Astron. Soc.* **441**, 2440 (2014).
- [43] Planck Collaboration, [arXiv:1502.01589](https://arxiv.org/abs/1502.01589).
- [44] C. Heymans *et al.*, *Mon. Not. R. Astron. Soc.* **432**, 2433 (2013).
- [45] A. Vikhlinin, A. V. Kravtsov, R. A. Burenin, H. Ebeling, W. R. Forman, A. Hornstrup, C. Jones, S. S. Murray, D. Nagai, H. Quintana, and A. Voevodkin, *Astrophys. J.* **692**, 1060 (2009).
- [46] E. Macaulay, I. K. Wehus, and H. K. Eriksen, *Phys. Rev. Lett.* **111**, 161301 (2013).
- [47] C. Alcock and B. Paczynski, *Nature (London)* **281**, 358 (1979).
- [48] W. E. Ballinger, J. A. Peacock, and A. F. Heavens, *Mon. Not. R. Astron. Soc.* **282**, 877 (1996).
- [49] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, G. B. Poole, L. Campbell, Q. Parker, W. Saunders, and F. Watson, *Mon. Not. R. Astron. Soc.* **423**, 3430 (2012).
- [50] C. Blake *et al.*, *Mon. Not. R. Astron. Soc.* **425**, 405 (2012).
- [51] C.-H. Chuang, F. Prada, F. Beutler, D. J. Eisenstein, S. Escoffier, S. Ho, J.-P. Kneib, M. Manera, S. E. Nuza, D. J. Schlegel, D. P. Schneider, B. A. Weaver, J. R. Brownstein, K. S. Dawson, C. Maraston, and D. Thomas, [arXiv:1312.4889](https://arxiv.org/abs/1312.4889).
- [52] L. Samushia *et al.*, *Mon. Not. R. Astron. Soc.* **439**, 3504 (2014).
- [53] F. Beutler, S. Saito, H.-J. Seo, J. Brinkmann, K. S. Dawson, D. J. Eisenstein, A. Font-Ribera, S. Ho, C. K. McBride, F. Montesano, W. J. Percival, A. J. Ross, N. P. Ross, L. Samushia, D. J. Schlegel, A. G. Sánchez, J. L. Tinker, and B. A. Weaver, *Mon. Not. R. Astron. Soc.* **443**, 1065 (2014).
- [54] L. Guzzo *et al.*, *Nature (London)* **451**, 541 (2008).
- [55] C. Howlett, A. J. Ross, L. Samushia, W. J. Percival, and M. Manera, *Mon. Not. R. Astron. Soc.* **449**, 848 (2015).
- [56] L. Samushia, W. J. Percival, and A. Raccanelli, *Mon. Not. R. Astron. Soc.* **420**, 2102 (2012).
- [57] S. de la Torre *et al.*, *Astron. Astrophys.* **557**, A54 (2013).
- [58] P. A. R. Ade, N. Aghanim, M. Arnaud, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday *et al.* and Planck Collaboration, [arXiv:1502.01592](https://arxiv.org/abs/1502.01592).
- [59] T. Delubac *et al.*, *Astron. Astrophys.* **574**, A59 (2015).
- [60] H.-J. Seo, M. Sato, S. Dodelson, B. Jain, and M. Takada, *Astrophys. J. Lett.* **729**, L11 (2011).
- [61] B. Joachimi, A. N. Taylor, and A. Kiessling, *Mon. Not. R. Astron. Soc.* **418**, 145 (2011).
- [62] Y. Yu, P. Zhang, W. Lin, W. Cui, and J. N. Fry, *Phys. Rev. D* **84**, 023523 (2011).
- [63] D. J. Eisenstein, H.-J. Seo, E. Sirko, and D. N. Spergel, *Astrophys. J.* **664**, 675 (2007).
- [64] N. Padmanabhan, X. Xu, D. J. Eisenstein, R. Scalzo, A. J. Cuesta, K. T. Mehta, and E. Kazin, *Mon. Not. R. Astron. Soc.* **427**, 2132 (2012).
- [65] L. Anderson *et al.*, *Mon. Not. R. Astron. Soc.* **441**, 24 (2014).