

Toward optimal cluster power spectrum analysis

Robert E. Smith^{1,2*} and Laura Marian^{1,3}

(1) *Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str.1, Postfach 1523, 85740 Garching, Germany*

(2) *Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK and*

(3) *Argelander-Institute for Astronomy, Auf dem Hügel 71, D-53121 Bonn, Germany*

The power spectrum of galaxy clusters is an important probe of the cosmological model. In this paper we determine the optimal weighting scheme for maximizing the signal-to-noise ratio for such measurements. We find a closed form analytic expression for the optimal weights. Our expression takes into account: cluster mass, finite survey volume effects, survey masking, and a flux limit. The optimal weights are: $w(M, \chi) \propto b(M, \chi) / [1 + \bar{n}_h(\chi) \overline{b_N^2}(\chi) \overline{P}(k)]$, where $b(M, \chi)$ is the bias of clusters of mass M at radial position $\chi(z)$, $\bar{n}_h(\chi)$ and $\overline{b_N^2}(\chi)$ are the expected space density and bias squared of all clusters, and $\overline{P}(k)$ is the matter power spectrum at wavenumber k . The implementation of this weighting scheme requires knowledge of the measured cluster masses, and analytic models for the bias and space-density of clusters as a function of mass and redshift. Recent studies have suggested that the optimal method for reconstruction of the matter density field from a set of clusters is mass-weighting [1, 31, 32]. We compare our optimal weighting scheme with this approach and also with the original power spectrum scheme of Feldman et al. [2]. We show that our optimal weighting scheme outperforms these approaches for both volume- and flux-limited cluster surveys. Finally, we present a new expression for the Fisher information matrix for cluster power spectrum analysis. Our expression shows that for an optimally weighted cluster survey the cosmological information content is boosted, relative to the standard approach of Tegmark [3].

PACS numbers:

I. INTRODUCTION

The number counts of massive galaxy clusters has long been known to provide strong constraints on the cosmological model, provided one understands how to map from observed mass proxies to a theoretical halo mass [for a recent review see 4]. Since the first measurements of the clustering of Abell clusters was performed in the 70's and early 80's, it has been understood that the clustering of clusters contains additional vital information about the cosmological model [5, 6]. In particular, these studies were able to show that the clustering of clusters was stronger than that of the galaxies. This quickly led to the realization that galaxies and clusters could not both be unbiased tracers of the mass distribution [7]. One of the major attractions of the cold dark matter (hereafter, CDM) framework, is that 'biased' clustering naturally emerges within it. In Kaiser's seminal work, he showed that the peaks and troughs of a Gaussian random field were correlated more strongly than the correlation function of the unconstrained field. Under the assumption that the Abell clusters formed out of the high peaks of a Gaussian Random Field one would then expect the Abell clusters to be more strongly correlated than galaxies. Further theoretical support comes from the excursion set formalism, which showed that initially overdense patches of a CDM universe would collapse to form dark matter haloes, and that these would, in general, be positively biased with respect to the underlying matter [8–10]. One important consequence of these developments was that around the mid 90s, it was also realized that, if one combined measurements of the clustering of clusters with measurements of their abundances, one could break the degeneracies in cosmological parameters that were inherent in one single method [11–18].

Some notable measurements of the clustering of clusters are: in the X-rays, initial measurements of the cluster correlation function for ROSAT data were performed by Romer et al. [19], these were later improved upon by Collins et al. [20] using the 344 clusters in the REFLEX survey (see also [21] for results from the XBACS survey). In the optical, cluster samples tend to be orders of magnitude larger [for a review of early results see 22]. The APM galaxy survey was able to identify several hundred clusters for which clustering was computed [23–25]. In the past decade, the Sloan Digital Sky Survey (SDSS) has produced, by far, the largest homogeneous sample of optical clusters: the MaxBCG sample whose clusters are detected via the 'red sequence' cluster detection method in multi-band imaging data [26]. This sample contains 13,823 clusters with velocity dispersions $\gtrsim 400 \text{ km s}^{-1}$ and covers an area of $\gtrsim 7000$ square degrees. The cluster correlation functions were explored by Bahcall et al. [27] and Estrada et al. [28], and the

*E-mail me at: res@mpa-garching.mpg.de, r.e.smith@sussex.ac.uk

power spectrum analysis was performed by Hütsi [29].

In the future, X-ray cluster surveys, such as eROSITA, should produce homogeneous cluster samples with numbers of clusters on the order of $\sim 100,000$ [30]. Deep multi-band optical surveys such as the Dark Energy Survey [43] should also produce tens of thousands of high signal-to-noise ratio (hereafter, \mathcal{S}/\mathcal{N}) clusters. The question then arises how one should perform an optimal measurement of the clustering of galaxy clusters? In a series of recent theoretical studies [1, 31], it was claimed that if galaxy clusters were weighted by their mass, then the shot-noise on cluster power spectra measurements would be significantly reduced, hence yielding improved cosmological information. A subsequent study by [32], proposed that the optimal way to reconstruct the mass distribution from a set of clusters was to weight the galaxy clusters by some linear combination of their mass and bias. Taken at face value, these works would lead us to the conclusion that a maximal \mathcal{S}/\mathcal{N} measurement of the clustering of clusters would result from weighting by mass. However, the caveat to the above analysis was that neither work directly demonstrated that the \mathcal{S}/\mathcal{N} would be maximized. In this paper we shall directly perform this task. As we shall show, our analysis generalizes the galaxy power spectrum methods developed by Feldman et al. [2, hereafter FKP] and Percival et al. [33, hereafter PVP].

This paper is broken down as follows: In §II we overview the specifications of a cluster survey, the construction of the density field of clusters, and its basic statistical properties. In §III we detail the estimators of the two-point correlation function and the power spectrum. In §IV we write down the covariance matrix of the power spectrum estimator in the most general form and the Gaussian limit. In §V we provide details of the derivation of the optimal weighting scheme. In §VI we compare various weighting schemes with our optimal weights for the cases of volume limited and flux-limited cluster surveys. We also present a new expression for the Fisher information matrix, which may be used for predicting the cosmological information content of optimally weighted measurements of the cluster power spectrum. Finally, in §VII we summarize our findings and conclude.

II. SURVEY SPECIFICATIONS AND THE \mathcal{F}_c -FIELD

A. A generic cluster survey

Let us begin by defining our fiducial cluster survey: suppose that we have observed N clusters and to the i th cluster we assign a mass M_i , redshift z_i and angular position on the sky $\mathbf{\Omega}_i = \mathbf{\Omega}(\theta_i, \phi_i)$. The cluster selection function depends on both position and cluster mass and in general, is a complex function of the survey flux limit, and the cluster detection procedure [for an example of the complexities involved in computing this for the eROSITA mission see 30]. However, it may be simplified in the following ways. Firstly, provided the flux-limit is homogeneous across the survey area, the angular and radial parts of the selection function are separable:

$$\Theta(\mathbf{x}|M) = \Theta_{\Omega}(\mathbf{\Omega})\Theta_{\chi}(\chi|M), \quad (1)$$

where $\chi = \chi(z)$ is the radial comoving geodesic distance to redshift z . The angular selection function may be written:

$$\Theta_{\Omega}(\mathbf{\Omega}) = \begin{cases} 1; & [\mathbf{\Omega} \in \mathbf{\Omega}_{\mu}] \\ 0; & [\text{otherwise}] \end{cases}, \quad (2)$$

where $\mathbf{\Omega}_{\mu}$ defines the survey mask. The radial selection function may be written:

$$\Theta_{\chi}(\chi|M) = \begin{cases} 1; & [\chi \in \chi_{\max}(M)] \\ 0; & [\text{otherwise}] \end{cases}, \quad (3)$$

where $\chi_{\max}(M)$ is the maximum comoving geodesic distance out to which a cluster of mass M could have been detected. This last relation may be inverted to obtain a very useful relation, which is the minimum detectable cluster mass at radial position $\chi(z)$ in the survey. We shall denote this quantity as $M_{\text{Lim}}(\chi)$.

Secondly, if the survey is volume limited, the minimum detectable mass is independent of position and we may write:

$$\Theta_{\chi}^{\text{VL}}(\chi|M) = \Theta_{\chi}^{\text{VL}}(\chi)\Theta_M^{\text{VL}}(M), \quad (4)$$

where

$$\Theta_{\chi}^{\text{VL}}(\chi|M) = \begin{cases} 1; & [\chi \leq \chi_{\max}] \\ 0; & [\text{otherwise}] \end{cases}; \quad \Theta_M^{\text{VL}}(M) = \begin{cases} 1; & [M \geq M_{\text{Lim}}] \\ 0; & [\text{otherwise}] \end{cases}. \quad (5)$$

The survey volume may now be defined as the integral of the selection function over all space:

$$V_\mu(M) = \int_0^{\chi_{\max}(M)} \Theta(\mathbf{x}|M) dV(\chi, \boldsymbol{\theta}) = \Omega_\mu \int_0^{\chi_{\max}(M)} d\chi D_A^2(\chi), \quad (6)$$

where $dV(\chi, \boldsymbol{\theta})$ is the comoving volume element at position vector $\mathbf{x}(\chi(z), \boldsymbol{\theta})$, $D_A(\chi)$ is the comoving angular diameter distance. For a flat space-time geometry the survey volume simplifies to,

$$V_\mu(M) = \frac{\Omega_\mu}{3} \chi_{\max}^3(M). \quad (7)$$

B. The cluster density and over density field expansion

In general the spatial density distribution of clusters, per unit mass, at position $\mathbf{x}(\chi, \boldsymbol{\Omega})$ may be written as a sum over Dirac delta functions:

$$n_c(M, \mathbf{x}) = \sum_{i=1}^N \delta^D(M - M_i) \delta^D(\mathbf{x} - \mathbf{x}_i). \quad (8)$$

If the selection function is inhomogeneous, then the mean density of clusters varies spatially over the survey. Next, in analogy with PVP, we define a field \mathcal{F}_c , which is related to the over-density of clusters. This can be written:

$$\mathcal{F}_c(\mathbf{x}) = \int dM \frac{w(\mathbf{x}, M)}{\sqrt{A}} [n_c(\mathbf{x}, M) - \alpha n_s(\mathbf{x}, M)], \quad (9)$$

where $n_s(\mathbf{x}, M)$ represents the number density of clusters in a mock sample that has no intrinsic spatial correlations, and whose density is $1/\alpha$ times that of the true cluster field at that mass. Note that whilst the field $n_s(\mathbf{x}, M)$ has no intrinsic spatial correlations it does possess a spectrum of masses, which is closely related to the mass spectrum of the field $\langle n_c(\mathbf{x}, M) \rangle$. The choice for the normalization parameter A will be given later. The quantity $w(\mathbf{x}, M)$ denotes a weight function that, in general, may depend on both the spatial position and mass of the cluster. It is this quantity that we shall aim to determine in an optimal way.

C. Statistical properties of the cluster density field

Determination of the optimal weight function will require statistical analysis on the field \mathcal{F}_c , let us therefore introduce the necessary statistical tools. As a simple example let us compute the ensemble average value of the field \mathcal{F}_c , which can be written,

$$\langle \mathcal{F}_c(\mathbf{x}) \rangle = \int dM \frac{w(\mathbf{x}, M)}{\sqrt{A}} [\langle n_c(\mathbf{x}, M) \rangle - \alpha \langle n_s(\mathbf{x}, M) \rangle], \quad (10)$$

where the angled brackets denote an ensemble average in the following sense:

$$\langle B(\{\mathbf{x}_i, M_i\}) \rangle \equiv \int d^3\mathbf{x}_1 \dots d^3\mathbf{x}_N dM_1 \dots dM_N p(\mathbf{x}_1, \dots, \mathbf{x}_N, M_1, \dots, M_N) B(\{\mathbf{x}_i, M_i\}). \quad (11)$$

in the above $p(\mathbf{x}_1, \dots, \mathbf{x}_N, M_1, \dots, M_N)$ is the N -point joint probability distribution for the N clusters being located at the set of spatial positions $\{\mathbf{x}_i\}$ and having the set of masses $\{M_i\}$. Thus, the first expectation on the right-hand-side of Eq. (10) can be written as:

$$\begin{aligned} \langle n_c(\mathbf{x}, M) \rangle &= \int \prod_{i=1}^N \{d^3\mathbf{x}_i dM_i\} p(\mathbf{x}_1, \dots, \mathbf{x}_N, M_1, \dots, M_N) \sum_{i=1}^N \delta^D(\mathbf{x} - \mathbf{x}_i) \delta^D(M - M_i) \\ &= \sum_{i=1}^N p(\mathbf{x}, M) = N p(\mathbf{x}, M). \end{aligned} \quad (12)$$

On the first line we inserted the expansion of the cluster density field from Eq.(8) and to obtain the second we integrated over the sum of Dirac delta functions. The quantity $Np(\mathbf{x}, M)$ can be written in a more transparent way if we make use of the survey selection function:

$$Np(\mathbf{x}, M) \equiv \bar{n}_c(\mathbf{x}, M) \equiv \bar{n}(M)\Theta(\mathbf{x}|M) , \quad (13)$$

where $\bar{n}_c(\mathbf{x}, M)$ is the mean spatial density of clusters in the survey, which varies with position, and where $\bar{n}(M)$ is the intrinsic number density of clusters per unit mass – i.e. the dark matter halo mass function.

Turning to the second expectation value, we note that the only difference between $\langle n_c(\mathbf{x}, M) \rangle$ and $\langle n_s(\mathbf{x}, M) \rangle$ is the artificially increased space-density of clusters and the absence of any intrinsic clustering. Hence, we also have,

$$\alpha \langle n_c(\mathbf{x}, M) \rangle = \bar{n}(M)\Theta(\mathbf{x}|M) . \quad (14)$$

Putting this all together, we arrive at the result:

$$\langle \mathcal{F}_c(\mathbf{x}) \rangle = 0 . \quad (15)$$

Hence the \mathcal{F}_c -field, like the over-density field of matter, truly is a mean zero field.

Note that we have neglected to take into account the statistical properties of obtaining the N clusters in the survey volume. In what follows we shall assume that the survey volumes are sufficiently large that this may be essentially treated as a deterministic quantity. However, it can be taken into account [e.g. see 34–36].

III. CLUSTERING ESTIMATORS

A. An estimator for the two-point correlation function

The two-point correlation function of the field \mathcal{F}_c can be computed directly as:

$$\begin{aligned} \langle \mathcal{F}_c(\mathbf{r}_1)\mathcal{F}_c(\mathbf{r}_2) \rangle &= \frac{1}{A} \int dM dM' w(\mathbf{r}_1, M) w(\mathbf{r}_2, M') \left\langle [n_c(\mathbf{r}_1, M) - \alpha n_s(\mathbf{r}_1, M)] [n_c(\mathbf{r}_2, M') - \alpha n_s(\mathbf{r}_2, M')] \right\rangle \\ &= \frac{1}{A} \int dM dM' w(\mathbf{r}_1, M) w(\mathbf{r}_2, M') \left[\langle n_c(\mathbf{r}_1, M) n_c(\mathbf{r}_2, M') \rangle - \alpha \langle n_c(\mathbf{r}_1, M) n_s(\mathbf{r}_2, M') \rangle \right. \\ &\quad \left. - \alpha \langle n_c(\mathbf{r}_2, M') n_s(\mathbf{r}_1, M) \rangle + \alpha^2 \langle n_s(\mathbf{r}_1, M) n_s(\mathbf{r}_2, M') \rangle \right] . \end{aligned} \quad (16)$$

The terms in square bracket on the right-hand side of the above equation can be evaluated as was done when computing $\langle \mathcal{F}_c \rangle$. The first term is:

$$\begin{aligned} \langle n_c(\mathbf{r}_1, M) n_c(\mathbf{r}_2, M') \rangle &= \int \prod_{l=1}^N \{d^3 \mathbf{x}_l dM_l\} p(\mathbf{x}_1, \dots, \mathbf{x}_N, M_1, \dots, M_N) \\ &\quad \times \sum_{i,j} \delta^D(\mathbf{r}_1 - \mathbf{x}_i) \delta^D(M - M_i) \delta^D(\mathbf{r}_2 - \mathbf{x}_j) \delta^D(M' - M_j) . \end{aligned} \quad (17)$$

The double sum can be broken up into the terms where $(i \neq j)$ and the terms where $(i = j)$, whereupon

$$\begin{aligned} \langle n_c(\mathbf{r}_1, M) n_c(\mathbf{r}_2, M') \rangle &= \int \prod_{l=1}^N \{d^3 \mathbf{x}_l dM_l\} p(\mathbf{x}_1, \dots, \mathbf{x}_N, M_1, \dots, M_N) \\ &\quad \times \left[\sum_{i \neq j} \delta^D(\mathbf{r}_1 - \mathbf{x}_i) \delta^D(M - M_i) \delta^D(\mathbf{r}_2 - \mathbf{x}_j) \delta^D(M' - M_j) \right. \\ &\quad \left. + \sum_{i=j} \delta^D(\mathbf{r}_1 - \mathbf{x}_i) \delta^D(\mathbf{r}_2 - \mathbf{x}_i) \delta^D(M - M_i) \delta^D(M' - M_i) \right] \\ &= \sum_{i \neq j} p(\mathbf{r}_1, \mathbf{r}_2, M, M') + \sum_{i=j} p(\mathbf{r}_1, M) \delta^D(\mathbf{r}_1 - \mathbf{r}_2) \delta^D(M - M') . \end{aligned} \quad (18)$$

In order to proceed further, we need to specify the joint probability density distribution for obtaining clusters at positions \mathbf{r}_1 and \mathbf{r}_2 and with masses M_1 and M_2 . This we do through the introduction of correlation functions:

$$p(\mathbf{r}_1, \mathbf{r}_2, M, M') \equiv p(\mathbf{r}_1, M)p(\mathbf{r}_2, M') [1 + \xi^c(\mathbf{r}_1, \mathbf{r}_2, M, M')] , \quad (19)$$

where $\xi^c(\mathbf{r}_1, \mathbf{r}_2, M, M')$ is the two-point cross-correlation function of clusters with masses M and M' . On use of the above definition in Eq. (18), we find

$$\begin{aligned} \langle n_c(\mathbf{r}_1, M)n_c(\mathbf{r}_2, M') \rangle &= \bar{n}(N-1)p(\mathbf{r}_1, M)p(\mathbf{r}_2, M') [1 + \xi^c(\mathbf{r}_1, \mathbf{r}_2, M, M')] + Np(\mathbf{r}_1, M)\delta^D(\mathbf{r}_1 - \mathbf{r}_2)\delta^D(M - M') \\ &\approx \bar{n}(\mathbf{r}_1, M)\bar{n}(\mathbf{r}_2, M') [1 + \xi^c(\mathbf{r}_1, \mathbf{r}_2, M, M')] + \bar{n}(\mathbf{r}_1, M)\delta^D(\mathbf{r}_1 - \mathbf{r}_2)\delta^D(M - M') , \end{aligned} \quad (20)$$

where in arriving at the last equality we have assumed that $N \gg 1$. We may now write down directly the remaining expectation values that enter Eq. (16), whereupon:

$$\langle n_c(\mathbf{r}_1, M)n_s(\mathbf{r}_2, M') \rangle = \bar{n}(\mathbf{r}_1, M)\bar{n}_s(\mathbf{r}_2, M') = \alpha^{-1}\bar{n}(\mathbf{r}_1, M)\bar{n}(\mathbf{r}_2, M') ; \quad (21)$$

$$\begin{aligned} \langle n_s(\mathbf{r}_1, M)n_s(\mathbf{r}_2, M') \rangle &= \bar{n}_s(\mathbf{r}_1, M)\bar{n}_s(\mathbf{r}_2, M') + \bar{n}_s(\mathbf{r}_1, M)\delta^D(\mathbf{r}_1 - \mathbf{r}_2)\delta^D(M - M') \\ &= \alpha^{-2}\bar{n}(\mathbf{r}_1, M)\bar{n}(\mathbf{r}_2, M') + \alpha^{-1}\bar{n}(\mathbf{r}_1, M)\delta^D(\mathbf{r}_1 - \mathbf{r}_2)\delta^D(M - M') . \end{aligned} \quad (22)$$

On substituting the above relations back into Eq. (16), and on relabelling $M = M_1$ and $M' = M_2$, we find

$$\begin{aligned} \xi_{\mathcal{F}_c}(\mathbf{r}_1, \mathbf{r}_2) &= \frac{1}{A} \int dM_1 dM_2 w(\mathbf{r}_1, M_1)w(\mathbf{r}_2, M_2)\bar{n}(M_1)\bar{n}(M_2)\Theta(\mathbf{r}_1|M_1)\Theta(\mathbf{r}_2|M_2)\xi^c(\mathbf{r}_1, \mathbf{r}_2, M_1, M_2) \\ &\quad + \frac{(1+\alpha)}{A} \int dM_1 w^2(\mathbf{r}_1, M_1)\bar{n}(M_1)\Theta(\mathbf{r}_1|M_1)\delta^D(\mathbf{r}_1 - \mathbf{r}_2) , \end{aligned} \quad (23)$$

where we defined $\xi_{\mathcal{F}_c}(\mathbf{r}_1, \mathbf{r}_2) \equiv \langle \mathcal{F}_c(\mathbf{r}_1)\mathcal{F}_c(\mathbf{r}_2) \rangle$. If we assume that the cluster density field is some local function of the underlying dark matter density [9, 37–39], the cross-correlation function of clusters of masses M_1 and M_2 , at leading order, can be written:

$$\xi^c(|\mathbf{r}_1 - \mathbf{r}_2|, M_1, M_2) = b(M_1)b(M_2)\xi(|\mathbf{r}_1 - \mathbf{r}_2|) , \quad (24)$$

where $\xi(r)$ is the correlation of the underlying matter fluctuations. On inserting this relation into Eq. (23) we find:

$$\xi_{\mathcal{F}_c}(\mathbf{r}_1, \mathbf{r}_2) = \mathcal{G}_{(1,1)}(\mathbf{r}_1)\mathcal{G}_{(1,1)}(\mathbf{r}_2)\xi(|\mathbf{r}_1 - \mathbf{r}_2|) + (1+\alpha)\mathcal{G}_{(2,0)}(\mathbf{r}_1)\delta^D(\mathbf{r}_1 - \mathbf{r}_2) , \quad (25)$$

where we have defined the weighted selection function

$$\mathcal{G}_{(l,m)}(\mathbf{r}) \equiv \frac{1}{A^{l/2}} \int dM \bar{n}(M)b^m(M)w^l(\mathbf{r}, M)\Theta(\mathbf{r}|M) . \quad (26)$$

One possible estimator for the matter correlation function from the \mathcal{F}_c field, therefore, is

$$\hat{\xi}_{\mathcal{F}_c}(\mathbf{r}) \equiv \int d^3\mathbf{x} \mathcal{F}_c(\mathbf{x})\mathcal{F}_c(\mathbf{x} + \mathbf{r}) \quad ; \quad (\mathbf{r} \neq \mathbf{0}) . \quad (27)$$

If we now compute the expectation of this estimator we find:

$$\langle \hat{\xi}_{\mathcal{F}_c}(\mathbf{r}) \rangle = \int d^3\mathbf{x} \langle \mathcal{F}_c(\mathbf{x})\mathcal{F}_c(\mathbf{x} + \mathbf{r}) \rangle = \xi(\mathbf{r}) \int d^3\mathbf{x} \mathcal{G}_{(1,1)}(\mathbf{x})\mathcal{G}_{(1,1)}(\mathbf{x} + \mathbf{r}) \quad ; \quad (\mathbf{r} \neq \mathbf{0}) . \quad (28)$$

Thus $\hat{\xi}_{\mathcal{F}_c}(\mathbf{r})$ is not an unbiased estimator of the matter correlation function, nevertheless, we may construct one:

$$\hat{\xi}_0(\mathbf{r}) \equiv \hat{\xi}_{\mathcal{F}_c}(\mathbf{r})/\Xi_0(\mathbf{r}) \quad ; \quad \Xi_0(\mathbf{r}) \equiv \int d^3\mathbf{x} \mathcal{G}_{(1,1)}(\mathbf{x})\mathcal{G}_{(1,1)}(\mathbf{x} + \mathbf{r}) . \quad (29)$$

B. An estimator for the cluster power spectrum

We may now compute the Fourier space equivalent of the two-point correlation function, the power spectrum. In what follows we shall adopt the following Fourier transform conventions:

$$\mathcal{A}(\mathbf{k}) = \int d^3\mathbf{x} \mathcal{A}(\mathbf{x}) e^{i\mathbf{k}\cdot\mathbf{x}} \quad \Leftrightarrow \quad \mathcal{A}(\mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \mathcal{A}(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} . \quad (30)$$

We also define the power spectrum, $P_{\mathcal{A}\mathcal{A}}(k)$, of any infinite statistically homogeneous random field $\mathcal{A}(\mathbf{k})$ to be:

$$\langle \mathcal{A}(\mathbf{k}_1) \mathcal{A}(\mathbf{k}_2) \rangle \equiv (2\pi)^3 \delta^{\text{D}}(\mathbf{k}_1 + \mathbf{k}_2) P_{\mathcal{A}\mathcal{A}}(\mathbf{k}_1) . \quad (31)$$

Note, if the field \mathcal{A} were statistically isotropic, the power spectrum would simply be a function of the scalar k_1 .

With the above definitions in hand, the covariance of the Fourier modes of the cluster field \mathcal{F}_c can be written:

$$\begin{aligned} \langle \mathcal{F}_c(\mathbf{k}_1) \mathcal{F}_c(\mathbf{k}_2) \rangle &= \frac{1}{A} \int \prod_{i=1}^2 \{ d^3\mathbf{r}_i dM_i w(\mathbf{r}_i, M_i) \bar{n}(M_i) \Theta(\mathbf{r}_i | M_i) e^{\mathbf{k}_i \cdot \mathbf{r}_i} \} \xi^c(\mathbf{r}_1, \mathbf{r}_2, M_1, M_2) \\ &\quad + (1 + \alpha) \int d^3\mathbf{r}_1 dM_1 \bar{n}(M_1) \Theta(\mathbf{r}_1 | M_1) w^2(\mathbf{r}_1 | M_1) e^{i(\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r}_1} . \end{aligned} \quad (32)$$

For any infinite homogeneous random field, the two-point correlation function and the power spectrum form a Fourier pair, hence we may write:

$$\xi^c(\mathbf{r}_1, \mathbf{r}_2, M_1, M_2) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} P_{cc}(\mathbf{k}, M_1, M_2) e^{-i\mathbf{k} \cdot (\mathbf{r}_1 - \mathbf{r}_2)} . \quad (33)$$

If we assume the linear biasing relation of Eq. (24), then the cross-power spectrum of clusters of different masses M_1 and M_2 , at leading order, can be written

$$P(\mathbf{k}, M_1, M_2) = b(M_1) b(M_2) P(\mathbf{k}) , \quad (34)$$

where $P(\mathbf{k})$ is the matter power spectrum. On using these expressions in Eq. (32) and considering the case $\mathbf{k}_1 = \mathbf{k}_2$ we find:

$$\langle |\mathcal{F}_c(\mathbf{k})|^2 \rangle = \int \frac{d^3\mathbf{q}}{(2\pi)^3} P(\mathbf{q}) |\mathcal{G}_{(1,1)}(\mathbf{k} - \mathbf{q})|^2 + (1 + \alpha) \mathcal{G}_{(2,0)}(\mathbf{0}) , \quad (35)$$

where in the above expression $\mathcal{G}_{(l,m)}(\mathbf{k})$ is the Fourier transform of the weighted survey selection function from Eq. (26). We thus see that, as in the case of galaxies [2], the expectation of the square amplitude of the Fourier modes of \mathcal{F}_c are given by the convolution of the matter power spectrum with the modulus square of the Fourier modes of the survey window function, plus a constant shot noise.

In the limit that the survey volume is large, the functions $\mathcal{G}_{(k,l)}(\mathbf{k})$ will all be very narrowly peaked around $\mathbf{k} = \mathbf{0}$. Provided the matter power spectrum is a smoothly varying function of scale, the window functions $\mathcal{G}_{(k,l)}(\mathbf{k})$ take on Dirac delta-function-like behaviour. Consider again the first term on the right-hand-side of Eq. (35), in the large-survey volume limit, this becomes:

$$\int \frac{d^3\mathbf{q}}{(2\pi)^3} P(\mathbf{q}) |\mathcal{G}_{(1,1)}(\mathbf{k} - \mathbf{q})|^2 \approx P(\mathbf{k}) \int \frac{d^3\mathbf{q}}{(2\pi)^3} |\mathcal{G}_{(1,1)}(\mathbf{k} - \mathbf{q})|^2 . \quad (36)$$

Let us focus on the integral factor on the right-hand-side of the above expression. If we now perform the transformation of variables $\mathbf{q}' \rightarrow \mathbf{k} - \mathbf{q}$ and use Parseval's theorem, then we find:

$$\int \frac{d^3\mathbf{q}}{(2\pi)^3} |\mathcal{G}_{(1,1)}(\mathbf{k} - \mathbf{q})|^2 = \int \frac{d^3\mathbf{q}'}{(2\pi)^3} |\mathcal{G}_{(1,1)}(\mathbf{q}')|^2 = \int d^3\mathbf{r} \mathcal{G}_{(1,1)}^2(\mathbf{r}) . \quad (37)$$

Upon back-substitution of Eq. (26), the above expression expands to:

$$\int d^3\mathbf{r} \mathcal{G}_{(1,1)}^2(\mathbf{r}) = \frac{1}{A} \int d^3\mathbf{r} \left[\int dM \bar{n}(M) b(M) w(\mathbf{r}, M) \Theta(\mathbf{r} | M) \right]^2 . \quad (38)$$

Recall that we have not yet specified the parameter A , let us now define this to be:

$$A \equiv \int d^3\mathbf{r} \left[\int dM \bar{n}(M) b(M) w(\mathbf{r}, M) \Theta(\mathbf{r}, M) \right]^2 . \quad (39)$$

On adopting this normalization we see that Eq. (38) is simply unity. Hence, for the case of large homogeneous survey volumes, an unbiased estimator for the dark matter power spectrum is:

$$\hat{P}(\mathbf{k}) \approx |\mathcal{F}_c(\mathbf{k})|^2 - P_c^{\text{shot}} , \quad (40)$$

where we defined

$$P_c^{\text{shot}} \equiv (1 + \alpha) \mathcal{G}_{(2,0)}(\mathbf{0}) = (1 + \alpha) \frac{\int d^3\mathbf{r} \int dM \bar{n}(M) \Theta(\mathbf{r}, M) w^2(\mathbf{r}, M)}{\int d^3\mathbf{r} \left[\int dM \bar{n}(M) b(M) \Theta(\mathbf{r}, M) w(\mathbf{r}, M) \right]^2} . \quad (41)$$

In fact it is not precisely the above estimator that we are interested in, since this gives the power at a particular mode, but instead, the band-power estimates of the power spectrum. Thus, our final estimator is:

$$\hat{\bar{P}}(k_i) = \frac{1}{V_i} \int_{V_i} d^3\mathbf{k} \hat{P}(\mathbf{k}) = \frac{1}{V_i} \int_{V_i} d^3\mathbf{k} |F_c(\mathbf{k})|^2 - P_c^{\text{shot}} , \quad (42)$$

where in the above we have summed all modes over a shell in k -space of thickness Δk , i.e.

$$V_i \equiv \int_{V_i} d^3\mathbf{k} = 4\pi \int_{k_i - \Delta k/2}^{k_i + \Delta k/2} k^2 dk = 4\pi k_i^2 \Delta k \left[1 + \frac{1}{12} \left(\frac{\Delta k}{k_i} \right)^2 \right] . \quad (43)$$

If the survey window function possesses small scale structure, then the matter power spectrum can only safely be recovered by deconvolution of the window function $|\mathcal{G}_{(1,1)}(\mathbf{k})|^2$, or alternatively one must convolve theory predictions with the window function. Otherwise, Eq. (40) is a biased estimator.

IV. STATISTICAL FLUCTUATIONS IN THE CLUSTER POWER SPECTRUM

In order to obtain the optimal estimator we need to know how the \mathcal{S}/\mathcal{N} varies when we vary the shape of our weight function $w(\mathbf{x}, M)$. Hence, we need to understand the noise properties of our power spectrum estimator.

A. The covariance of the cluster power spectrum estimator

In general, the covariance matrix of the band-power spectrum estimator can be written as:

$$\begin{aligned} \text{Cov}[\hat{\bar{P}}(k_i), \hat{\bar{P}}(k_j)] &\equiv \langle \hat{\bar{P}}(k_i) \hat{\bar{P}}(k_j) \rangle - \langle \hat{\bar{P}}(k_i) \rangle \langle \hat{\bar{P}}(k_j) \rangle \\ &= \frac{1}{V_i} \int_{V_i} d^3\mathbf{k}_1 \frac{1}{V_j} \int_{V_j} d^3\mathbf{k}_2 \text{Cov}[\hat{P}(\mathbf{k}_1), \hat{P}(\mathbf{k}_2)] . \end{aligned} \quad (44)$$

This in turn may be written in terms of the covariance of the power spectrum of two different Fourier modes:

$$\text{Cov}[\hat{P}(\mathbf{k}_1), \hat{P}(\mathbf{k}_2)] = \text{Cov}[|\mathcal{F}_c(\mathbf{k}_1)|^2, |\mathcal{F}_c(\mathbf{k}_2)|^2] + \text{Var}[P_c^{\text{shot}}] - \text{Cov}[|\mathcal{F}_c(\mathbf{k}_1)|^2, P_c^{\text{shot}}] - \text{Cov}[|\mathcal{F}_c(\mathbf{k}_2)|^2, P_c^{\text{shot}}] \quad (45)$$

where

$$\text{Cov}[|\mathcal{F}_c(\mathbf{k}_1)|^2, |\mathcal{F}_c(\mathbf{k}_2)|^2] \equiv \langle |\mathcal{F}_c(\mathbf{k}_1)|^2 |\mathcal{F}_c(\mathbf{k}_2)|^2 \rangle - \langle |\mathcal{F}_c(\mathbf{k}_1)|^2 \rangle \langle |\mathcal{F}_c(\mathbf{k}_2)|^2 \rangle ; \quad (46)$$

$$\text{Var}[P_c^{\text{shot}}] \equiv \langle [P_c^{\text{shot}}]^2 \rangle - \langle P_c^{\text{shot}} \rangle^2 ; \quad (47)$$

$$\text{Cov}[|\mathcal{F}_c(\mathbf{k}_i)|^2, P_c^{\text{shot}}] \equiv \langle |\mathcal{F}_c(\mathbf{k}_i)|^2 P_c^{\text{shot}} \rangle - \langle |\mathcal{F}_c(\mathbf{k}_i)|^2 \rangle \langle P_c^{\text{shot}} \rangle ; \quad i \in (1, 2) . \quad (48)$$

Assuming that the total number density of clusters is a deterministic quantity, all of the covariance terms involving P_c^{shot} vanish. Thus we are left with the task of determining the covariance of the modulus square of the Fourier modes of the \mathcal{F}_c field. In order to write this succinctly, we have made use of the following short-hand-notation:

$$\bar{n}_i \equiv \bar{n}(M_i) ; \quad w_i \equiv w(\mathbf{r}_i, M_i) ; \quad b_i \equiv b(M_i) ; \quad \Theta_i \equiv \Theta(\mathbf{r}_i | M_i) ; \quad \delta_{r,ij}^D \equiv \delta^D(\mathbf{r}_i - \mathbf{r}_j) ; \quad \delta_{M,ij}^D \equiv \delta^D(M_i - M_j) . \quad (49)$$

In Appendix A 1 we present the derivation of the covariance, and the final result is:

$$\begin{aligned}
\text{Cov}[|\mathcal{F}_c(\mathbf{k}_1)|^2, |\mathcal{F}_c(\mathbf{k}_2)|^2] &= \frac{1}{A^2} \int d^3\mathbf{k}_3 d^3\mathbf{k}_4 \delta^D(\mathbf{k}_1 + \mathbf{k}_3) \delta^D(\mathbf{k}_2 + \mathbf{k}_4) \int \prod_{j=1}^4 \{d^3\mathbf{r}_j dM_j w_j \bar{n}_j \Theta_j e^{i\mathbf{k}_j \cdot \mathbf{r}_j}\} \\
&\left\{ \eta_{1234}^c + \left[\xi_{12}^c + \frac{(1+\alpha)}{\bar{n}_2 \Theta_2} \delta_{r,12}^D \delta_{M,12}^D \right] \left[\xi_{34}^c + \frac{(1+\alpha)}{\bar{n}_4 \Theta_4} \delta_{r,34}^D \delta_{M,34}^D \right] \right. \\
&+ \left[\xi_{14}^c + \frac{(1+\alpha)}{\bar{n}_4 \Theta_4} \delta_{r,14}^D \delta_{M,14}^D \right] \left[\xi_{23}^c + \frac{(1+\alpha)}{\bar{n}_3 \Theta_3} \delta_{r,23}^D \delta_{M,23}^D \right] + \zeta_{134}^c \frac{\delta_{r,12}^D \delta_{M,12}^D}{\bar{n}_2 \Theta_2} + \zeta_{124}^c \frac{\delta_{r,23}^D \delta_{M,23}^D}{\bar{n}_3 \Theta_3} \\
&+ \zeta_{123}^c \frac{\delta_{r,14}^D \delta_{M,14}^D}{\bar{n}_4 \Theta_4} + \zeta_{124}^c \frac{\delta_{r,13}^D \delta_{M,13}^D}{\bar{n}_3 \Theta_3} + \zeta_{123}^c \frac{\delta_{r,24}^D \delta_{M,24}^D}{\bar{n}_4 \Theta_4} + \zeta_{123}^c \frac{\delta_{r,34}^D \delta_{M,34}^D}{\bar{n}_4 \Theta_4} \\
&+ \zeta_{13}^c \frac{\delta_{r,12}^D \delta_{r,34}^D \delta_{M,12}^D \delta_{M,34}^D}{\bar{n}_2 \bar{n}_4 \Theta_2 \Theta_4} + \zeta_{12}^c \frac{\delta_{r,13}^D \delta_{r,24}^D \delta_{M,13}^D \delta_{M,24}^D}{\bar{n}_3 \bar{n}_4 \Theta_3 \Theta_4} + \zeta_{12}^c \frac{\delta_{r,14}^D \delta_{r,23}^D \delta_{M,14}^D \delta_{M,23}^D}{\bar{n}_3 \bar{n}_4 \Theta_3 \Theta_4} \\
&+ \zeta_{14}^c \frac{\delta_{r,12}^D \delta_{r,13}^D \delta_{M,12}^D \delta_{M,13}^D}{\bar{n}_2 \bar{n}_3 \Theta_2 \Theta_3} + \zeta_{13}^c \frac{\delta_{r,12}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,14}^D}{\bar{n}_2 \bar{n}_4 \Theta_2 \Theta_4} + \zeta_{12}^c \frac{\delta_{r,13}^D \delta_{r,14}^D \delta_{M,13}^D \delta_{M,14}^D}{\bar{n}_3 \bar{n}_4 \Theta_3 \Theta_4} \\
&\left. + \zeta_{12}^c \frac{\delta_{r,23}^D \delta_{M,23}^D \delta_{r,24}^D \delta_{M,24}^D}{\bar{n}_3 \bar{n}_4 \Theta_3 \Theta_4} + \frac{(1+\alpha^3) \delta_{r,12}^D \delta_{r,13}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,13}^D \delta_{M,14}^D}{\bar{n}_2 \bar{n}_3 \bar{n}_4 \Theta_2 \Theta_3 \Theta_4} \right\}. \quad (50)
\end{aligned}$$

In the above ξ_{ij}^c , ζ_{ijk}^c and η_{ijkl}^c represent the true three- and four-point connected correlation functions of galaxy clusters, respectively. Under the assumption of linear biasing these may be written in terms of the connected correlation functions of the matter as:

$$\xi_{ij}^c \equiv \xi_c(\mathbf{r}_i, \mathbf{r}_j, M_i, M_j) = b(M_i) b(M_j) \xi_{ij} = b_i b_j \xi_{ij}; \quad (51)$$

$$\zeta_{ijk}^c \equiv \zeta_c(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k, M_i, M_j, M_k) = b(M_i) b(M_j) b(M_k) \zeta_{ijk} = b_i b_j b_k \zeta_{ijk}; \quad (52)$$

$$\eta_{ijkl}^c \equiv \eta_c(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k, \mathbf{r}_l, M_i, M_j, M_k, M_l) = b(M_i) b(M_j) b(M_k) b(M_l) \eta_{ijkl} = b_i b_j b_k b_l \eta_{ijkl}. \quad (53)$$

In Appendix A 2 we provide a general relation for the covariance of the \mathcal{F}_c power spectrum in terms of Fourier space quantities. These results generalize the expressions for the covariance matrix of the cluster power spectrum presented in [34], extending that work to the case of finite survey geometry and an arbitrary weighting scheme that depends on mass and position.

Under the assumption that the matter density field is Gaussianly distributed, i.e. all connected correlation functions beyond the two-point vanish $\zeta = \eta = 0$. The general expression simplifies to:

$$\begin{aligned}
\text{Cov}[|\mathcal{F}_c(\mathbf{k}_1)|^2, |\mathcal{F}_c(\mathbf{k}_2)|^2] &= \left| \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} P(\mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_2 + \mathbf{q}_1) + (1+\alpha) \mathcal{G}_{(2,0)}(\mathbf{k}_1 + \mathbf{k}_2) \right|^2 \\
&+ \left| \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} P(\mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_2 + \mathbf{q}_1) + (1+\alpha) \mathcal{G}_{(2,0)}(\mathbf{k}_1 - \mathbf{k}_2) \right|^2 \\
&+ \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} P(\mathbf{q}_1) \left\{ |\mathcal{G}_{(2,1)}(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}_1)|^2 + |\mathcal{G}_{(2,1)}(\mathbf{q}_1)|^2 + |\mathcal{G}_{(2,1)}(\mathbf{k}_1 - \mathbf{k}_2 - \mathbf{q}_1)|^2 \right. \\
&+ \mathcal{G}_{(3,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_1 + \mathbf{q}_1) + \mathcal{G}_{(3,1)}(\mathbf{k}_2 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_2 + \mathbf{q}_1) \\
&\left. + \mathcal{G}_{(3,1)}(-\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_1 + \mathbf{q}_1) + \mathcal{G}_{(3,1)}(-\mathbf{k}_2 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_2 + \mathbf{q}_1) \right\} + (1+\alpha^3) \mathcal{G}_{(4,0)}(0). \quad (54)
\end{aligned}$$

Let us now consider the covariance for the case where the survey volume is large and hence the $\mathcal{G}_{j,k}(\mathbf{k})$ functions are

all very narrowly peaked around $\mathbf{k} = 0$, whereupon

$$\begin{aligned}
\text{Cov}[|\mathcal{F}_c(\mathbf{k}_1)|^2, |\mathcal{F}_c(\mathbf{k}_2)|^2] &\approx \left| P(\mathbf{k}_1) \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \mathcal{G}_{(1,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_2 + \mathbf{q}_1) + (1 + \alpha) \mathcal{G}_{(0,2)}(\mathbf{k}_1 + \mathbf{k}_2) \right|^2 \\
&+ \left| P(\mathbf{k}_1) \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \mathcal{G}_{(1,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_2 + \mathbf{q}_1) + (1 + \alpha) \mathcal{G}_{(0,2)}(\mathbf{k}_1 - \mathbf{k}_2) \right|^2 \\
&+ P(\mathbf{k}_1 + \mathbf{k}_2) \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} |\mathcal{G}_{(2,1)}(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}_1)|^2 + P(\mathbf{0}) \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} |\mathcal{G}_{(2,1)}(\mathbf{q}_1)|^2 \\
&+ P(\mathbf{k}_1 - \mathbf{k}_2) \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} |\mathcal{G}_{(2,1)}(\mathbf{k}_1 - \mathbf{k}_2 - \mathbf{q}_1)|^2 + P(\mathbf{k}_1) \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \mathcal{G}_{(3,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_1 + \mathbf{q}_1) \\
&+ P(\mathbf{k}_2) \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \mathcal{G}_{(3,1)}(\mathbf{k}_2 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_2 + \mathbf{q}_1) + P(-\mathbf{k}_1) \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \mathcal{G}_{(3,1)}(-\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_1 + \mathbf{q}_1) \\
&+ P(-\mathbf{k}_2) \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \mathcal{G}_{(3,1)}(-\mathbf{k}_2 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_2 + \mathbf{q}_1) + (1 + \alpha^3) \mathcal{G}_{(4,0)}(\mathbf{0}) . \tag{55}
\end{aligned}$$

It is now useful to prove a general relation concerning the $\mathcal{G}_{(k,l)}$ functions. Let us define the following function:

$$\mathcal{Q}_{(j_1, \dots, j_n)}^{(i_1, \dots, i_n)}(\mathbf{x}) \equiv \mathcal{G}_{(i_1, j_1)}(\mathbf{x}) \dots \mathcal{G}_{(i_n, j_n)}(\mathbf{x}) \tag{56}$$

On Fourier transforming the above relation, and recalling the convolution theorem, the above expression becomes

$$\mathcal{Q}_{(j_1, \dots, j_n)}^{(i_1, \dots, i_n)}(\mathbf{k}) \equiv \int \frac{d^3 \mathbf{q}_1}{(2\pi)^3} \dots \frac{d^3 \mathbf{q}_{n-1}}{(2\pi)^3} \mathcal{G}_{(i_1, j_1)}(\mathbf{q}_1) \dots \mathcal{G}_{(i_{n-1}, j_{n-1})}(\mathbf{q}_{n-1}) \mathcal{G}_{(i_n, j_n)}(\mathbf{k} - \mathbf{q}_1 - \dots - \mathbf{q}_{n-1}) \tag{57}$$

Using this definition, noting that $P(\mathbf{0}) = 0$, and after performing a transformation of variables in the integrals over \mathbf{q}_1 , we find that Eq. (55) can be expressed more compactly as:

$$\begin{aligned}
\text{Cov}[|\mathcal{F}_c(\mathbf{k}_1)|^2, |\mathcal{F}_c(\mathbf{k}_2)|^2] &\approx \left| P(\mathbf{k}_1) \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{k}_1 + \mathbf{k}_2) + (1 + \alpha) \mathcal{Q}_{(0)}^{(2)}(\mathbf{k}_1 + \mathbf{k}_2) \right|^2 \\
&+ \left| P(\mathbf{k}_1) \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{k}_1 - \mathbf{k}_2) + (1 + \alpha) \mathcal{Q}_{(0)}^{(2)}(\mathbf{k}_1 - \mathbf{k}_2) \right|^2 \\
&+ \mathcal{Q}_{(1,1)}^{(2,2)}(\mathbf{0}) \left[P(\mathbf{k}_1 + \mathbf{k}_2) + P(\mathbf{k}_1 - \mathbf{k}_2) \right] \\
&+ \mathcal{Q}_{(1,1)}^{(3,1)}(\mathbf{0}) \left[P(\mathbf{k}_1) + P(\mathbf{k}_2) + P(-\mathbf{k}_1) + P(-\mathbf{k}_2) \right] + (1 + \alpha^3) \mathcal{Q}_{(0)}^{(4)}(\mathbf{0}) . \tag{58}
\end{aligned}$$

We may now compute the bin-averaged estimates of the \mathcal{F}_c power spectra:

$$\begin{aligned}
\text{Cov}[|\mathcal{F}_c(k_i)|^2, |\mathcal{F}_c(k_j)|^2] &= 2 \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \left| P(\mathbf{k}_1) \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{k}_1 + \mathbf{k}_2) + (1 + \alpha) \mathcal{Q}_{(0)}^{(2)}(\mathbf{k}_1 + \mathbf{k}_2) \right|^2 \\
&+ 2 \mathcal{Q}_{(1,1)}^{(2,2)}(\mathbf{0}) \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} P(\mathbf{k}_1 + \mathbf{k}_2) \\
&+ 2 \mathcal{Q}_{(1,1)}^{(3,1)}(\mathbf{0}) \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \left[P(\mathbf{k}_1) + P(\mathbf{k}_2) \right] + (1 + \alpha^3) \mathcal{Q}_{(0)}^{(4)}(\mathbf{0}) . \tag{59}
\end{aligned}$$

On expanding the modulus squared term and assuming that the k -space shells are sufficiently narrow that the power spectrum can be considered constant over the shell, then the above equation may be approximated by,

$$\begin{aligned}
\text{Cov}[|\mathcal{F}_c(k_i)|^2, |\mathcal{F}_c(k_j)|^2] &\approx 2 \bar{P}^2(k_i) \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{k}_1 + \mathbf{k}_2) \mathcal{Q}_{(1,1)}^{(1,1)}(-\mathbf{k}_1 - \mathbf{k}_2) \\
&+ 4(1 + \alpha) \bar{P}(k_i) \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{k}_1 + \mathbf{k}_2) \mathcal{Q}_{(0)}^{(2)}(-\mathbf{k}_1 - \mathbf{k}_2) \\
&+ 2(1 + \alpha)^2 \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \mathcal{Q}_{(0)}^{(2)}(\mathbf{k}_1 + \mathbf{k}_2) \mathcal{Q}_{(0)}^{(2)}(-\mathbf{k}_1 - \mathbf{k}_2) \\
&+ 2 \mathcal{Q}_{(1,1)}^{(2,2)}(\mathbf{0}) \bar{P}[k_i, k_j] + 2 \mathcal{Q}_{(1,1)}^{(3,1)}(\mathbf{0}) \left[\bar{P}(k_i) + \bar{P}(k_j) \right] + (1 + \alpha^3) \mathcal{Q}_{(0)}^{(4)}(\mathbf{0}) . \tag{60}
\end{aligned}$$

where we defined

$$\bar{P}[k_i, k_j] \equiv \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} P(\mathbf{k}_1 + \mathbf{k}_2) . \quad (61)$$

In Appendix A 3 we prove a general result concerning the k -space shell averaging of products of the \mathcal{Q} functions. However, to simplify the analysis let us consider the first term in Eq. (60), this can be written:

$$\mathcal{P}_1[k_i, k_j] \equiv \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{k}_1 + \mathbf{k}_2) \mathcal{Q}_{(1,1)}^{(1,1)}(-\mathbf{k}_1 - \mathbf{k}_2) \quad (62)$$

On Fourier transforming the \mathcal{Q} functions and integrating the exponential functions over the angles $\hat{\mathbf{k}}_1$ and $\hat{\mathbf{k}}_2$ the above expression simplifies to:

$$\begin{aligned} \mathcal{P}_1[k_i, k_j] &= \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \int d^3 \mathbf{x}_1 d^3 \mathbf{x}_2 \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{x}_1) \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{x}_2) e^{i(\mathbf{k}_1 + \mathbf{k}_2) \cdot (\mathbf{x}_1 - \mathbf{x}_2)} \\ &= \int d^3 \mathbf{x}_1 d^3 \mathbf{x}_2 \bar{j}_0(k_i |\mathbf{x}_1 - \mathbf{x}_2|) \bar{j}_0(k_j |\mathbf{x}_1 - \mathbf{x}_2|) \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{x}_1) \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{x}_2) , \end{aligned} \quad (63)$$

where in the above equation we defined the shell-average of the spherical Bessel function as

$$\bar{j}_0(k_i x) \equiv \frac{1}{V_i} \int_{k_i - \Delta k/2}^{k_i + \Delta k/2} dk_1 k_1^2 4\pi j_0(k_1 x) . \quad (64)$$

On performing a change of variables $\mathbf{x}_{21} = \mathbf{x}_2 - \mathbf{x}_1$, then the above integral further simplifies to:

$$\mathcal{P}_1[k_i, k_j] = \int d^3 \mathbf{x}_{21} \bar{j}_0(k_i x_{21}) \bar{j}_0(k_j x_{21}) \Xi_1(x_{21}) , \quad (65)$$

where in the above we have defined the correlation function of the weighted survey window function to be,

$$\begin{aligned} \Xi_1(x_{21}) &\equiv \int \frac{d^2 \hat{\mathbf{x}}_{21}}{4\pi} \int d^3 \mathbf{x}_1 \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{x}_1) \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{x}_{21} + \mathbf{x}_1) \\ &= \int \frac{d^2 \hat{\mathbf{x}}_{21}}{4\pi} \int d^3 \mathbf{x}_1 \mathcal{G}_{(1,1)}^2(\mathbf{x}_1) \mathcal{G}_{(1,1)}^2(\mathbf{x}_{21} + \mathbf{x}_1) . \end{aligned} \quad (66)$$

In the limit that the survey volume is large, the weighted survey window correlation function is very slowly varying over nearly all length scales of interest, and so can be approximated by its value at zero-lag. Hence,

$$\mathcal{P}_1[k_i, k_j] \approx \Xi_1(0) \int d^3 \mathbf{x}_{21} \bar{j}_0(x_{21} k_i) \bar{j}_0(x_{21} k_j) . \quad (67)$$

The spherical Bessel functions obey the following orthogonality relationship:

$$\int_0^\infty dx x^2 j_\alpha(ux) j_\alpha(vx) = \frac{\pi}{2u^2} \delta^D(u - v) . \quad (68)$$

Using this relationship in Eq. (67) gives,

$$\begin{aligned} \mathcal{P}_1[k_i, k_j] &\approx \Xi_1(0) \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \int_0^\infty dr_{21} 4\pi r_{21}^2 j_0(r_{21} k_1) j_0(r_{21} k_2) \\ &= \Xi_1(0) \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \frac{2\pi^2}{k_1^2} \delta^D(k_1 - k_2) . \end{aligned} \quad (69)$$

We note that the delta function will only give a non-zero value when the two k -space shells are overlapping, i.e. when $V_i = V_j$. Thus the above expression simplifies to:

$$\mathcal{P}_1[k_i, k_j] \approx \Xi_1(0) \delta_{i,j}^K \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i^2} 4\pi k_1^2 \frac{2\pi^2}{k_1^2} = \frac{(2\pi)^3}{V_i} \Xi_1(0) \delta_{i,j}^K . \quad (70)$$

The second and third terms in Eq. (60) can be computed in a similar fashion and are presented in Appendix A 3.

Returning now to our expression for the covariance matrix of the power spectra of the field \mathcal{F}_c given by Eq. (60), we see that the first three terms on the right-hand-side may now be recast:

$$\begin{aligned} \text{Cov}[|\mathcal{F}_c(k_i)|^2, |\mathcal{F}_c(k_j)|^2] \approx & \frac{2(2\pi)^3}{V_i} \left[\overline{P}^2(k_i) \Xi_1(0) + 2(1+\alpha) \overline{P}(k_i) \Xi_2(0) + (1+\alpha)^2 \Xi_3(0) \right] \delta_{i,j}^K \\ & + 2\mathcal{Q}_{(1,1)}^{(2,2)}(0) \overline{P}[k_i, k_j] + 2\mathcal{Q}_{(1,1)}^{(3,1)}(0) [\overline{P}(k_i) + \overline{P}(k_j)] + (1+\alpha^3) \mathcal{Q}_{(0)}^{(4)}(0), \end{aligned} \quad (71)$$

where $\Xi_2(0)$ and $\Xi_3(0)$ are defined in Appendix A 3. Finally, on taking the limit that $\bar{n}_h V_\mu \gg 1$, then the last three terms will be sub-dominant [34]. We may now use Eqs (56) and (66) to re-express the above equation in the more simplified form:

$$\text{Cov}[|F(k_i)|^2, |F(k_j)|^2] \approx \frac{2(2\pi)^3}{V_i} \overline{P}^2(k_i) \int d^3 \mathbf{x}_1 \left\{ [\mathcal{G}_{(1,1)}(\mathbf{x}_1)]^2 + \frac{(1+\alpha)}{\overline{P}(k_i)} \mathcal{G}_{(2,0)}(\mathbf{x}_1) \right\}^2 \delta_{i,j}^K. \quad (72)$$

Let us now make an important point about our derivation of the statistical fluctuations in the cluster power spectrum. The weights derived in the next section are only optimal under some important approximations: the underlying matter density is Gaussian; the sampling fluctuations for a given realization of the cluster field are small; the survey volume is large and homogeneous. In future work it will be interesting to explore how the optimal weights change as each of these assumptions is gradually relaxed.

At this point it is also worth comparing our analysis with that of FKP and PVP. Our expressions for the correlation function Eq. (29) and the power spectrum Eq. (40) are analogous to those found in PVP for the luminosity dependence of the galaxy bias. However, our derivation is different, particularly in the way in which we treat the statistical properties of the cluster field, which we believe to be more general and transparent. In addition our choice of the normalization parameter A is different from PVP. Until $w(\mathbf{r}, M)$ has been selected this is of no relevance. Indeed, in the next section we will show that, once the optimal weights are selected, their and our choice for the field \mathcal{F}_c turn out to be equivalent. We also note that our derivation of the statistical fluctuations in the power and the approximations that we make in order to obtain a diagonal covariance matrix, are again different from the derivation of FKP. Again, we feel that they are more rigorous and transparent. For instance, the original FKP derivation makes use of Parseval's theorem to transform from a band-power k -space integral to an integral over the entirety of real space – that is they effectively go from Eq. (60) to Eq. (72) using Parseval's theorem. Owing to the fact that the band-power averages do not extend over the whole of k -space, indeed they may be limited to lie only in a very narrow shell, the application of Parseval's theorem here appears incorrect. As a more minor point, their derivation also misses a factor of 2, which arises due to the Hermitian nature of the Fourier modes.

V. AN OPTIMAL WEIGHTING SCHEME

A. Optimal weights as a functional problem

Our aim is to find the optimal weighting scheme that will maximize the S/\mathcal{N} ratio on a given band power estimate of the cluster power spectrum. To begin, note that maximizing the S/\mathcal{N} ratio is equivalent to minimizing its inverse, the noise-over-signal ratio \mathcal{N}/S . This can be expressed as:

$$F[w(\mathbf{x})] \equiv \frac{\sigma_P^2(k_i)}{\overline{P}^2(k_i)} = \frac{2(2\pi)^3}{V_i} \int d^3 \mathbf{x}_1 \left\{ [\mathcal{G}_{(1,1)}(\mathbf{x}_1)]^2 + \frac{(1+\alpha)}{\overline{P}(k_i)} \mathcal{G}_{(2,0)}(\mathbf{x}_1) \right\}^2. \quad (73)$$

In the above expression we have written the quantity $F[w]$ as a functional of the weights $w(\mathbf{x}, M)$. Let us consider the right-hand-side of this expression in more detail, and let us define four functionals:

$$S_1[w] \equiv \int d^3 \mathbf{x} \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \right]^4; \quad (74)$$

$$S_2[w] \equiv \int d^3 \mathbf{x} \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \right]^2 \tilde{\mathcal{G}}_{(2,0)}(\mathbf{x}); \quad (75)$$

$$S_3[w] \equiv \int d^3 \mathbf{x} \left[\tilde{\mathcal{G}}_{(2,0)}(\mathbf{x}) \right]^2; \quad (76)$$

$$S_4[w] \equiv \left\{ \int d^3 \mathbf{x} \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \right]^2 \right\}^2, \quad (77)$$

where in the above we have introduced a scaled version of the survey window function:

$$\tilde{\mathcal{G}}_{(i,j)} \equiv A^{i/2} \mathcal{G}_{(i,j)} . \quad (78)$$

The right-hand-side of Eq. (73) can now be re-expressed in terms of our set of functionals $S_i[w]$ in the following manner:

$$F[w(\mathbf{x})] = \frac{C_1}{S_4[w]} \{ S_1[w] + 2C_2 S_2[w] + C_2^2 S_3[w] \} , \quad (79)$$

where we have introduced the two parameters:

$$C_1 \equiv 2(2\pi)^3/V_i ; \quad C_2 \equiv (1 + \alpha)/\bar{P}(k_i) . \quad (80)$$

Eq. (79) expresses $F[w]$ as a function of four functionals of the weights $w(\mathbf{x}, M)$. The standard way for finding the optimal weights, is to perform the functional variation of F with respect to the weights w .

B. The optimal weight equation

Operationally, the functional variation of $F[w]$ may be defined:

$$dF[w(\mathbf{x}, M)] = F[w(\mathbf{x}, M)] - F[w(\mathbf{x}, M) + \delta w(\mathbf{x}, M)] = \int d^3 \mathbf{x} dM \left\{ \frac{\delta F}{\delta w(\mathbf{x}, M)} \right\} \delta w(\mathbf{x}, M) . \quad (81)$$

The extremisation condition is that the functional derivative is stationary for the optimal weights:

$$\frac{\delta F}{\delta w(\mathbf{x}, M)} = 0 . \quad (82)$$

In Appendix B we show that, at the extremum $dF[w] = 0$, the functional derivatives of the $S_i[w]$ must be related to one another through the expression:

$$\frac{\delta S_1/\delta w + 2C_2 \delta S_2/\delta w + C_2^2 \delta S_3/\delta w}{S_1 + 2C_2 S_2 + C_2^2 S_3} = \frac{\delta S_4}{S_4} . \quad (83)$$

Also, in Appendix B we provide details of the functional derivatives of the $S_i[w]$, whose final expressions are:

$$\frac{\delta S_1}{\delta w(\mathbf{x}, M)} = 4 \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \right]^3 \bar{n}(M) b(M) \Theta(\mathbf{x}|M) ; \quad (84)$$

$$\frac{\delta S_2}{\delta w(\mathbf{x}, M)} = 2\bar{n}(M) b(M) \Theta(\mathbf{x}|M) \left\{ \tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \tilde{\mathcal{G}}_{(2,0)}(\mathbf{x}) + \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \right]^2 \frac{w(\mathbf{x}, M)}{b(M)} \right\} ; \quad (85)$$

$$\frac{\delta S_3}{\delta w(\mathbf{x}, M)} = 4\tilde{\mathcal{G}}_{(2,0)}(\mathbf{x}) \bar{n}(M) \Theta(\mathbf{x}|M) w(\mathbf{x}, M) ; \quad (86)$$

$$\frac{\delta S_4}{\delta w(\mathbf{x}, M)} = 4\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \bar{n}(M) b(M) \Theta(\mathbf{x}|M) \int d^3 \mathbf{x}' \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}') \right]^2 . \quad (87)$$

Consider the numerator on the left-hand-side of Eq. (83), and let us write this in the following manner:

$$\frac{\delta S_1}{\delta w} + 2C_2 \frac{\delta S_2}{\delta w} + C_2^2 \frac{\delta S_3}{\delta w} \equiv 4\bar{n}(M) b(M) \Theta(\mathbf{x}, M) \tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \mathcal{K}(\mathbf{x}, M) , \quad (88)$$

where we have defined,

$$\mathcal{K}(\mathbf{x}, M) \equiv \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \right]^2 + C_2 \left[\tilde{\mathcal{G}}_{(2,0)}(\mathbf{x}) + \tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \frac{w(\mathbf{x}, M)}{b(M)} \right] + C_2^2 \frac{\tilde{\mathcal{G}}_{(2,0)}(\mathbf{x}) w(\mathbf{x}, M)}{\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) b(M)} . \quad (89)$$

On substituting Eqns (84), (85), (86), (87) and (88) into Eq. (83) we find,

$$\frac{4\bar{n}(M) b(M) \Theta(\mathbf{x}, M) \tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \mathcal{K}(\mathbf{x}, M)}{S_1[w] + C_2 S_2[w] + C_3 S_3[w]} = \frac{4\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \bar{n}(M) b(M) \Theta(\mathbf{x}, M) \int d^3 \mathbf{x}' \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}') \right]^2}{S_4[w]} . \quad (90)$$

On canceling terms and rearranging this becomes:

$$\mathcal{K}(\mathbf{x}, M) = \frac{S_1[w] + C_2 S_2[w] + C_3 S_3[w]}{S_4[w]} \int d^3 \mathbf{x}' \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}') \right]^2 . \quad (91)$$

Consider the terms $S_1[w]$, $S_2[w]$, $S_3[w]$ and $\int d^3 \mathbf{x}' \left[\mathcal{G}_{(1)}^{(1)}(\mathbf{x}') \right]^2$, these are simply numbers and so have no explicit space or mass dependence. Their action is thus to simply renormalise the overall amplitude of the weight function. Therefore, without any loss of generality, we may set the right-hand-side of Eq. (91) to the constant value β . Hence, our optimal weight equation now reads:

$$\mathcal{K}(\mathbf{x}, M) = \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \right]^2 + C_2 \left[\tilde{\mathcal{G}}_{(2,0)}(\mathbf{x}) + \tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}) \frac{w(\mathbf{x}, M)}{b(M)} \right] + C_2^2 \frac{\tilde{\mathcal{G}}_{(2,0)}(\mathbf{x})}{\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x})} \frac{w(\mathbf{x}, M)}{b(M)} = \beta . \quad (92)$$

C. The optimal weights

We are now in a position to derive the optimal weights. Consider Eq. (92), we see that there are terms which are functions of position alone and terms that are a mixture of both mass and position. Our first insight is that the weight equation must still be valid if we treat it simply as a function of mass for a constant position in the survey \mathbf{x}_0 . On taking $\mathbf{x} = \mathbf{x}_0$ and rearranging Eq. (92) we find for any arbitrary point the weights are given by:

$$w(\mathbf{x}_0, M) = b(M) \tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}_0) \left[\frac{\beta - \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}_0) \right]^2 - C_2 \tilde{\mathcal{G}}_{(2,0)}(\mathbf{x}_0)}{C_2 \left[\tilde{\mathcal{G}}_{(1,1)}(\mathbf{x}_0) \right]^2 + C_2^2 \tilde{\mathcal{G}}_{(2,0)}(\mathbf{x}_0)} \right] . \quad (93)$$

Thus, the mass dependence of the weights is simply proportional to the mass dependence of the bias of the clusters. Without any loss of generality we may now take the general weight function to be separable in position and mass,

$$w(\mathbf{x}, M) = b(M) \tilde{w}(\mathbf{x}) . \quad (94)$$

On back substituting this separable relation for the weights into Eq. (26) and with Eq. (78) we see that:

$$\tilde{\mathcal{G}}_{(i,j)}(\mathbf{x}) = \tilde{w}^i(\mathbf{x}) \int dM \bar{n}(M) b^{i+j}(M) \Theta(\mathbf{x}|M) . \quad (95)$$

On substituting Eqs (94) and (95) into Eq. (92) we arrive at the relation for the spatial dependence of the weight function:

$$\tilde{w}^2(\mathbf{x}) \left\{ \left[\int dM \bar{n}(M) b^2(M) \Theta(\mathbf{x}|M) \right]^2 + 2C_2 \int dM \bar{n}(M) b^2(M) \Theta(\mathbf{x}|M) + C_2^2 \right\} = \beta \quad (96)$$

On rearranging the above expression to isolate the space dependent weight and square-rooting the expression we find:

$$\tilde{w}(\mathbf{x}) = \frac{\beta^{1/2}}{\left[\int dM \bar{n}(M) b^2(M) \Theta(\mathbf{x}|M) + C_2^2 \right]} . \quad (97)$$

On putting together Eqs (94) and (97), and on back-substituting for the constants C_1 and C_2 , and dropping the arbitrary normalization constant β , we arrive at our expression for the optimal weights for achieving maximal \mathcal{S}/\mathcal{N} on a given band-power estimate of the cluster power spectrum:

$$w(\mathbf{x}, M) = \frac{b(M)}{(1 + \alpha) + \int dM \bar{n}(M) b^2(M) \Theta(\mathbf{x}|M) \bar{P}_i} . \quad (98)$$

In a final step, we generalize the above relation to the case where both $b(M)$ and $\bar{n}(M)$ are time dependent quantities. This may be achieved by simply making them a function of conformal time η or equivalently χ . If one propagates these transformations through the entire derivation, then one finds that the weights may be written:

$$w_{\text{OPT}}(\mathbf{x}(\chi, \boldsymbol{\Omega}), M) = \frac{b(M, \chi)}{(1 + \alpha) + \int dM \bar{n}(M, \chi) b^2(M, \chi) \Theta(\mathbf{x}|M) \bar{P}_i} . \quad (99)$$

As for the case of the original FKP weights, we see that the answer is somewhat circular, in that in order to estimate the cluster power spectrum optimally, we already need to have some reasonably good estimate of the underlying matter power spectrum. Indeed, in order to fully implement our scheme we are also required to have knowledge of the functions $b(M)$ and $\bar{n}(M)$. These two functions are theoretically very well known for dark matter haloes. They may also be measured directly from the data – albeit with noise. One also should have very good understanding of the selection function $\Theta(\mathbf{x}|M)$. The parameter α is determined directly from the density of the random cluster sample.

We note that the above scheme is virtually identical to the result found by PVP for the case of luminosity dependent galaxy bias. However, they provided no analytic proof for their result, but simply proposed a conjecture for the general weight solution and showed that it satisfied their weight equation. Our derivation, on the other hand, is more elegant and easily verified. Finally, as was pointed out earlier, once the optimal weights are determined, our choice of field \mathcal{F}_c and theirs, are virtually equivalent. Their inverse weighting of the cluster field by the bias, therefore appears to be an unnecessary step.

VI. CASE STUDIES

We shall now compare how the \mathcal{S}/\mathcal{N} ratios vary as a function of minimum cluster mass, for both volume- and flux-limited limited samples of clusters. We shall compare the optimal weighting scheme derived in the previous section with the standard FKP weighting and also the mass weighting scheme advocated by Seljak et al. [1], Hamaus et al. [31] and Cai et al. [32]. We shall denote the standard FKP weighting as,

$$w_{\text{FKP}}(\mathbf{x}(\chi, \boldsymbol{\Omega}), M) = w_{\text{FKP}}(\chi) = \frac{1}{(1 + \alpha) + \bar{n}_h(\chi)P_i} . \quad (100)$$

We shall denote the mass weighting of the clusters [1, 31, 32] combined with FKP's space weighting as,

$$w_{\text{M+FKP}}(\mathbf{x}(\chi, \boldsymbol{\Omega}), M) = w_{\text{M+FKP}}(\chi, M) = Mw_{\text{FKP}}(\chi) . \quad (101)$$

Before proceeding further, it will be very useful to introduce the following quantities:

$$\bar{n}_h(\chi) \equiv \int_{M_{\text{Lim}}(\chi)}^{\infty} dM \bar{n}(M, \chi) ; \quad (102)$$

$$\bar{\rho}_h(\chi) \equiv \int_{M_{\text{Lim}}(\chi)}^{\infty} dM \bar{n}(M, \chi) M ; \quad (103)$$

$$\overline{b_N^i}(\chi) \equiv \frac{1}{\bar{n}_h(\chi)} \int_{M_{\text{Lim}}(\chi)}^{\infty} dM \bar{n}(M, \chi) b^i(M, \chi) ; \quad (104)$$

$$\overline{b_M^i}(\chi) \equiv \frac{1}{\bar{\rho}_h(\chi)} \int_{M_{\text{Lim}}(\chi)}^{\infty} dM \bar{n}(M, \chi) M b^i(M, \chi) ; \quad (105)$$

$$\langle M^i(x) \rangle \equiv \frac{1}{\bar{n}_h(\chi)} \int_{M_{\text{Lim}}(\chi)}^{\infty} dM \bar{n}(M, \chi) M^i . \quad (106)$$

A. Volume-limited samples

To begin, we adopt the selection function for a volume-limited sample as described in §II. We also note that for a volume-limited sample, the weight function possesses no spatial dependence and so we are free to take the weights simply as

$$w(\mathbf{x}, M) = w(M) . \quad (107)$$

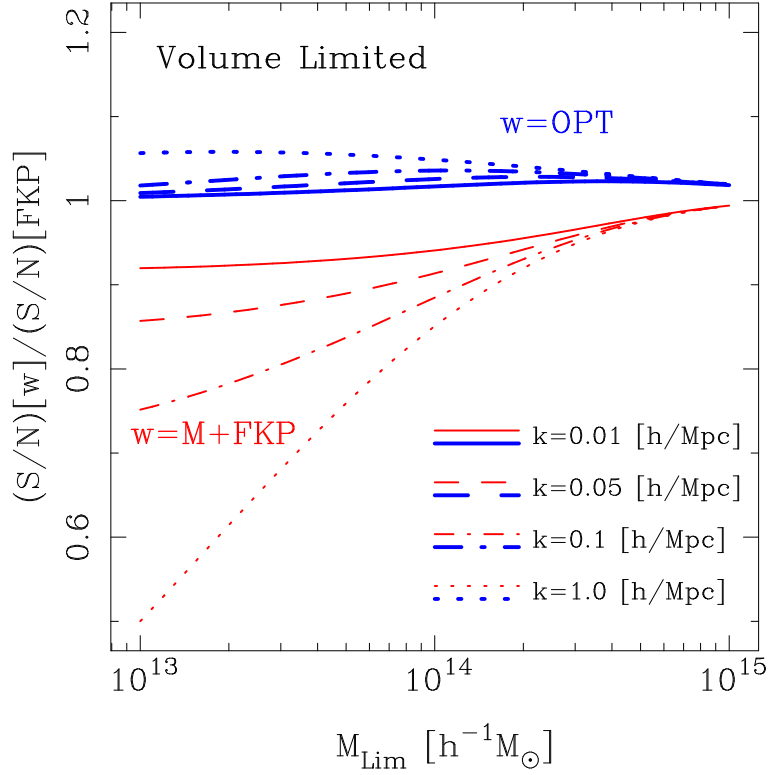


FIG. 1: Signal-to-noise ratios for the optimal and mass weighted cluster power spectrum relative to the FKP signal-to-noise, as a function of the minimum detectable mass in a volume-limited cluster survey. Thick blue lines and thin red lines denote the optimal and mass dependent weighting schemes, respectively. The solid, dash, dot-dash and dotted lines show the results for band-power wavenumbers $k = \{0.01, 0.05, 0.1, 1.0\} h \text{ Mpc}^{-1}$.

With these modifications in hand the expressions for $S_1[w]$, $S_2[w]$, $S_3[w]$ and $S_4[w]$, given in Eqns (74), (75), (76) and (77) can be much simplified and are:

$$S_1^{\text{VL}}[w] = V_\mu \left[\int_{M_{\text{Lim}}}^{\infty} \bar{n}(M)b(M)w(M) \right]^4 ; \quad (108)$$

$$S_2^{\text{VL}}[w] = V_\mu \left[\int_{M_{\text{Lim}}}^{\infty} \bar{n}(M)b(M)w(M) \right]^2 \left[\int_{M_{\text{Lim}}}^{\infty} \bar{n}(M)w^2(M) \right] ; \quad (109)$$

$$S_3^{\text{VL}}[w] = V_\mu \left[\int_{M_{\text{Lim}}}^{\infty} \bar{n}(M)w^2(M) \right]^2 ; \quad (110)$$

$$S_4^{\text{VL}}[w] = V_\mu^2 \left[\int_{M_{\text{Lim}}}^{\infty} \bar{n}(M)b(M)w(M) \right]^4 . \quad (111)$$

Let us now consider how the \mathcal{S}/\mathcal{N} behaves for the three schemes mentioned above, as a function of the minimum detectable mass M_{Lim} . Note that in the volume limited sample we shall assume that $M_{\text{Lim}}(\chi) = M_{\text{Lim}}(\chi_0)$ is a constant throughout the survey.

- **FKP weighting:** $w(M) = 1$. On adopting this weighting scheme we find that Eqns (109), (109), (110) and (111) become Eqn (C1), (C2), (C3) and (C4), and that the \mathcal{S}/\mathcal{N} can be written:

$$\left(\frac{\mathcal{S}}{\mathcal{N}} \right)_{\text{FKP}}^2 = \frac{V_\mu V_i}{2(2\pi)^3} \left[\frac{\bar{b}_N^{-2} \bar{n}_h \bar{P}_i}{(1 + \alpha) + \bar{b}_N^{-2} \bar{n}_h \bar{P}_i} \right]^2 . \quad (112)$$

- **Optimal weighting:** $w(M) = b(M)$. On adopting this weighting scheme we find that Eqns (109), (109), (110)

and (111) become Eqn (C5), (C6), (C7) and (C8), and that the \mathcal{S}/\mathcal{N} can be written:

$$\left(\frac{\mathcal{S}}{\mathcal{N}}\right)_{\text{OPT}}^2 = \frac{V_\mu V_i}{2(2\pi)^3} \left[\frac{\overline{b_N^2 \bar{n}_h \bar{P}_i}}{(1+\alpha) + \overline{b_N^2 \bar{n}_h \bar{P}_i}} \right]^2. \quad (113)$$

- **Mass weighting:** $w(M) = M$. On adopting this weighting scheme we find that Eqns (109), (109), (110) and (111) become Eqn (C9), (C10), (C11) and (C12), and that the \mathcal{S}/\mathcal{N} can be written:

$$\left(\frac{\mathcal{S}}{\mathcal{N}}\right)_{\text{M+FKP}}^2 = \frac{V_\mu V_i}{2(2\pi)^3} \left[\frac{r^2(M) \overline{b_M^2 \bar{n}_h \bar{P}_i}}{(1+\alpha) + r^2(M) \overline{b_M^2 \bar{n}_h \bar{P}_i}} \right]^2, \quad (114)$$

where we defined $r^2(M) \equiv \langle M^2 \rangle / \langle M \rangle^2$.

The \mathcal{S}/\mathcal{N} values obtained from the three weighting schemes may be more easily compared if we take the ratio of the optimal and mass weighting scheme with respect to the original FKP weights. Whereupon,

$$\frac{(\mathcal{S}/\mathcal{N})_{\text{OPT}}}{(\mathcal{S}/\mathcal{N})_{\text{FKP}}} = \frac{1 + (1+\alpha)/\bar{n}_h \bar{P}_i \overline{b_N^2}}{1 + (1+\alpha)/\bar{n}_h \bar{P}_i \overline{b_N^2}}; \quad (115)$$

$$\frac{(\mathcal{S}/\mathcal{N})_{\text{M+FKP}}}{(\mathcal{S}/\mathcal{N})_{\text{FKP}}} = \frac{1 + (1+\alpha)/\bar{n}_h \bar{P}_i \overline{b_N^2}}{1 + (1+\alpha)/\bar{n}_h \bar{P}_i \overline{b_M^2} r^2(M)}. \quad (116)$$

There are two limiting cases of interest:

- $\bar{n}_h \bar{P}_i \gg 1$: in this limit Eqs (115) and (116) become,

$$\frac{(\mathcal{S}/\mathcal{N})_{\text{OPT}}}{(\mathcal{S}/\mathcal{N})_{\text{FKP}}} \approx 1 + \frac{(1+\alpha)}{\bar{n}_h \bar{P}_i} \left[\frac{1}{\overline{b_N^2}} - \frac{1}{\overline{b_N^2}} \right] \geq 1; \quad (117)$$

$$\frac{(\mathcal{S}/\mathcal{N})_{\text{M+FKP}}}{(\mathcal{S}/\mathcal{N})_{\text{FKP}}} \approx 1 + \frac{(1+\alpha)}{(\bar{n}_h \bar{P}_i)} \left[\frac{1}{\overline{b_N^2}} - \frac{1}{\overline{b_M^2} r^2(M)} \right]. \quad (118)$$

- $\bar{n}_h \bar{P}_i \ll 1$: in this limit Eqs (115) and (116) become,

$$\frac{(\mathcal{S}/\mathcal{N})_{\text{OPT}}}{(\mathcal{S}/\mathcal{N})_{\text{FKP}}} \approx \frac{\overline{b_N^2}}{\overline{b_N^2}} \geq 1; \quad (119)$$

$$\frac{(\mathcal{S}/\mathcal{N})_{\text{M+FKP}}}{(\mathcal{S}/\mathcal{N})_{\text{FKP}}} \approx \frac{1}{r^2(M)} \left(\frac{\overline{b_M^2}}{\overline{b_N^2}} \right)^2. \quad (120)$$

The inequalities given by Eqs (117) and (119) both follow from the fact that $\overline{b_N^2} \leq \overline{b_N^2}$ (for a proof of this relation see Appendix D). In order to determine whether Eqs (118) and (120) are greater or less than unity, one must examine the product of the quantities $\overline{b_N^2}/\overline{b_M^2} \leq 1$ and $\langle M^2 \rangle / \langle M \rangle^2 \geq 1$. Unfortunately, this is not so easy to determine and we therefore turn to numerical evaluation of the expressions.

In Fig. 1 we present the evolution of the \mathcal{S}/\mathcal{N} for the optimal and mass weighting schemes, relative to the \mathcal{S}/\mathcal{N} obtained from the FKP weighting, as a function of M_{Lim} . The blue and red lines denote the optimal and mass weighting schemes, respectively and results are presented for several k -band powers. Here we have used the LCDM model as a particular example and have therefore adopted the bias and dark matter halo mass function formula of Sheth and Tormen [10] to evaluate Eqs (115) and (116). Several important points may be noted: the optimal weighting scheme does indeed maximize the \mathcal{S}/\mathcal{N} ; the mass weighting scheme is inferior to the optimal and FKP weighting schemes; the overall gains of the optimal weighting scheme relative to the FKP scheme appear modest $\sim 10\%$.

B. Flux-limited samples

To begin, we adopt the selection function for a flux-limited sample as described in §II. As described in §V the weight function in this case can be written:

$$w(\mathbf{x}(\chi, \mathbf{\Omega}), M) = w_\chi(\chi) w_M(M, \chi). \quad (121)$$

With this information in hand, our expressions for $S_1[w]$, $S_2[w]$, $S_3[w]$ and $S_4[w]$, given in Eqns (74), (75), (76) and (77) can be written more simply as:

$$S_1^{\text{FL}}[w] \equiv \Omega_\mu \int_0^\infty d\chi \chi^2 w_\chi^4(\chi) \left[\int_{M_{\text{Lim}}(\chi)}^\infty dM_1 \bar{n}(M_1) b(M_1) w_M(M_1) \right]^4 ; \quad (122)$$

$$S_2^{\text{FL}}[w] \equiv \Omega_\mu \int_0^\infty d\chi \chi^2 w_\chi^4(\chi) \left[\int_{M_{\text{Lim}}(\chi)}^\infty dM_1 \bar{n}(M_1) b(M_1) w_M(M_1) \right]^2 \left[\int_{M_{\text{Lim}}(\chi)}^\infty dM_2 \bar{n}(M_2) w_M^2(M_2) \right] ; \quad (123)$$

$$S_3^{\text{FL}}[w] \equiv \Omega_\mu \int_0^\infty d\chi \chi^2 w_\chi^4(\chi) \left[\int_{M_{\text{Lim}}(\chi)}^\infty dM_1 \bar{n}(M_1) w_M^2(M_1) \right]^2 ; \quad (124)$$

$$S_4^{\text{FL}}[w] \equiv \left\{ \Omega_\mu \int_0^\infty d\chi \chi^2 w_\chi^2(\chi) \left[\int_{M_{\text{Lim}}(\chi)}^\infty dM_1 \bar{n}(M_1) b(M_1) w_M(M_1) \right]^2 \right\}^2 , \quad (125)$$

where $M_{\text{Lim}}(\chi)$ is the minimum detectable mass at radial distance χ from the observer and Ω_μ is the survey area. Note that for compactness, we have suppressed the dependence of $\bar{n}(M)$, $b(M)$ and $w_M(M)$ on χ .

We again consider the three weighting schemes from the previous section, for which the \mathcal{S}/\mathcal{N} values can be written:

- **FKP weighting:** On adopting this weighting scheme we find that Eqns (122), (123), (124) and (125) become Eqn (C13), (C14), (C15) and (C16), and that the \mathcal{S}/\mathcal{N} can be written:

$$\left(\frac{\mathcal{S}}{\mathcal{N}} \right)_{\text{FKP}}^2 = \frac{\Omega_\mu V_i}{2(2\pi)^3} \frac{\left\{ \int_0^\infty d\chi \chi^2 \left[\frac{\bar{b}_N(\chi) \bar{n}_h(\chi) \bar{P}_i}{(1+\alpha) + \bar{n}_h(\chi) \bar{P}_i} \right]^2 \right\}^2}{\int_0^\infty d\chi \chi^2 \bar{n}_h^2(\chi) \bar{P}_i^2 \frac{[(1+\alpha) + \bar{b}_N^2(\chi) \bar{n}_h(\chi) \bar{P}_i]^2}{[(1+\alpha) + \bar{n}_h(\chi) \bar{P}_i]^4}} . \quad (126)$$

- **Optimal weighting:** On adopting this weighting scheme we find that Eqns (122), (123), (124) and (125) become Eqn (C17), (C18), (C19) and (C20), and that the \mathcal{S}/\mathcal{N} can be written:

$$\left(\frac{\mathcal{S}}{\mathcal{N}} \right)_{\text{OPT}}^2 = \frac{\Omega_\mu V_i}{2(2\pi)^3} \int_0^\infty d\chi \chi^2 \left[\frac{\bar{n}_h(\chi) \bar{b}_N^2(\chi) \bar{P}_i}{(1+\alpha) + \bar{n}_h(\chi) \bar{b}_N^2(\chi) \bar{P}_i} \right]^2 . \quad (127)$$

- **Mass+FKP weighting:** On adopting this weighting scheme we find that Eqns (122), (123), (124) and (125) become Eqn (C21), (C22), (C23) and (C24), and that the \mathcal{S}/\mathcal{N} can be written:

$$\left(\frac{\mathcal{S}}{\mathcal{N}} \right)_{\text{M+FKP}}^2 = \frac{\Omega_\mu V_i}{2(2\pi)^3} \frac{\left\{ \int_0^\infty d\chi \chi^2 \left[\frac{\langle M(\chi) \rangle \bar{b}_M(\chi) \bar{n}_h(\chi) \bar{P}_i}{(1+\alpha) + \bar{n}_h(\chi) \bar{P}_i} \right]^2 \right\}^2}{\int_0^\infty d\chi \chi^2 \bar{n}_h^2(\chi) \bar{P}_i^2 \frac{[(1+\alpha) \langle M^2(\chi) \rangle + \langle M(\chi) \rangle^2 \bar{b}_M^2(\chi) \bar{n}_h(\chi) \bar{P}_i]^2}{[(1+\alpha) + \bar{n}_h(\chi) \bar{P}_i]^4}} . \quad (128)$$

Further analytic developments are non-trivial, and so we numerically evaluate them for the particular case of LCDM. In order to accomplish this we require knowledge of the minimum detectable mass as a function of χ . As mentioned in §II, $M_{\text{Lim}}(\chi)$, in general, is a complicated function of the survey flux-limit and the cluster identification algorithm. For simplicity, we shall assume that this can be written as:

$$M_{\text{Lim}}(\chi) = M_{\text{Lim}}(0) e^{\gamma z(\chi)} . \quad (129)$$

If we adopt the value $\gamma = 1$, the above functional form roughly matches the cluster selection as a function of redshift, which one finds for weak lensing detected cluster surveys [40]. We evaluate the above \mathcal{S}/\mathcal{N} ratios as a function of $M_{\text{Lim}}(0)$, with $M_{\text{Lim}}(0)$ varying in the range $[10^{13}, 10^{15}] h^{-1} M_\odot$.

In Fig. 2 we present the evolution of the \mathcal{S}/\mathcal{N} relative to the \mathcal{S}/\mathcal{N} obtained from the usual FKP weighting as a function of the minimum detectable mass for a flux-limited cluster survey. The blue and red lines denote the optimal and mass weighting schemes, respectively and results are presented for several k -band powers. Again, we have used the LCDM model as an example and have adopted the bias and dark matter halo mass functions from Sheth and

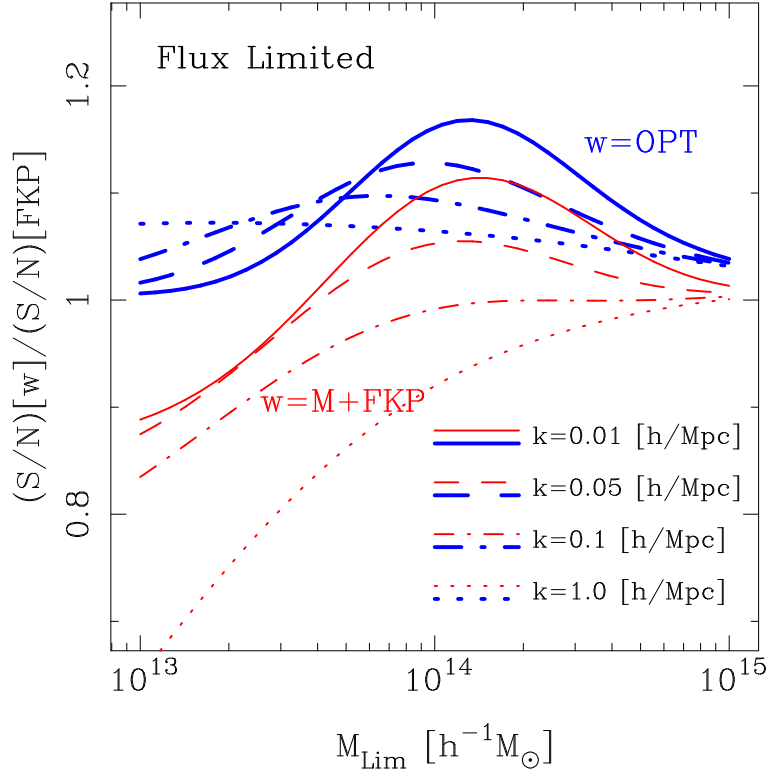


FIG. 2: Same as Fig.1 except this time for a flux-limited survey.

Tormen [10] to evaluate Eqns (126), (127) and (128). Several important points may be noted: again, the optimal weighting scheme always maximizes the signal-to-noise ratio; the mass weighting scheme is inferior to the optimal scheme for all scales; the mass weighting scheme can be more optimal than FKP for massive clusters on large-scales, otherwise it is always less efficient than FKP; the overall S/N gains for the optimal weighting are $\sim 17\%$ improved over the FKP approach.

C. The Fisher matrix

As a final corollary to this section, we explore the cosmological information content of an optimally weighted cluster power spectrum analysis. Following Tegmark [3], the Fisher information matrix for a power spectrum analysis can be defined as:

$$\begin{aligned}
 \mathcal{F}_{\alpha\beta} &= \sum_{i,j} \frac{\partial \bar{P}_i}{\partial \alpha} C_{ij}^{-1} \frac{\partial \bar{P}_j}{\partial \beta} \\
 &= \sum_{i,j} \frac{\partial \log \bar{P}_i}{\partial \alpha} \bar{P}_i C_{ij}^{-1} \bar{P}_j \frac{\partial \log \bar{P}_j}{\partial \beta}, \quad (130)
 \end{aligned}$$

where $\partial/\partial\alpha \equiv \partial/\partial\theta_\alpha$ denote partial derivatives with respect to the cosmological parameters θ_α , and where here we have neglected the information content in the covariance matrix [41]. On inserting our earlier expression for the covariance matrix, as given by Eq. (72), into the above expression, then the Fisher matrix becomes:

$$\begin{aligned}
 \mathcal{F}_{\alpha\beta} &= \sum_{i,j} \frac{\partial \log \bar{P}_i}{\partial \alpha} \bar{P}_i \frac{\delta_{ij}^K}{\sigma_P^2(k_i)} \bar{P}_j \frac{\partial \log \bar{P}_j}{\partial \beta} \\
 &= \sum_i \frac{\partial \log \bar{P}_i}{\partial \alpha} \frac{\partial \log \bar{P}_i}{\partial \beta} \left(\frac{S}{N} \right)_{\text{OPT}}^2 (k_i). \quad (131)
 \end{aligned}$$

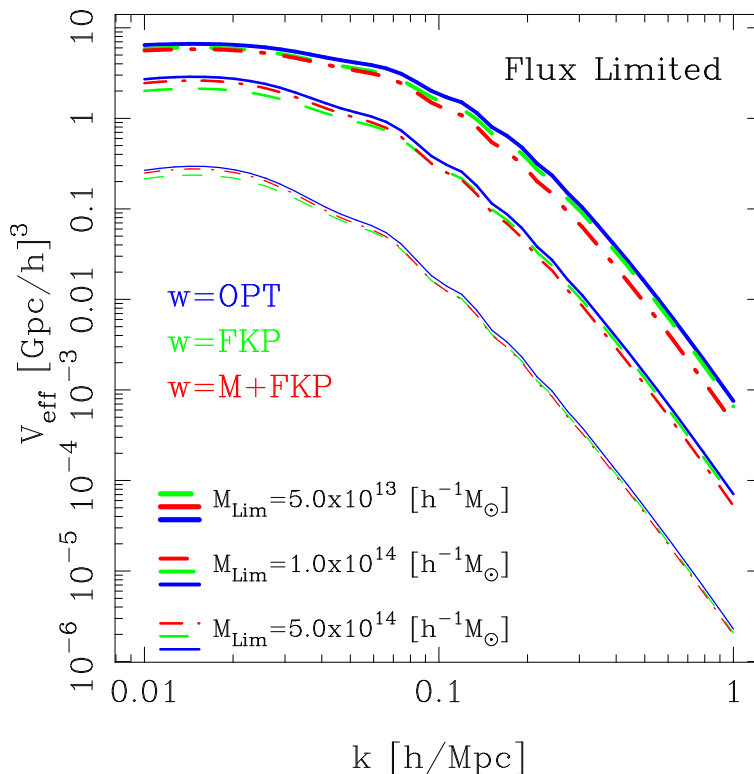


FIG. 3: Effective survey volume as a function of the power spectrum wavenumber probed. Lines of decreasing thickness correspond to a minimum detectable mass $M_{\text{Lim}}(0)$, of $5.0 \times 10^{13} h^{-1} M_{\odot}$, $1.0 \times 10^{14} h^{-1} M_{\odot}$, $5.0 \times 10^{14} h^{-1} M_{\odot}$, respectively. The solid, dashed and dot-dashed lines correspond to V_{eff} for optimal, FKP and M+FKP weighting, respectively.

On inserting our expression for the \mathcal{S}/\mathcal{N} for the optimal weights given by Eq. (127), the Fisher matrix, in the continuum limit of Fourier modes, can be written:

$$\mathcal{F}_{\alpha\beta} = \frac{1}{2} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{\partial \log P(k)}{\partial \alpha} \frac{\partial \log P(k)}{\partial \beta} V_{\text{eff}}(k), \quad (132)$$

where the effective survey volume has the new form:

$$V_{\text{eff}}(k) = \int_0^{\infty} d\chi \chi^2 \left[\frac{\bar{n}_h(\chi) \overline{b_N^2}(\chi) P(k)}{(1 + \alpha) + \bar{n}_h(\chi) \overline{b_N^2}(\chi) P(k)} \right]^2. \quad (133)$$

In Fig. 3 we show the effective survey volume as a function of wavenumber, for a flux-limited survey of similar type to that described in §VIB. Here we consider the three cases where the minimum detectable mass normalization parameter from, c.f. Eq. (129), has the values $5.0 \times 10^{13} h^{-1} M_{\odot}$, $1.0 \times 10^{14} h^{-1} M_{\odot}$, $5.0 \times 10^{14} h^{-1} M_{\odot}$, respectively. In all cases the optimal weighting increases V_{eff} . Thus we conclude that the cosmological information extractable from an optimal weighted cluster survey will exceed that from the sub-optimal strategies, such as FKP or M+FKP weighting. Eq. (132) may thus be used as the starting point for exploring the cosmological information content of optimally weighted cluster power spectra [for a review of the Fisher matrix approach, see 42].

VII. CONCLUSIONS

In this paper we have developed the optimal weighting scheme for estimating the cluster power spectrum, and used it to make inferences about the information in the matter power spectrum. Our derivation generalizes the original approach of FKP for galaxies, and is analogous to the derivation by PVP for examining the impact of luminosity dependent galaxy bias on the weights. Our derivation provides for the first time an analytic proof for the optimal weight equation, and it also corrects several errors that were found in these earlier works.

In §II we described the generic properties of a cluster survey, taking into account finite survey geometry, arbitrary weighting of position and mass, and a flux limit; we also introduced our statistical treatment for describing the cluster density field using delta function expansions.

In §III we presented estimators for recovering the matter clustering from the computation of the two-point correlation function and the power spectrum of the cluster field. We demonstrated that in order to extract information from the matter clustering one must deconvolve the survey window function. For large homogeneous survey volumes, provided an appropriate choice for the normalization of the cluster field is taken, this estimate was shown to be an unbiased estimator of the dark matter power spectrum, modulo a shot-noise correction.

In §IV we explored the statistical fluctuations in the power spectrum of clusters. We derived general expressions for the covariance of the cluster sample, including all non-Gaussian terms arising from the nonlinear evolution of matter fluctuations and discreteness effects. This generalized the result of [34]. We then proved the necessary conditions for the covariance matrix to be diagonal.

In §V we have provided an analytic derivation of the optimal weights for a cluster power spectrum analysis. We show in general terms that the optimal weights are separable functions of mass and space.

In §VI we present a comparison of the optimal weighting scheme with the original FKP weighting scheme and with the mass weighting scheme advocated by Seljak et al. [1], Hamaus et al. [31] and Cai et al. [32]. For the case of both volume- and flux-limited cluster surveys the optimal weighting scheme outperforms both alternate weighting schemes. On large scales the differences are modest $\sim 10\%$. However, on intermediate scales the mass-dependent weighting causes a significant loss in signal-to-noise. Whilst the mass-dependent weighting may be useful for reconstructing the matter field from a cluster sample, we recommend that it should not be used to extract cosmological information from cluster power spectrum analysis. We also presented a new expression for the Fisher information matrix, for an optimally-weighted cluster power spectrum measurement. We showed that the cosmological constraining power of our optimally-weighted power spectra is superior to an FKP analysis of clusters.

In this paper we have derived the optimal weights for measuring the cluster power spectrum under certain conditions, if these conditions are relaxed then the weights are no longer optimal. It will be interesting in future work to explore whether a more general weighting scheme can be derived for the more realistic situations where a non-diagonal covariance matrix is considered. We also expect that the weights that we have derived for the power spectrum should also be used to obtain optimal measurements of the cluster correlation function. However, we have not yet demonstrated this explicitly. Another question for future consideration is: why the optimal density field reconstruction weights presented by Hamaus et al. [31] and also by Cai et al. [32] do not appear to optimize the cluster power spectrum estimation?

Acknowledgments

We thank Ravi Sheth and Simon White for useful discussions. RES acknowledges support from ERC Advanced grant 246797 GALFORMOD. LM thanks MPA for its kind hospitality while this work was being performed. LM was partly supported by the DFG through the grant MA 4967/1-2.

-
- [1] U. Seljak, N. Hamaus, and V. Desjacques, *Physical Review Letters* **103**, 091303 (2009), 0904.2963.
 - [2] H. A. Feldman, N. Kaiser, and J. A. Peacock, *ApJ* **426**, 23 (1994), arXiv:astro-ph/9304022.
 - [3] M. Tegmark, *Physical Review Letters* **79**, 3806 (1997), arXiv:astro-ph/9706198.
 - [4] S. W. Allen, A. E. Evrard, and A. B. Mantz, *ArXiv e-prints* (2011), 1103.4829.
 - [5] M. G. Hauser and P. J. E. Peebles, *ApJ* **185**, 757 (1973).
 - [6] N. A. Bahcall and R. M. Soneira, *ApJ* **270**, 20 (1983).
 - [7] N. Kaiser, *ApJL* **284**, L9 (1984).
 - [8] S. Cole and N. Kaiser, *MNRAS* **237**, 1127 (1989).
 - [9] H. J. Mo and S. D. M. White, *MNRAS* **282**, 347 (1996), arXiv:astro-ph/9512127.
 - [10] R. K. Sheth and G. Tormen, *MNRAS* **308**, 119 (1999), arXiv:astro-ph/9901122.
 - [11] H. J. Mo, Y. P. Jing, and S. D. M. White, *MNRAS* **282**, 1096 (1996), astro-ph/9602052.
 - [12] S. Majumdar and J. J. Mohr, *ApJ* **613**, 41 (2004), astro-ph/0305341.
 - [13] M. Lima and W. Hu, *PRD* **70**, 043504 (2004), arXiv:astro-ph/0401559.
 - [14] M. Lima and W. Hu, *PRD* **72**, 043006 (2005), arXiv:astro-ph/0503363.
 - [15] M. Oguri, *Physical Review Letters* **102**, 211301 (2009), 0905.0920.
 - [16] C. Cunha, D. Huterer, and O. Doré, *PRD* **82**, 023004 (2010), 1003.2416.
 - [17] R. E. Smith and L. Mariani, *MNRAS* **418**, 729 (2011), 1106.1665.
 - [18] M. Oguri and M. Takada, *PRD* **83**, 023008 (2011), 1010.0744.

- [19] A. K. Romer, C. A. Collins, H. Böhringer, R. G. Cruddace, H. Ebeling, H. T. MacGillivray, and W. Voges, *Nature* **372**, 75 (1994).
- [20] C. A. Collins, L. Guzzo, H. Böhringer, P. Schuecker, G. Chincarini, R. Cruddace, S. De Grandi, H. T. MacGillivray, D. M. Neumann, S. Schindler, et al., *MNRAS* **319**, 939 (2000), astro-ph/0008245.
- [21] M. G. Abadi, D. G. Lambas, and H. Muriel, *ApJ* **507**, 526 (1998), astro-ph/9806274.
- [22] N. A. Bahcall, *Annual Reviews of Astronomy & Astrophys.* **26**, 631 (1988).
- [23] G. B. Dalton, G. Efstathiou, S. J. Maddox, and W. J. Sutherland, *ApJL* **390**, L1 (1992).
- [24] C. J. Miller, R. C. Nichol, and D. J. Batuski, *ApJ* **555**, 68 (2001), astro-ph/0103018.
- [25] C. J. Miller and D. J. Batuski, *ApJ* **551**, 635 (2001), astro-ph/0002295.
- [26] B. P. Koester, T. A. McKay, J. Annis, R. H. Wechsler, A. Evrard, L. Bleem, M. Becker, D. Johnston, E. Sheldon, R. Nichol, et al., *ApJ* **660**, 239 (2007), astro-ph/0701265.
- [27] N. A. Bahcall, F. Dong, L. Hao, P. Bode, J. Annis, J. E. Gunn, and D. P. Schneider, *ApJ* **599**, 814 (2003), astro-ph/0307102.
- [28] J. Estrada, E. Sefusatti, and J. A. Frieman, *ApJ* **692**, 265 (2009), 0801.3485.
- [29] G. Hütsi, *MNRAS* **401**, 2477 (2010), 0910.0492.
- [30] A. Pillepich, C. Porciani, and T. H. Reiprich, *MNRAS* **422**, 44 (2012), 1111.6587.
- [31] N. Hamaus, U. Seljak, V. Desjacques, R. E. Smith, and T. Baldauf, *PRD* **82**, 043515 (2010), 1004.5377.
- [32] Y.-C. Cai, G. Bernstein, and R. K. Sheth, *MNRAS* **412**, 995 (2011), 1007.3500.
- [33] W. J. Percival, L. Verde, and J. A. Peacock, *MNRAS* **347**, 645 (2004), arXiv:astro-ph/0306511.
- [34] R. E. Smith, *MNRAS* pp. 1337–+ (2009), 0810.1960.
- [35] R. K. Sheth and G. Lemson, *MNRAS* **304**, 767 (1999), arXiv:astro-ph/9808138.
- [36] R. E. Smith and P. I. R. Watts, *MNRAS* **360**, 203 (2005), arXiv:astro-ph/0412441.
- [37] J. N. Fry and E. Gaztanaga, *ApJ* **413**, 447 (1993), arXiv:astro-ph/9302009.
- [38] R. E. Smith, R. Scoccimarro, and R. K. Sheth, *PRD* **75**, 063512 (2007), arXiv:astro-ph/0609547.
- [39] H. J. Mo, Y. P. Jing, and S. D. M. White, *MNRAS* **284**, 189 (1997), arXiv:astro-ph/9603039.
- [40] L. Mariani and G. M. Bernstein, *PRD* **73**, 123525 (2006), arXiv:astro-ph/0605746.
- [41] M. Tegmark, A. N. Taylor, and A. F. Heavens, *ApJ* **480**, 22 (1997), arXiv:astro-ph/9603021.
- [42] A. Heavens, *ArXiv e-prints* (2009), 0906.0664.
- [43] www.darkenergysurvey.org

Appendix A: Statistical fluctuations in the \mathcal{F}_c power spectrum

1. Derivation of Eq. (50)

In this section we give the necessary steps to derive Eq. (50). The covariance matrix of the power in the field \mathcal{F}_c can be written:

$$\text{Cov}[|\mathcal{F}_c(\mathbf{k}_1)|^2, |\mathcal{F}_c(\mathbf{k}_2)|^2] = \int d^3\mathbf{k}_3 d^3\mathbf{k}_4 \delta^D(\mathbf{k}_1 + \mathbf{k}_3) \delta^D(\mathbf{k}_2 + \mathbf{k}_4) \left[\langle \mathcal{F}_c(\mathbf{k}_1) \dots \mathcal{F}_c(\mathbf{k}_4) \rangle - \langle \mathcal{F}_c(\mathbf{k}_1) \mathcal{F}_c(\mathbf{k}_3) \rangle \langle \mathcal{F}_c(\mathbf{k}_2) \mathcal{F}_c(\mathbf{k}_4) \rangle \right] \quad (\text{A1})$$

In order to proceed further we see that we must compute the four-point correlation function of the Fourier modes of the field $\mathcal{F}_c(\mathbf{k})$. On Fourier transforming the fields on the right-hand-side of the above expression we see that this is equivalent to specifying the four-point correlation function of the field \mathcal{F}_c :

$$\langle \mathcal{F}_c(\mathbf{k}_1) \dots \mathcal{F}_c(\mathbf{k}_4) \rangle = \int d^3\mathbf{x}_1 \dots d^3\mathbf{x}_4 \langle \mathcal{F}_c(\mathbf{x}_1) \dots \mathcal{F}_c(\mathbf{x}_4) \rangle e^{i\mathbf{k}_1 \cdot \mathbf{x}_1 + \dots + \mathbf{k}_4 \cdot \mathbf{x}_4} . \quad (\text{A2})$$

If we now substitute Eq. (9) into the expression for the four-point correlation function then we find:

$$\langle \mathcal{F}_c(\mathbf{x}_1) \dots \mathcal{F}_c(\mathbf{x}_4) \rangle = \frac{1}{A^2} \int \prod_{i=1}^4 \{dM_i w(\mathbf{x}_i, M_i)\} \left\langle [n_c(\mathbf{x}_1, M_1) - \alpha n_s(\mathbf{x}_1, M_1)] \dots [n_c(\mathbf{x}_4, M_4) - \alpha n_s(\mathbf{x}_4, M_4)] \right\rangle \quad (\text{A3})$$

Expanding the term in angled braces on the right-hand-side gives,

$$\begin{aligned} \langle [n_{c,1} - \alpha n_{s,1}] \dots [n_{c,4} - \alpha n_{s,4}] \rangle &= \langle n_{c,1} \dots n_{c,4} \rangle - \alpha \left[\langle n_{c,1} n_{c,2} n_{c,3} n_{s,4} \rangle + 3\text{cyc} \right] + \alpha^2 \left[\langle n_{c,1} n_{c,2} n_{s,3} n_{s,4} \rangle + 5\text{perm} \right] \\ &\quad - \alpha^3 \left[\langle n_{c,1} n_{s,2} n_{s,3} n_{s,4} \rangle + 3\text{cyc} \right] + \alpha^4 \langle n_{s,1} \dots n_{s,4} \rangle , \end{aligned} \quad (\text{A4})$$

where we have introduced the following short-hand notation, $n_{c,1} \equiv n_c(\mathbf{x}_1, M_1)$ and $n_{s,1} \equiv n_s(\mathbf{x}_1, M_1)$. Focusing on the first term in curly brackets on the right-hand-side, and if we insert our delta function expansion for the number

density field we find:

$$\begin{aligned}
\mathcal{L}^{\text{cccc}} &\equiv \left\langle n_{\text{c}}(\mathbf{x}_1, \tilde{M}_1) \dots n_{\text{c}}(\mathbf{x}_4, \tilde{M}_4) \right\rangle \\
&= \left\langle \sum_{i,j,k,l=1}^N \delta^{\text{D}}(\mathbf{r}_1 - \mathbf{x}_i) \dots \delta^{\text{D}}(\mathbf{r}_4 - \mathbf{x}_l) \delta^{\text{D}}(\tilde{M}_1 - M_i) \dots \delta^{\text{D}}(\tilde{M}_4 - M_l) \right\rangle \\
&= \int \prod_{p=1}^N \{d^3 \mathbf{x}_p dM_p\} p(\mathbf{x}_1, \dots, \mathbf{x}_N, M_1, \dots, M_N) (\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5)
\end{aligned} \tag{A5}$$

where in the above we have defined the following terms:

$$\begin{aligned}
\beta_1 &\equiv \sum_{i \neq j \neq k \neq l} \delta^{\text{D}}(\mathbf{r}_1 - \mathbf{x}_i) \dots \delta^{\text{D}}(\mathbf{r}_4 - \mathbf{x}_l) \delta^{\text{D}}(\tilde{M}_1 - M_i) \dots \delta^{\text{D}}(\tilde{M}_4 - M_l) \\
\beta_2 &\equiv \sum_{i \neq j \neq k=l} \delta^{\text{D}}(\mathbf{r}_1 - \mathbf{x}_i) \delta^{\text{D}}(\mathbf{r}_2 - \mathbf{x}_j) \delta^{\text{D}}(\tilde{M}_1 - M_i) \delta^{\text{D}}(\tilde{M}_2 - M_j) \prod_{p=3}^4 \left\{ \delta^{\text{D}}(\mathbf{r}_p - \mathbf{x}_k) \delta^{\text{D}}(\tilde{M}_p - M_k) \right\} + 5\text{perms} \\
\beta_3 &\equiv \sum_{i=j \neq k=l} \prod_{p=1}^2 \left\{ \delta^{\text{D}}(\mathbf{r}_p - \mathbf{x}_i) \delta^{\text{D}}(\tilde{M}_p - M_i) \right\} \prod_{q=3}^4 \left\{ \delta^{\text{D}}(\mathbf{r}_q - \mathbf{x}_j) \delta^{\text{D}}(\tilde{M}_q - M_j) \right\} + 2\text{perms} \\
\beta_4 &\equiv \sum_{i \neq j=k=l} \delta^{\text{D}}(\mathbf{r}_1 - \mathbf{x}_i) \delta^{\text{D}}(\tilde{M}_1 - M_i) \prod_{p=2}^4 \left\{ \delta^{\text{D}}(\mathbf{r}_p - \mathbf{x}_j) \delta^{\text{D}}(\tilde{M}_p - M_j) \right\} + 3\text{perms} \\
\beta_5 &\equiv \sum_{i=j=k=l} \prod_{p=1}^4 \left\{ \delta^{\text{D}}(\mathbf{r}_p - \mathbf{x}_i) \delta^{\text{D}}(\tilde{M}_p - M_i) \right\}
\end{aligned} \tag{A6}$$

Let us break up Eq. (A5) and define $\mathcal{L}_i^{\text{cccc}}$ to be the contribution coming from the term β_i .

• $\mathcal{L}_1^{\text{cccc}}$: On performing the integrations over the delta functions we have:

$$\begin{aligned}
\mathcal{L}_1^{\text{cccc}} &= \sum_{i \neq j \neq k \neq l} p(\mathbf{r}_1, \dots, \mathbf{r}_4, M_1, \dots, M_4) \\
&= \bar{n}(N-1)(N-2)(N-3)p(\mathbf{r}_1, M_1) \dots p(\mathbf{r}_4, M_4) \left[1 + \xi_{12}^{\text{c}} + \xi_{13}^{\text{c}} + \xi_{14}^{\text{c}} + \xi_{23}^{\text{c}} + \xi_{24}^{\text{c}} + \xi_{34}^{\text{c}} \right. \\
&\quad \left. + \zeta_{123}^{\text{c}} + \zeta_{234}^{\text{c}} + \zeta_{341}^{\text{c}} + \zeta_{412}^{\text{c}} + \xi_{12}^{\text{c}} \xi_{34}^{\text{c}} + \xi_{13}^{\text{c}} \xi_{24}^{\text{c}} + \xi_{14}^{\text{c}} \xi_{23}^{\text{c}} + \eta_{1234}^{\text{c}} \right] \\
&\approx \bar{n}_1 \dots \bar{n}_4 \Theta_1 \dots \Theta_4 \left[1 + \xi_{12}^{\text{c}} + \xi_{13}^{\text{c}} + \xi_{14}^{\text{c}} + \xi_{23}^{\text{c}} + \xi_{24}^{\text{c}} + \xi_{34}^{\text{c}} + \zeta_{123}^{\text{c}} + \zeta_{234}^{\text{c}} + \zeta_{341}^{\text{c}} + \zeta_{412}^{\text{c}} \right. \\
&\quad \left. + \xi_{12}^{\text{c}} \xi_{34}^{\text{c}} + \xi_{13}^{\text{c}} \xi_{24}^{\text{c}} + \xi_{14}^{\text{c}} \xi_{23}^{\text{c}} + \eta_{1234}^{\text{c}} \right],
\end{aligned} \tag{A7}$$

where the last equality holds in the limit that $N \gg 1$. The terms ζ^{c} and η^{c} are the three- and four-point correlation functions of the haloes, respectively, and in the above we have also made use of the short-hand notation:

$$\xi_{ij}^{\text{c}} \equiv \xi^{\text{c}}(\mathbf{r}_i, \mathbf{r}_j, M_i, M_j) \quad ; \quad \zeta_{ijk}^{\text{c}} \equiv \zeta(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k, M_i, M_j, M_k) \quad ; \quad \eta_{ijkl}^{\text{c}} \equiv \eta^{\text{c}}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k, \mathbf{r}_l, M_i, M_j, M_k, M_l) \quad . \tag{A8}$$

• $\mathcal{L}_2^{\text{cccc}}$: On integration over the delta functions and defining the short-hand-notation $\delta_{r,ij}^{\text{D}} \equiv \delta^{\text{D}}(\mathbf{r}_i - \mathbf{r}_j)$ and $\delta_{M,ij}^{\text{D}} \equiv \delta^{\text{D}}(M_i - M_j)$, the second term may be written:

$$\begin{aligned}
\mathcal{L}_2^{\text{cccc}} &\equiv \sum_{i \neq j \neq k \neq l} p(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, M_1, M_2, M_3) \delta_{r,34}^{\text{D}} \delta_{M,34}^{\text{D}} + \sum_{i \neq j=k \neq l} p(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_4, M_1, M_2, M_4) \delta_{r,23}^{\text{D}} \delta_{M,23}^{\text{D}} \\
&\quad + \sum_{i=j \neq k \neq l} p(\mathbf{r}_1, \mathbf{r}_3, \mathbf{r}_4, M_1, M_3, M_4) \delta_{r,12}^{\text{D}} \delta_{M,12}^{\text{D}} + \sum_{i=k \neq j \neq l} p(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_4, M_1, M_2, M_4) \delta_{r,13}^{\text{D}} \delta_{M,13}^{\text{D}} \\
&\quad + \sum_{i=l \neq j \neq k} p(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, M_1, M_2, M_3) \delta_{r,14}^{\text{D}} \delta_{M,14}^{\text{D}} + \sum_{j=l \neq i \neq k} p(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, M_1, M_2, M_3) \delta_{r,24}^{\text{D}} \delta_{M,24}^{\text{D}}
\end{aligned} \tag{A9}$$

On performing the summations and taking the limit of large numbers the above expression reduces to:

$$\begin{aligned} \mathcal{L}_2^{\text{cccc}} = & \bar{n}_1 \bar{n}_3 \bar{n}_4 \Theta_1 \Theta_3 \Theta_4 [1 + \xi_{13}^c + \xi_{14}^c + \xi_{34}^c + \zeta_{134}^c] \delta_{r,12}^D \delta_{M,12}^D + \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 [1 + \xi_{12}^c + \xi_{24}^c + \xi_{24}^c + \zeta_{124}^c] \delta_{r,23}^D \delta_{M,23}^D \\ & + \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 [1 + \xi_{12}^c + \xi_{13}^c + \xi_{23}^c + \zeta_{123}^c] \delta_{r,14}^D \delta_{M,14}^D + \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 [1 + \xi_{12}^c + \xi_{14}^c + \xi_{24}^c + \zeta_{124}^c] \delta_{r,23}^D \delta_{M,23}^D \\ & + \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 [1 + \xi_{12}^c + \xi_{23}^c + \xi_{31}^c + \zeta_{123}^c] \delta_{r,24}^D \delta_{M,24}^D + \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 [1 + \xi_{12}^c + \xi_{23}^c + \xi_{31}^c + \zeta_{123}^c] \delta_{r,34}^D \delta_{M,34}^D \end{aligned} \quad (\text{A10})$$

• $\mathcal{L}_3^{\text{cccc}}$: On performing the integrations over the delta functions we have:

$$\begin{aligned} \mathcal{L}_3^{\text{cccc}} \equiv & \sum_{i=j \neq k=l} p(\mathbf{r}_1, \mathbf{r}_3, M_1, M_3) \delta_{r,12}^D \delta_{r,34}^D \delta_{M,12}^D \delta_{M,34}^D + \sum_{i=k \neq j=l} p(\mathbf{r}_1, \mathbf{r}_2, M_1, M_2) \delta_{r,13}^D \delta_{r,24}^D \delta_{M,13}^D \delta_{M,24}^D \\ & + \sum_{i=l \neq j=k} p(\mathbf{r}_1, \mathbf{r}_2, M_1, M_2) \delta_{r,14}^D \delta_{r,23}^D \delta_{M,14}^D \delta_{M,23}^D. \end{aligned} \quad (\text{A11})$$

On performing the summations and taking the limit of large numbers of clusters we find:

$$\mathcal{L}_3 = \bar{n}_1 \bar{n}_3 \Theta_1 \Theta_3 [1 + \xi_{13}^c] \delta_{r,12}^D \delta_{M,12}^D \delta_{r,34}^D \delta_{M,34}^D + \bar{n}_1 \bar{n}_2 \Theta_1 \Theta_2 [1 + \xi_{12}^c] \left[\delta_{r,13}^D \delta_{r,24}^D \delta_{M,13}^D \delta_{M,24}^D + \delta_{r,14}^D \delta_{r,23}^D \delta_{M,14}^D \delta_{M,23}^D \right]. \quad (\text{A12})$$

• \mathcal{L}_4 : On performing the integrations over the delta functions we have:

$$\begin{aligned} \mathcal{L}_4^{\text{cccc}} \equiv & \sum_{i=j=k \neq l} p(\mathbf{r}_1, \mathbf{r}_4, M_1, M_4) \delta_{r,12}^D \delta_{r,13}^D \delta_{M,12}^D \delta_{M,13}^D + \sum_{i=j=l \neq k} p(\mathbf{r}_1, \mathbf{r}_3, M_1, M_3) \delta_{r,12}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,14}^D \\ & + \sum_{i=k=l \neq j} p(\mathbf{r}_1, \mathbf{r}_2, M_1, M_2) \delta_{r,13}^D \delta_{r,14}^D \delta_{M,13}^D \delta_{M,14}^D + \sum_{j=k=l \neq i} p(\mathbf{r}_1, \mathbf{r}_2, M_1, M_2) \delta_{r,23}^D \delta_{r,24}^D \delta_{M,23}^D \delta_{M,24}^D \end{aligned} \quad (\text{A13})$$

On performing the summations and taking the limit of large numbers of clusters we find:

$$\begin{aligned} \mathcal{L}_4^{\text{cccc}} = & \bar{n}_1 \bar{n}_4 \Theta_1 \Theta_4 [1 + \xi_{14}^c] \delta_{r,12}^D \delta_{r,13}^D \delta_{M,12}^D \delta_{M,13}^D + \bar{n}_1 \bar{n}_3 \Theta_1 \Theta_3 [1 + \xi_{13}^c] \delta_{r,12}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,14}^D \\ & + \bar{n}_1 \bar{n}_2 \Theta_1 \Theta_2 [1 + \xi_{12}^c] \{ \delta_{r,13}^D \delta_{r,14}^D \delta_{M,13}^D \delta_{M,14}^D + \delta_{r,23}^D \delta_{r,24}^D \delta_{M,23}^D \delta_{M,24}^D \} \end{aligned} \quad (\text{A14})$$

• $\mathcal{L}_5^{\text{cccc}}$: On performing the integrations over the delta functions we have:

$$\begin{aligned} \mathcal{L}_5^{\text{cccc}} \equiv & \sum_{i=j=k=l} p(\mathbf{r}_1, M_1) \delta_{r,12}^D \delta_{r,13}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,13}^D \delta_{M,14}^D \\ = & \bar{n}_1 \Theta_1 \delta_{r,12}^D \delta_{r,13}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,13}^D \delta_{M,14}^D \end{aligned} \quad (\text{A15})$$

On summation of the terms $\mathcal{L}_i^{\text{cccc}}$ we find:

$$\begin{aligned} \mathcal{L}^{\text{cccc}} = & \bar{n}_1 \Theta_1 \dots \bar{n}_4 \Theta_4 \left[1 + \xi_{12}^c + \xi_{13}^c + \xi_{14}^c + \xi_{23}^c + \xi_{24}^c + \xi_{34}^c + \zeta_{123}^c + \zeta_{234}^c + \zeta_{341}^c + \zeta_{412}^c + \xi_{12}^c \xi_{34}^c + \xi_{13}^c \xi_{24}^c + \xi_{14}^c \xi_{23}^c + \eta_{1234}^c \right] \\ & + \bar{n}_1 \bar{n}_3 \bar{n}_4 \Theta_1 \Theta_3 \Theta_4 \left[1 + \xi_{13}^c + \xi_{14}^c + \xi_{34}^c + \zeta_{134}^c \right] \delta_{r,12}^D \delta_{M,12}^D + \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \left[1 + \xi_{12}^c + \xi_{24}^c + \xi_{41}^c + \zeta_{124}^c \right] \delta_{r,23}^D \delta_{M,23}^D \\ & + \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left[1 + \xi_{12}^c + \xi_{13}^c + \xi_{23}^c + \zeta_{123}^c \right] \delta_{r,14}^D \delta_{M,14}^D + \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \left[1 + \xi_{12}^c + \xi_{14}^c + \xi_{24}^c + \zeta_{124}^c \right] \delta_{r,13}^D \delta_{M,13}^D \\ & + \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left[1 + \xi_{12}^c + \xi_{23}^c + \xi_{31}^c + \zeta_{123}^c \right] \delta_{r,24}^D \delta_{M,24}^D + \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left[1 + \xi_{12}^c + \xi_{23}^c + \xi_{31}^c + \zeta_{123}^c \right] \delta_{r,34}^D \delta_{M,34}^D \\ & + \bar{n}_1 \bar{n}_3 \Theta_1 \Theta_3 \left[1 + \xi_{13}^c \right] \delta_{r,12}^D \delta_{r,34}^D \delta_{M,12}^D \delta_{M,34}^D + \bar{n}_1 \bar{n}_2 \Theta_1 \Theta_2 \left[1 + \xi_{12}^c \right] \delta_{r,13}^D \delta_{r,24}^D \delta_{M,13}^D \delta_{M,24}^D \\ & + \bar{n}_1 \bar{n}_2 \Theta_1 \Theta_2 \left[1 + \xi_{12}^c \right] \delta_{r,14}^D \delta_{r,23}^D \delta_{M,14}^D \delta_{M,23}^D + \bar{n}_1 \bar{n}_4 \Theta_1 \Theta_4 \left[1 + \xi_{14}^c \right] \delta_{r,12}^D \delta_{r,13}^D \delta_{M,12}^D \delta_{M,13}^D \\ & + \bar{n}_1 \bar{n}_3 \Theta_1 \Theta_3 \left[1 + \xi_{13}^c \right] \delta_{r,12}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,14}^D + \bar{n}_1 \bar{n}_2 \Theta_1 \Theta_2 \left[1 + \xi_{12}^c \right] \{ \delta_{r,13}^D \delta_{r,14}^D \delta_{M,13}^D \delta_{M,14}^D + \delta_{r,23}^D \delta_{r,24}^D \delta_{M,23}^D \delta_{M,24}^D \} \\ & + \bar{n}_1 \Theta_1 \delta_{r,12}^D \delta_{r,13}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,13}^D \delta_{M,14}^D \end{aligned} \quad (\text{A16})$$

From Eq. (A16) we are now able to immediately write down all of the terms entering Eq. (A4). Thus we have:

$$\begin{aligned} \mathcal{L}^{\text{cccs}} \equiv & \alpha^{-1} \bar{n}_1 \Theta_1 \dots \bar{n}_4 \Theta_4 \left[1 + \xi_{12}^c + \xi_{13}^c + \xi_{23}^c + \zeta_{123}^c \right] \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_3 \bar{n}_4 \Theta_1 \Theta_3 \Theta_4 \left[1 + \xi_{13}^c \right] \delta_{r,12}^D \delta_{M,12}^D + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \left[1 + \xi_{12}^c \right] \delta_{r,23}^D \delta_{M,23}^D \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \left[1 + \xi_{12}^c \right] \delta_{r,13}^D \delta_{M,13}^D + \alpha^{-1} \bar{n}_1 \bar{n}_4 \Theta_1 \Theta_4 \delta_{r,12}^D \delta_{r,13}^D \delta_{M,12}^D \delta_{M,13}^D ; \end{aligned} \quad (\text{A17})$$

$$\begin{aligned} \mathcal{L}^{\text{ccss}} \equiv & \alpha^{-2} \bar{n}_1 \Theta_1 \dots \bar{n}_4 \Theta_4 \left[1 + \xi_{12}^c \right] + \alpha^{-2} \bar{n}_1 \bar{n}_3 \bar{n}_4 \Theta_1 \Theta_3 \Theta_4 \delta_{r,12}^D \delta_{M,12}^D + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left[1 + \xi_{12}^c \right] \delta_{r,34}^D \delta_{M,34}^D \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_3 \Theta_1 \Theta_3 \delta_{r,12}^D \delta_{r,34}^D \delta_{M,12}^D \delta_{M,34}^D ; \end{aligned} \quad (\text{A18})$$

$$\begin{aligned} \mathcal{L}^{\text{csss}} \equiv & \alpha^{-3} \bar{n}_1 \Theta_1 \dots \bar{n}_4 \Theta_4 + \alpha^{-2} \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \delta_{r,23}^D \delta_{M,23}^D + \alpha^{-2} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \delta_{r,24}^D \delta_{M,24}^D \\ & + \alpha^{-2} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \delta_{r,34}^D \delta_{M,34}^D + \alpha^{-1} \bar{n}_1 \bar{n}_2 \Theta_1 \Theta_2 \delta_{r,23}^D \delta_{r,24}^D \delta_{M,23}^D \delta_{M,24}^D ; \end{aligned} \quad (\text{A19})$$

$$\begin{aligned} \mathcal{L}^{\text{ssss}} \equiv & \alpha^{-4} \bar{n}_1 \Theta_1 \dots \bar{n}_4 \Theta_4 + \alpha^{-3} \bar{n}_1 \bar{n}_3 \bar{n}_4 \Theta_1 \Theta_3 \Theta_4 \delta_{r,12}^D \delta_{M,12}^D + \alpha^{-3} \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \delta_{r,23}^D \delta_{M,23}^D \\ & + \alpha^{-3} \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \delta_{r,13}^D \delta_{M,13}^D + \alpha^{-3} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \delta_{r,14}^D \delta_{M,14}^D + \alpha^{-3} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \delta_{r,24}^D \delta_{M,24}^D \\ & + \alpha^{-3} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \delta_{r,34}^D \delta_{M,34}^D + \alpha^{-2} \bar{n}_1 \bar{n}_3 \Theta_1 \Theta_3 \delta_{r,12}^D \delta_{r,34}^D \delta_{M,12}^D \delta_{M,34}^D + \alpha^{-2} \bar{n}_1 \bar{n}_4 \Theta_1 \Theta_4 \delta_{r,12}^D \delta_{r,13}^D \delta_{M,12}^D \delta_{M,13}^D \\ & + \alpha^{-2} \bar{n}_1 \bar{n}_3 \Theta_1 \Theta_3 \delta_{r,12}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,14}^D + \alpha^{-2} \bar{n}_1 \bar{n}_2 \Theta_1 \Theta_2 \left[\delta_{r,13}^D \delta_{r,24}^D \delta_{M,13}^D \delta_{M,24}^D + \delta_{r,14}^D \delta_{r,23}^D \delta_{M,14}^D \delta_{M,23}^D \right] \\ & + \alpha^{-2} \bar{n}_1 \bar{n}_2 \Theta_1 \Theta_2 \left[\delta_{r,13}^D \delta_{r,14}^D \delta_{M,13}^D \delta_{M,14}^D + \delta_{r,23}^D \delta_{r,24}^D \delta_{M,23}^D \delta_{M,24}^D \right] + \alpha^{-1} \bar{n}_1 \Theta_1 \delta_{r,12}^D \delta_{r,13}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,13}^D \delta_{M,14}^D \end{aligned} \quad (\text{A20})$$

Hence, for the second term in Eq. (A4) we have:

$$\begin{aligned} \mathcal{L}^{\text{cccs}} + 3\text{perms} = & \alpha^{-1} \bar{n}_1 \Theta_1 \dots \bar{n}_4 \Theta_4 \left[4 + 2\xi_{12}^c + 2\xi_{13}^c + 2\xi_{23}^c + 2\xi_{14}^c + 2\xi_{24}^c + 2\xi_{34}^c + \zeta_{123}^c + \zeta_{124}^c + \zeta_{134}^c + \zeta_{234}^c \right] \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_3 \bar{n}_4 \Theta_1 \Theta_3 \Theta_4 \left[1 + \xi_{13}^c \right] \delta_{r,12}^D \delta_{M,12}^D + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \left[1 + \xi_{12}^c \right] \delta_{r,23}^D \delta_{M,23}^D \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \left[1 + \xi_{12}^c \right] \delta_{r,31}^D \delta_{M,31}^D + \alpha^{-1} \bar{n}_1 \bar{n}_3 \bar{n}_4 \Theta_1 \Theta_3 \Theta_4 \left[1 + \xi_{14}^c \right] \delta_{r,12}^D \delta_{M,12}^D \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left[1 + \xi_{12}^c \right] \delta_{r,24}^D \delta_{M,24}^D + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left[1 + \xi_{12}^c \right] \delta_{r,41}^D \delta_{M,41}^D \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \left[1 + \xi_{14}^c \right] \delta_{r,13}^D \delta_{M,13}^D + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left[1 + \xi_{13}^c \right] \delta_{r,14}^D \delta_{M,14}^D \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left[1 + \xi_{13}^c \right] \delta_{r,34}^D \delta_{M,34}^D + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \left[1 + \xi_{24}^c \right] \delta_{r,23}^D \delta_{M,23}^D \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left[1 + \xi_{23}^c \right] \delta_{r,24}^D \delta_{M,24}^D + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left[1 + \xi_{23}^c \right] \delta_{r,34}^D \delta_{M,34}^D \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_4 \Theta_1 \Theta_4 \delta_{r,12}^D \delta_{r,13}^D \delta_{M,12}^D \delta_{M,13}^D + \alpha^{-1} \bar{n}_1 \bar{n}_3 \Theta_1 \Theta_3 \delta_{r,12}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,14}^D \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_2 \Theta_1 \Theta_2 \delta_{r,13}^D \delta_{r,14}^D \delta_{M,13}^D \delta_{M,14}^D + \alpha^{-1} \bar{n}_1 \bar{n}_2 \Theta_1 \Theta_2 \delta_{r,23}^D \delta_{r,24}^D \delta_{M,23}^D \delta_{M,24}^D . \end{aligned} \quad (\text{A21})$$

Hence, for the third term in Eq. (A4) we have:

$$\begin{aligned} \mathcal{L}^{\text{ccss}} + 5\text{perms} = & \alpha^{-2} \bar{n}_1 \Theta_1 \dots \bar{n}_4 \Theta_4 \left[6 + \xi_{12}^c + \xi_{13}^c + \xi_{14}^c + \xi_{23}^c + \xi_{24}^c + \xi_{34}^c \right] \\ & + \alpha^{-2} \bar{n}_1 \bar{n}_3 \bar{n}_4 \Theta_1 \Theta_3 \Theta_4 \delta_{r,12}^D \delta_{M,12}^D + \alpha^{-2} \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \left[\delta_{r,13}^D \delta_{M,13}^D + \delta_{r,23}^D \delta_{M,23}^D \right] \\ & + \alpha^{-2} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left[\delta_{r,14}^D \delta_{M,14}^D + \delta_{r,24}^D \delta_{M,24}^D + \delta_{r,34}^D \delta_{M,34}^D \right] \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_3 \Theta_1 \Theta_2 \Theta_3 \left\{ \left[1 + \xi_{12}^c \right] \delta_{r,34}^D \delta_{M,34}^D + \left[1 + \xi_{13}^c \right] \delta_{r,24}^D \delta_{M,24}^D + \left[1 + \xi_{23}^c \right] \delta_{r,14}^D \delta_{M,14}^D \right\} \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_4 \Theta_1 \Theta_2 \Theta_4 \left\{ \left[1 + \xi_{14}^c \right] \delta_{r,23}^D \delta_{M,23}^D + \left[1 + \xi_{24}^c \right] \delta_{r,13}^D \delta_{M,13}^D \right\} \\ & + \alpha^{-1} \bar{n}_1 \bar{n}_3 \bar{n}_4 \Theta_1 \Theta_3 \Theta_4 \left[1 + \xi_{34}^c \right] \delta_{r,12}^D \delta_{M,12}^D + 2\alpha^{-1} \bar{n}_1 \bar{n}_3 \Theta_1 \Theta_3 \delta_{r,12}^D \delta_{r,34}^D \delta_{M,12}^D \delta_{M,34}^D \\ & + 2\alpha^{-1} \bar{n}_1 \bar{n}_2 \Theta_1 \Theta_2 \left[\delta_{r,13}^D \delta_{r,24}^D \delta_{M,13}^D \delta_{M,24}^D + \delta_{r,14}^D \delta_{r,23}^D \delta_{M,14}^D \delta_{M,23}^D \right] . \end{aligned} \quad (\text{A22})$$

Hence, for the fourth term in Eq. (A4) we have:

$$\begin{aligned}
\mathcal{L}^{csss} + 3 \text{ perms} &= 4\alpha^{-3}\bar{n}_1\Theta_1 \dots \bar{n}_4\Theta_4 + 2\alpha^{-2}\bar{n}_1\bar{n}_2\bar{n}_4\Theta_1\Theta_2\Theta_4\delta_{r,23}^D\delta_{M,23}^D + 2\alpha^{-2}\bar{n}_1\bar{n}_2\bar{n}_3\Theta_1\Theta_2\Theta_4\delta_{r,24}^D\delta_{M,24}^D \\
&+ 2\alpha^{-2}\bar{n}_1\bar{n}_2\bar{n}_3\Theta_1\Theta_2\Theta_4\delta_{r,34}^D\delta_{M,34}^D + 2\alpha^{-2}\bar{n}_1\bar{n}_2\bar{n}_4\Theta_1\Theta_2\Theta_4\delta_{r,13}^D\delta_{M,13}^D \\
&+ 2\alpha^{-2}\bar{n}_1\bar{n}_2\bar{n}_3\Theta_1\Theta_2\Theta_3\delta_{r,14}^D\delta_{M,14}^D + 2\alpha^{-2}\bar{n}_1\bar{n}_3\bar{n}_4\Theta_1\Theta_3\Theta_4\delta_{r,12}^D\delta_{M,12}^D \\
&+ \alpha^{-1}\bar{n}_1\bar{n}_2\Theta_1\Theta_2\delta_{r,23}^D\delta_{r,24}^D\delta_{M,23}^D\delta_{M,24}^D + \alpha^{-1}\bar{n}_1\bar{n}_2\Theta_1\Theta_2\delta_{r,13}^D\delta_{r,14}^D\delta_{M,13}^D\delta_{M,14}^D \\
&+ \alpha^{-1}\bar{n}_1\bar{n}_3\Theta_1\Theta_3\delta_{r,12}^D\delta_{r,14}^D\delta_{M,12}^D\delta_{M,14}^D + \alpha^{-1}\bar{n}_1\bar{n}_4\Theta_1\Theta_4\delta_{r,12}^D\delta_{r,13}^D\delta_{M,12}^D\delta_{M,13}^D. \tag{A23}
\end{aligned}$$

On collecting all of the terms and after a little algebra we find that Eq. (A3) can be written:

$$\begin{aligned}
\langle \mathcal{F}_c(\mathbf{r}_1) \dots \mathcal{F}_c(\mathbf{r}_4) \rangle &= \frac{1}{A^2} \prod_{p=1}^4 \left\{ \int dM_p w(\mathbf{x}_p, M_p) \bar{n}(M_p) \Theta(\mathbf{x}_p, M_p) \right\} \left\{ \eta_{1234}^c \right. \\
&+ \left[\xi_{12}^c + \frac{(1+\alpha)}{\bar{n}_2\Theta_2} \delta_{r,12}^D \delta_{M,12}^D \right] \left[\xi_{34}^c + \frac{(1+\alpha)}{\bar{n}_4\Theta_4} \delta_{r,34}^D \delta_{M,34}^D \right] + \left[\xi_{13}^c + \frac{(1+\alpha)}{\bar{n}_3\Theta_3} \delta_{r,13}^D \delta_{M,13}^D \right] \left[\xi_{24}^c + \frac{(1+\alpha)}{\bar{n}_4\Theta_4} \delta_{r,24}^D \delta_{M,24}^D \right] \\
&+ \left[\xi_{14}^c + \frac{(1+\alpha)}{\bar{n}_4\Theta_4} \delta_{r,14}^D \delta_{M,14}^D \right] \left[\xi_{23}^c + \frac{(1+\alpha)}{\bar{n}_3\Theta_3} \delta_{r,23}^D \delta_{M,23}^D \right] + \zeta_{134}^c \frac{\delta_{r,12}^D \delta_{M,12}^D}{\bar{n}_2\Theta_2} + \zeta_{124}^c \frac{\delta_{r,23}^D \delta_{M,23}^D}{\bar{n}_3\Theta_3} + \zeta_{123}^c \frac{\delta_{r,14}^D \delta_{M,14}^D}{\bar{n}_4\Theta_4} \\
&+ \zeta_{124}^c \frac{\delta_{r,13}^D \delta_{M,13}^D}{\bar{n}_3\Theta_3} + \zeta_{123}^c \frac{\delta_{r,24}^D \delta_{M,24}^D}{\bar{n}_4\Theta_4} + \zeta_{123}^c \frac{\delta_{r,34}^D \delta_{M,34}^D}{\bar{n}_4\Theta_4} + \xi_{13}^c \frac{\delta_{r,12}^D \delta_{r,34}^D \delta_{M,12}^D \delta_{M,34}^D}{\bar{n}_2\bar{n}_4\Theta_2\Theta_4} + \xi_{12}^c \frac{\delta_{r,13}^D \delta_{r,24}^D \delta_{M,13}^D \delta_{M,24}^D}{\bar{n}_3\bar{n}_4\Theta_3\Theta_4} \\
&+ \xi_{12}^c \frac{\delta_{r,14}^D \delta_{r,23}^D \delta_{M,14}^D \delta_{M,23}^D}{\bar{n}_3\bar{n}_4\Theta_3\Theta_4} + \xi_{14}^c \frac{\delta_{r,12}^D \delta_{r,13}^D \delta_{M,12}^D \delta_{M,13}^D}{\bar{n}_2\bar{n}_3\Theta_2\Theta_3} + \xi_{13}^c \frac{\delta_{r,12}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,14}^D}{\bar{n}_2\bar{n}_4\Theta_2\Theta_4} + \xi_{12}^c \frac{\delta_{r,13}^D \delta_{r,14}^D \delta_{M,13}^D \delta_{M,14}^D}{\bar{n}_3\bar{n}_4\Theta_3\Theta_4} \\
&+ \xi_{12}^c \frac{\delta_{r,23}^D \delta_{M,23}^D \delta_{r,24}^D \delta_{M,24}^D}{\bar{n}_3\bar{n}_4\Theta_3\Theta_4} + \left. \frac{(1+\alpha^3)\delta_{r,12}^D \delta_{r,13}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,13}^D \delta_{M,14}^D}{\bar{n}_2\bar{n}_3\bar{n}_4\Theta_2\Theta_3\Theta_4} \right\} \tag{A24}
\end{aligned}$$

The last factor on the right-hand-side of Eq. (A1) can be written:

$$\begin{aligned}
\langle \mathcal{F}_c(\mathbf{x}_1) \mathcal{F}_c(\mathbf{x}_3) \rangle \langle \mathcal{F}_c(\mathbf{x}_2) \mathcal{F}_c(\mathbf{x}_4) \rangle &= \prod_{p=1}^4 \left\{ \int dM_p w(\mathbf{x}_p, M_p) \bar{n}(M_p) \Theta(\mathbf{x}_p, M_p) \right\} \\
&\left[\xi_{13}^c + \frac{(1+\alpha)\delta_{r,13}^D \delta_{M,13}^D}{\bar{n}_3\Theta(M_3, \mathbf{x}_3)} \right] \left[\xi_{24}^c + \frac{(1+\alpha)\delta_{r,24}^D \delta_{M,24}^D}{\bar{n}_4\Theta(M_4, \mathbf{x}_4)} \right]. \tag{A25}
\end{aligned}$$

On subtracting Eq. (A25) from Eq. (A24) we see that the argument of the bracket on the right-hand-side of Eq. (A1) can be expressed as:

$$\begin{aligned}
\langle \mathcal{F}_c(\mathbf{x}_1) \dots \mathcal{F}_c(\mathbf{x}_4) \rangle - \langle \mathcal{F}_c(\mathbf{x}_1) \mathcal{F}_c(\mathbf{x}_3) \rangle \langle \mathcal{F}_c(\mathbf{x}_2) \mathcal{F}_c(\mathbf{x}_4) \rangle &= \frac{1}{A^2} \prod_{p=1}^4 \left\{ \int dM_p w(\mathbf{x}_p, M_p) \bar{n}(M_p) \Theta(\mathbf{x}_p, M_p) \right\} \left\{ \eta_{1234}^c + \right. \\
&\times \left[\xi_{12}^c + \frac{(1+\alpha)}{\bar{n}_2\Theta_2} \delta_{r,12}^D \delta_{M,12}^D \right] \left[\xi_{34}^c + \frac{(1+\alpha)}{\bar{n}_4\Theta_4} \delta_{r,34}^D \delta_{M,34}^D \right] + \left[\xi_{14}^c + \frac{(1+\alpha)}{\bar{n}_4\Theta_4} \delta_{r,14}^D \delta_{M,14}^D \right] \left[\xi_{23}^c + \frac{(1+\alpha)}{\bar{n}_3\Theta_3} \delta_{r,23}^D \delta_{M,23}^D \right] \\
&+ \zeta_{134}^c \frac{\delta_{r,12}^D \delta_{M,12}^D}{\bar{n}_2\Theta_2} + \zeta_{124}^c \frac{\delta_{r,23}^D \delta_{M,23}^D}{\bar{n}_3\Theta_3} + \zeta_{123}^c \frac{\delta_{r,14}^D \delta_{M,14}^D}{\bar{n}_4\Theta_4} + \zeta_{124}^c \frac{\delta_{r,13}^D \delta_{M,13}^D}{\bar{n}_3\Theta_3} + \zeta_{123}^c \frac{\delta_{r,24}^D \delta_{M,24}^D}{\bar{n}_4\Theta_4} + \zeta_{123}^c \frac{\delta_{r,34}^D \delta_{M,34}^D}{\bar{n}_4\Theta_4} \\
&+ \xi_{13}^c \frac{\delta_{r,12}^D \delta_{r,34}^D \delta_{M,12}^D \delta_{M,34}^D}{\bar{n}_2\bar{n}_4\Theta_2\Theta_4} + \xi_{12}^c \frac{\delta_{r,13}^D \delta_{r,24}^D \delta_{M,13}^D \delta_{M,24}^D}{\bar{n}_3\bar{n}_4\Theta_3\Theta_4} + \xi_{12}^c \frac{\delta_{r,14}^D \delta_{r,23}^D \delta_{M,14}^D \delta_{M,23}^D}{\bar{n}_3\bar{n}_4\Theta_3\Theta_4} + \xi_{14}^c \frac{\delta_{r,12}^D \delta_{r,13}^D \delta_{M,12}^D \delta_{M,13}^D}{\bar{n}_2\bar{n}_3\Theta_2\Theta_3} \\
&+ \xi_{13}^c \frac{\delta_{r,12}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,14}^D}{\bar{n}_2\bar{n}_4\Theta_2\Theta_4} + \xi_{12}^c \frac{\delta_{r,13}^D \delta_{r,14}^D \delta_{M,13}^D \delta_{M,14}^D}{\bar{n}_3\bar{n}_4\Theta_3\Theta_4} + \xi_{12}^c \frac{\delta_{r,23}^D \delta_{M,23}^D \delta_{r,24}^D \delta_{M,24}^D}{\bar{n}_3\bar{n}_4\Theta_3\Theta_4} \\
&+ \left. \frac{(1+\alpha^3)\delta_{r,12}^D \delta_{r,13}^D \delta_{r,14}^D \delta_{M,12}^D \delta_{M,13}^D \delta_{M,14}^D}{\bar{n}_2\bar{n}_3\bar{n}_4\Theta_2\Theta_3\Theta_4} \right\}. \tag{A26}
\end{aligned}$$

On substituting the above relation into Eq. (A1) we recover Eq. (50).

2. General expression for the covariance matrix of the power spectrum of \mathcal{F}_c

Consider Eq. (50), we may Fourier transform all of the space dependent terms. For the case of the n -point correlation functions, these are Fourier dual with the n -point multi-spectra:

$$\begin{aligned}\xi_{12} &\equiv \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} \frac{d^3\mathbf{q}_2}{(2\pi)^3} (2\pi)^3 \delta^D(\mathbf{q}_1 + \mathbf{q}_2) P(\mathbf{q}_1, \mathbf{q}_2) ; \\ \zeta_{123} &\equiv \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} \frac{d^3\mathbf{q}_2}{(2\pi)^3} \frac{d^3\mathbf{q}_3}{(2\pi)^3} (2\pi)^3 \delta^D(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3) B(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \exp[-i\mathbf{q}_1 \cdot \mathbf{r}_1 - i\mathbf{q}_2 \cdot \mathbf{r}_2 - i\mathbf{q}_3 \cdot \mathbf{r}_3] ; \\ \eta_{1234} &\equiv \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} \dots \frac{d^3\mathbf{q}_4}{(2\pi)^3} (2\pi)^3 \delta^D(\mathbf{q}_1 + \dots + \mathbf{q}_4) T(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4) \exp[-i\mathbf{q}_1 \cdot \mathbf{r}_1 \dots - i\mathbf{q}_4 \cdot \mathbf{r}_4] ,\end{aligned}\quad (\text{A27})$$

where B and T are bispectrum and trispectrum, respectively. Note that, owing to the Dirac delta function in the above expressions the bispectrum and trispectrum are in fact functions of two and three k -vectors, respectively. Using these relations in Eq. (50), we find that the covariance matrix may be written in general as:

$$\begin{aligned}\text{Cov}[|\mathcal{F}_c(\mathbf{k}_1)|^2, |\mathcal{F}_c(\mathbf{k}_2)|^2] &= \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} \dots \frac{d^3\mathbf{q}_3}{(2\pi)^3} T(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \mathcal{G}_{(1,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_2 - \mathbf{q}_2) \\ &\quad \times \mathcal{G}_{(1,1)}(-\mathbf{k}_1 - \mathbf{q}_3) \mathcal{G}_{(1,1)}(-\mathbf{k}_2 + \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3) \\ &\quad + \left| \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} P(\mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_2 + \mathbf{q}_1) + (1 + \alpha) \mathcal{G}_{2,0}(\mathbf{k}_1 + \mathbf{k}_2) \right|^2 \\ &\quad + \left| \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} P(\mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_2 + \mathbf{q}_1) + (1 + \alpha) \mathcal{G}_{2,0}(\mathbf{k}_1 - \mathbf{k}_2) \right|^2 \\ &\quad + \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} \frac{d^3\mathbf{q}_2}{(2\pi)^3} B(\mathbf{q}_1, \mathbf{q}_2) \mathcal{G}_{(2,1)}(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_1 - \mathbf{q}_2) \mathcal{G}_{(1,1)}(-\mathbf{k}_2 + \mathbf{q}_1 + \mathbf{q}_2) \\ &\quad + \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} \frac{d^3\mathbf{q}_2}{(2\pi)^3} B(\mathbf{q}_1, \mathbf{q}_2) \mathcal{G}_{(2,1)}(\mathbf{k}_1 - \mathbf{k}_2 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(+\mathbf{k}_2 + \mathbf{q}_1 + \mathbf{q}_2) \\ &\quad + \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} \frac{d^3\mathbf{q}_2}{(2\pi)^3} B(\mathbf{q}_1, \mathbf{q}_2) \mathcal{G}_{(2,1)}(-\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_2 + \mathbf{q}_1 + \mathbf{q}_2) \\ &\quad + \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} \frac{d^3\mathbf{q}_2}{(2\pi)^3} B(\mathbf{q}_1, \mathbf{q}_2) \mathcal{G}_{(2,1)}(-\mathbf{k}_2 - \mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_2 + \mathbf{q}_1 + \mathbf{q}_2) \\ &\quad + \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} \frac{d^3\mathbf{q}_2}{(2\pi)^3} B(\mathbf{q}_1, \mathbf{q}_2) \mathcal{G}_{(2,1)}(-\mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_1 + \mathbf{q}_1 + \mathbf{q}_2) \\ &\quad + \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} \frac{d^3\mathbf{q}_2}{(2\pi)^3} B(\mathbf{q}_1, \mathbf{q}_2) \mathcal{G}_{(2,1)}(-\mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_2 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_2 + \mathbf{q}_1 + \mathbf{q}_2) \\ &\quad + \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} P(\mathbf{q}_1) \left\{ |\mathcal{G}_{(2,1)}(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}_1)|^2 + |\mathcal{G}_{(2,1)}(\mathbf{q}_1)|^2 + |\mathcal{G}_{(2,1)}(\mathbf{k}_1 - \mathbf{k}_2 - \mathbf{q}_1)|^2 \right\} \\ &\quad + \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} P(\mathbf{q}_1) \left\{ \mathcal{G}_{(3,1)}(\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_1 + \mathbf{q}_1) + \mathcal{G}_{(3,1)}(\mathbf{k}_2 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(-\mathbf{k}_2 + \mathbf{q}_1) \right. \\ &\quad \quad \left. + \mathcal{G}_{(3,1)}(-\mathbf{k}_1 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_1 + \mathbf{q}_1) + \mathcal{G}_{(3,1)}(-\mathbf{k}_2 - \mathbf{q}_1) \mathcal{G}_{(1,1)}(\mathbf{k}_2 + \mathbf{q}_1) \right\} \\ &\quad + (1 + \alpha^3) \mathcal{G}_{(4,0)}(0) .\end{aligned}\quad (\text{A28})$$

3. A result concerning the shell averaging of products of the \mathcal{Q} functions

Consider the following integral:

$$\begin{aligned}\mathcal{P}_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}[k_i, k_j] &\equiv \int_{V_i} \frac{d^3\mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3\mathbf{k}_2}{V_j} \mathcal{Q}_{(j_1 \dots j_n)}^{(i_1 \dots i_n)}(\mathbf{k}_1 + \mathbf{k}_2) \mathcal{Q}_{(m_1 \dots m_n)}^{(l_1 \dots l_n)}(-\mathbf{k}_1 - \mathbf{k}_2) \\ &= \int_{V_i} \frac{d^3\mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3\mathbf{k}_2}{V_j} \int d^3\mathbf{x}_1 d^3\mathbf{x}_2 \mathcal{Q}_{(j_1 \dots j_n)}^{(i_1 \dots i_n)}(\mathbf{x}_1) \mathcal{Q}_{(m_1 \dots m_n)}^{(l_1 \dots l_n)}(\mathbf{x}_2) e^{i(\mathbf{k}_1 + \mathbf{k}_2) \cdot (\mathbf{x}_1 - \mathbf{x}_2)} .\end{aligned}\quad (\text{A29})$$

On integrating the exponential functions over the angles $\hat{\mathbf{k}}_1$ and $\hat{\mathbf{k}}_2$ the above expression simplifies to:

$$\mathcal{P}_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}[k_i, k_j] = \int d^3 \mathbf{x}_1 d^3 \mathbf{x}_2 \overline{j_0}(k_i |\mathbf{x}_1 - \mathbf{x}_2|) \overline{j_0}(k_j |\mathbf{x}_1 - \mathbf{x}_2|) \mathcal{Q}_{(j_1 \dots j_n)}^{(i_1 \dots i_n)}(\mathbf{x}_1) \mathcal{Q}_{(m_1 \dots m_n)}^{(l_1 \dots l_n)}(\mathbf{x}_2), \quad (\text{A30})$$

where in the above equation we defined the shell-average of the spherical Bessel function as

$$\overline{j_0}(k_i x) \equiv \frac{1}{V_i} \int_{k_i - \Delta k/2}^{k_i + \Delta k/2} dk_1 k_1^2 4\pi j_0(k_1 x). \quad (\text{A31})$$

On performing a change of variables $\mathbf{x}_{21} = \mathbf{x}_2 - \mathbf{x}_1$, then the above integral further simplifies to:

$$\mathcal{P}_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}[k_i, k_j] = \int d^3 \mathbf{x}_{21} \overline{j_0}(k_i x_{21}) \overline{j_0}(k_j x_{21}) \Xi_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}(x_{21}), \quad (\text{A32})$$

where in the above we have defined the correlation function of the weighted survey window function to be,

$$\Xi_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}(x_{21}) \equiv \int \frac{d^2 \hat{\mathbf{x}}_{21}}{4\pi} \int d^3 \mathbf{x}_1 \mathcal{Q}_{(j_1 \dots j_n)}^{(i_1 \dots i_n)}(\mathbf{x}_1) \mathcal{Q}_{(m_1 \dots m_n)}^{(l_1 \dots l_n)}(\mathbf{x}_{21} + \mathbf{x}_1). \quad (\text{A33})$$

In the limit that the survey volume is large, the weighted survey window correlation function is very slowly varying over nearly all length scales of interest, and so can be approximated by its value at zero-lag. Hence,

$$\mathcal{P}_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}[k_i, k_j] \approx \Xi_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}(0) \int d^3 \mathbf{x}_{21} \overline{j_0}(x_{21} k_i) \overline{j_0}(x_{21} k_j). \quad (\text{A34})$$

The spherical Bessel functions obey the following orthogonality relationship:

$$\int_0^\infty dx x^2 j_\alpha(ux) j_\alpha(vx) = \frac{\pi}{2u^2} \delta^D(u - v). \quad (\text{A35})$$

Using this relationship in Eq. (A34) gives,

$$\begin{aligned} \mathcal{P}_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}[k_i, k_j] &\approx \Xi_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}(0) \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \int_0^\infty dr_{21} 4\pi r_{21}^2 j_0(r_{21} k_1) j_0(r_{21} k_2) \\ &= \Xi_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}(0) \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \frac{2\pi^2}{k_1^2} \delta^D(k_1 - k_2). \end{aligned} \quad (\text{A36})$$

We note that the delta function will only give a non-zero value when the two k -space shells are overlapping, i.e. when $V_i = V_j$. Thus the above expression simplifies to:

$$\begin{aligned} \mathcal{P}_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}[k_i, k_j] &\approx \Xi_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}(0) \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i^2} 4\pi k_1^2 \frac{2\pi^2}{k_1^2} \\ &= \frac{(2\pi)^3}{V_i} \Xi_{(j_1 \dots j_n | m_1 \dots m_n)}^{(i_1 \dots i_n | l_1 \dots l_n)}(0) \delta_{i,j}^K. \end{aligned} \quad (\text{A37})$$

The specific cases that we require to evaluate Eq. (60) are

$$\mathcal{P}_1[k_i, k_j] \equiv \mathcal{P}_{(11|11)}^{(11|11)}[k_i, k_j] = \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{k}_1 + \mathbf{k}_2) \mathcal{Q}_{(1,1)}^{(1,1)}(-\mathbf{k}_1 - \mathbf{k}_2); \quad (\text{A38})$$

$$\mathcal{P}_2[k_i, k_j] \equiv \mathcal{P}_{(11|0)}^{(11|2)}[k_i, k_j] = \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \mathcal{Q}_{(1,1)}^{(1,1)}(\mathbf{k}_1 + \mathbf{k}_2) \mathcal{Q}_{(0)}^{(2)}(-\mathbf{k}_1 - \mathbf{k}_2); \quad (\text{A39})$$

$$\mathcal{P}_3[k_i, k_j] \equiv \mathcal{P}_{(0|0)}^{(2|2)}[k_i, k_j] = \int_{V_i} \frac{d^3 \mathbf{k}_1}{V_i} \int_{V_j} \frac{d^3 \mathbf{k}_2}{V_j} \mathcal{Q}_{(0)}^{(2)}(\mathbf{k}_1 + \mathbf{k}_2) \mathcal{Q}_{(0)}^{(2)}(-\mathbf{k}_1 - \mathbf{k}_2). \quad (\text{A40})$$

These may be written using our result Eq. (A37), where:

$$\Xi_1(0) \equiv \Xi_{(11|11)}^{(11|11)}(0) = \frac{\int d^3\mathbf{r} [\int dM \bar{n}(M) b(M) w(\mathbf{r}, M) \Theta(\mathbf{r}, M)]^4}{\left\{ \int d^3\mathbf{r} [\int dM \bar{n}(M) b(M) w(\mathbf{r}, M) \Theta(\mathbf{r}, M)]^2 \right\}^2}; \quad (\text{A41})$$

$$\Xi_2(0) \equiv \Xi_{(11|0)}^{(11|2)}(0) = \frac{\int d^3\mathbf{r} [\int dM \bar{n}(M) b(M) w(\mathbf{r}, M) \Theta(\mathbf{r}, M)]^2 [\int dM \bar{n}(M) w^2(\mathbf{r}, M) \Theta(\mathbf{r}, M)]^2}{\left\{ \int d^3\mathbf{r} [\int dM \bar{n}(M) b(M) w(\mathbf{r}, M) \Theta(\mathbf{r}, M)]^2 \right\}^2}; \quad (\text{A42})$$

$$\Xi_3(0) \equiv \Xi_{(0|0)}^{(2|2)}(0) = \frac{\int d^3\mathbf{r} [\int dM \bar{n}(M) w^2(\mathbf{r}, M) \Theta(\mathbf{r}, M)]^4}{\left\{ \int d^3\mathbf{r} [\int dM \bar{n}(M) b(M) w(\mathbf{r}, M) \Theta(\mathbf{r}, M)]^2 \right\}^2}. \quad (\text{A43})$$

Appendix B: Functional derivatives

1. Functional derivative of $F[w(\mathbf{x}, M)]$

On making the variation of $w \rightarrow w + \delta w$ the functional $F[w]$ can be written as:

$$\begin{aligned} F[w + \delta w] &= C_1 \frac{S_1[w + \delta w]}{S_4[w + \delta w]} + 2C_1 C_2 \frac{S_2[w + \delta w]}{S_4[w + \delta w]} + C_1 C_2^2 \frac{S_3[w + \delta w]}{S_4[w + \delta w]} \\ &= C_1 \frac{S_1[w]}{S_4[w]} \left[1 + \frac{\delta S_1}{S_1} - \frac{\delta S_4}{S_4} \right] + 2C_1 C_2 \frac{S_2[w]}{S_4[w]} \left[1 + \frac{\delta S_2}{S_3} - \frac{\delta S_4}{S_4} \right] + C_1 C_2^2 \frac{S_3[w]}{S_4[w]} \left[1 + \frac{\delta S_3}{S_3} - \frac{\delta S_4}{S_4} \right] \\ &= F[w] + C_1 \left[\frac{\delta S_1}{S_4} + 2C_2 \frac{\delta S_2}{S_4} + C_2^2 \frac{\delta S_3}{S_4} \right] - C_1 (S_1 + 2C_2 S_2 + C_2^2 S_3) \frac{\delta S_4}{S_4^2}. \end{aligned} \quad (\text{B1})$$

2. Functional derivative of $S_1[w(\mathbf{x}, M)]$

Let us now compute the functional derivatives of S_1 , this can be written:

$$\begin{aligned} S_1[w + \delta w] &= \int d^3\mathbf{x} \left[\tilde{\mathcal{G}}_{(1)}^{(1)}[w + \delta w] \right]^4 \\ &= \int d^3\mathbf{x} \left[\int dM \bar{n}(M) b(M) \Theta(\mathbf{x}, M) [w + \delta w] \right]^4 \\ &= \int d^3\mathbf{x} \left[\tilde{\mathcal{G}}_{(1)}^{(1)} \right]^4 \left[1 + \frac{1}{\tilde{\mathcal{G}}_{(1)}^{(1)}} \int dM \bar{n}(M) b(M) \Theta(\mathbf{x}, M) \delta w \right]^4 \\ &\approx S_1[w] + \int d^3\mathbf{x} \int dM \left\{ 4 \left[\tilde{\mathcal{G}}_{(1)}^{(1)} \right]^3 \bar{n}(M) b(M) \Theta(\mathbf{x}, M) \right\} \delta w(\mathbf{x}, M). \end{aligned} \quad (\text{B2})$$

3. Functional derivative of $S_2[w(\mathbf{x}, M)]$

Let us now compute the functional derivatives of S_2 , this can be written:

$$\begin{aligned}
S_2[w + \delta w] &= \int d^3\mathbf{x} \left[\tilde{\mathcal{G}}_{(1)}^{(1)}[w + \delta w] \right]^2 \mathcal{G}_{(2)}^{(0)}[w + \delta w] \\
&= \int d^3\mathbf{x} \left\{ \int dM \bar{n}(M) b(M) \Theta(\mathbf{x}, M) [w(\mathbf{x}, M) + \delta w(\mathbf{x}, M)] \right\}^2 \int dM' \bar{n}(M') \Theta(\mathbf{x}, M') [w(\mathbf{x}, M') + \delta w(\mathbf{x}, M')]^2 \\
&= \int d^3\mathbf{x} \left\{ \left[\tilde{\mathcal{G}}_{(1)}^{(1)}(\mathbf{x}) \right]^2 + 2 \tilde{\mathcal{G}}_{(1)}^{(1)}(\mathbf{x}) \int dM \bar{n}(M) b(M) \Theta(\mathbf{x}, M) \delta w(\mathbf{x}, M) \right\} \\
&\quad \times \left[\tilde{\mathcal{G}}_{(2)}^{(0)}(\mathbf{x}) + 2 \int dM' \bar{n}(M') \Theta(\mathbf{x}, M') w(\mathbf{x}, M') \delta w(\mathbf{x}, M') \right] \\
&= S_2[w] + 2 \int d^3\mathbf{x} \tilde{\mathcal{G}}_{(1)}^{(1)}(\mathbf{x}) \tilde{\mathcal{G}}_{(2)}^{(0)}(\mathbf{x}) \int dM \bar{n}(M) b(M) \Theta(\mathbf{x}, M) \delta w(\mathbf{x}, M) \\
&\quad + 2 \int d^3\mathbf{x} \left[\tilde{\mathcal{G}}_{(1)}^{(1)}(\mathbf{x}) \right]^2 \int dM \bar{n}(M) \Theta(\mathbf{x}, M) w(\mathbf{x}, M) \delta w(\mathbf{x}, M) \\
&= S_2[w] + \int d^3\mathbf{x} \int dM 2\bar{n}(M) b(M) \Theta(\mathbf{x}, M) \left\{ \tilde{\mathcal{G}}_{(1)}^{(1)}(\mathbf{x}) \tilde{\mathcal{G}}_{(2)}^{(0)}(\mathbf{x}) + \left[\tilde{\mathcal{G}}_{(1)}^{(1)}(\mathbf{x}) \right]^2 \frac{w(\mathbf{x}, M)}{b(M)} \right\} \delta w(\mathbf{x}, M) . \tag{B3}
\end{aligned}$$

4. Functional derivative of $S_3[w(\mathbf{x}, M)]$

Let us now compute the functional derivatives of S_3 , this can be written:

$$\begin{aligned}
S_3[w + \delta w] &= \int d^3\mathbf{x} \left[\mathcal{G}_{(2)}^{(0)}[w + \delta w] \right]^2 \\
&= \int d^3\mathbf{x} \left\{ \int dM \bar{n}(M) \Theta(\mathbf{x}, M) [w + \delta w]^2 \right\}^2 \\
&= \int d^3\mathbf{x} \left\{ \mathcal{G}_{(2)}^{(0)} + 2 \int dM \bar{n}(M) \Theta(\mathbf{x}, M) w \delta w \right\}^2 \\
&= S_3[w] + \int d^3\mathbf{x} \int dM \left\{ 4\mathcal{G}_{(2)}^{(0)}(\mathbf{x}) \bar{n}(M) \Theta(\mathbf{x}, M) w(\mathbf{x}, M) \right\} \delta w . \tag{B4}
\end{aligned}$$

5. Functional derivative of $S_4[w(\mathbf{x}, M)]$

Let us now compute the functional derivatives of S_4 , this can be written:

$$\begin{aligned}
S_4[w + \delta w] &= \left\{ \int d^3\mathbf{x} \left[\mathcal{G}_{(1)}^{(1)}[w + \delta w] \right]^2 \right\}^2 \\
&= \left\{ \int d^3\mathbf{x} \left[\int dM \bar{n}(M) b(M) \Theta(\mathbf{x}, M) w(\mathbf{x}, M) \right]^2 \left(1 + \frac{2}{\mathcal{G}_{(1)}^{(1)}} \int dM \bar{n}(M) b(M) \delta w(\mathbf{x}, M) \right) \right\}^2 \\
&= S_4[w] + 4 \int d^3\mathbf{x}' \left[\mathcal{G}_{(1)}^{(1)}(\mathbf{x}') \right]^2 \int d^3\mathbf{x} \mathcal{G}_{(1)}^{(1)}(\mathbf{x}) \int dM \bar{n}(M) b(M) \Theta(\mathbf{x}, M) \delta w(\mathbf{x}, M) \\
&= S_4[w] + \int d^3\mathbf{x} \int dM \left\{ 4\mathcal{G}_{(1)}^{(1)}(\mathbf{x}) \bar{n}(M) b(M) \Theta(\mathbf{x}, M) \int d^3\mathbf{x}' \left[\mathcal{G}_{(1)}^{(1)}(\mathbf{x}') \right]^2 \right\} \delta w(\mathbf{x}, M) . \tag{B5}
\end{aligned}$$

Appendix C: Evaluation of the $S_i[w]$ functions for various weighting schemes

1. Volume-limited samples

The $S_i^{\text{VL}}[w]$ functions for the weighting scheme, $w(M) = 1$ are:

$$S_1^{\text{VL}}[w = 1] = V_\mu \bar{n}_h^4 \bar{b}_N^{-4} ; \quad (\text{C1})$$

$$S_2^{\text{VL}}[w = 1] = V_\mu \bar{n}_h^3 \bar{b}_N^{-2} ; \quad (\text{C2})$$

$$S_3^{\text{VL}}[w = 1] = V_\mu \bar{n}_h^2 ; \quad (\text{C3})$$

$$S_4^{\text{VL}}[w = 1] = V_\mu^2 \bar{n}_h^4 \bar{b}_N^{-4} . \quad (\text{C4})$$

The $S_i^{\text{VL}}[w]$ functions for the weighting scheme, $w(M) = b(M)$ are:

$$S_1^{\text{VL}}[w = b] = V_\mu \bar{n}_h^4 \bar{b}_N^{-4} ; \quad (\text{C5})$$

$$S_2^{\text{VL}}[w = b] = V_\mu \bar{n}_h^3 \bar{b}_N^{-3} ; \quad (\text{C6})$$

$$S_3^{\text{VL}}[w = b] = V_\mu \bar{n}_h^2 \bar{b}_N^{-2} ; \quad (\text{C7})$$

$$S_4^{\text{VL}}[w = b] = V_\mu^2 \bar{n}_h^4 \bar{b}_N^{-4} . \quad (\text{C8})$$

The $S_i^{\text{VL}}[w]$ functions for the weighting scheme, $w(M) = M$ are:

$$S_1^{\text{VL}}[w = M] = V_\mu \bar{n}_h^4 \bar{b}_M^{-4} \langle M \rangle^4 ; \quad (\text{C9})$$

$$S_2^{\text{VL}}[w = M] = V_\mu \bar{n}_h^3 \bar{b}_M^{-2} \langle M^2 \rangle \langle M \rangle^2 ; \quad (\text{C10})$$

$$S_3^{\text{VL}}[w = M] = V_\mu \bar{n}_h^2 \langle M^2 \rangle^2 ; \quad (\text{C11})$$

$$S_4^{\text{VL}}[w = M] = V_\mu^2 \bar{n}_h^4 \bar{b}_M^{-4} \langle M \rangle^4 . \quad (\text{C12})$$

2. Flux-limited surveys

The $S_i^{\text{FL}}[w]$ functions for the standard FKP weighting scheme are:

$$S_1[w = \text{FKP}] = \Omega_\mu \int_0^\infty dx x^2 \left[\frac{\bar{n}_h(x) \bar{b}_N(x)}{(1 + \alpha) + \bar{n}_h(x) P_i} \right]^4 ; \quad (\text{C13})$$

$$S_2[w = \text{FKP}] = \Omega_\mu \int_0^\infty dx x^2 \frac{\bar{n}_h^3(x) \bar{b}_N^{-2}(x)}{[(1 + \alpha) + \bar{n}_h(x) P_i]^4} ; \quad (\text{C14})$$

$$S_3[w = \text{FKP}] = \Omega_\mu \int_0^\infty dx x^2 \frac{\bar{n}_h^2(x)}{[(1 + \alpha) + \bar{n}_h(x) P_i]^4} ; \quad (\text{C15})$$

$$S_4[w = \text{FKP}] = \left\{ \Omega_\mu \int_0^\infty dx x^2 \left[\frac{\bar{n}_h(x) \bar{b}_N(x)}{(1 + \alpha) + \bar{n}_h(x) P_i} \right]^2 \right\}^2 . \quad (\text{C16})$$

The $S_i^{\text{FL}}[w]$ functions for the optimal weighting scheme derived in §99 are:

$$S_1[w = \text{OPT}] = \Omega_\mu \int_0^\infty dx x^2 \left[\frac{\bar{n}_h(\chi) \overline{b_N^2}(\chi)}{(1 + \alpha) + \bar{n}_h(\chi) \overline{b_N^2}(\chi) P_i} \right]^4 ; \quad (\text{C17})$$

$$S_2[w = \text{OPT}] = \Omega_\mu \int_0^\infty dx x^2 \frac{[\bar{n}_h(x) \overline{b_N} \bar{n}(x)]^3}{[(1 + \alpha) + \bar{n}_h(\chi) \overline{b_N^2}(\chi) P_i]^4} ; \quad (\text{C18})$$

$$S_3[w = \text{OPT}] = \Omega_\mu \int_0^\infty dx x^2 \frac{[\bar{n}_h(x) \overline{b_N^2}(\chi)]^2}{[(1 + \alpha) + \bar{n}_h(\chi) \overline{b_N^2}(\chi) P_i]^4} ; \quad (\text{C19})$$

$$S_4[w = \text{OPT}] = \left\{ \Omega_\mu \int_0^\infty dx x^2 \left[\frac{\bar{n}_h(\chi) \overline{b_N^2}(\chi)}{(1 + \alpha) + \bar{n}_h(\chi) \overline{b_N^2}(\chi) P_i} \right]^2 \right\}^2 . \quad (\text{C20})$$

The $S_i^{\text{FL}}[w]$ functions for the mass weighting plus FKP space weighting are:

$$S_1[w = M + \text{FKP}] = \Omega_\mu \int_0^\infty dx x^2 \left[\frac{\langle M(x) \rangle \bar{n}_h(\chi) \overline{b_M}(\chi)}{(1 + \alpha) + \bar{n}_h(\chi) P_i} \right]^4 ; \quad (\text{C21})$$

$$S_2[w = M + \text{FKP}] = \Omega_\mu \int_0^\infty dx x^2 \frac{\langle M(x) \rangle^2 \langle M^2(x) \rangle \bar{n}_h^3(x) \overline{b_M^2}(x)}{[(1 + \alpha) + \bar{n}_h(\chi) P_i]^4} ; \quad (\text{C22})$$

$$S_3[w = M + \text{FKP}] = \Omega_\mu \int_0^\infty dx x^2 \frac{\bar{n}_h^2(x) \langle M^2(x) \rangle^2}{[(1 + \alpha) + \bar{n}_h(\chi) P_i]^4} ; \quad (\text{C23})$$

$$S_4[w = M + \text{FKP}] = \left\{ \Omega_\mu \int_0^\infty dx x^2 \left[\frac{\langle M(x) \rangle \bar{n}_h(\chi) \overline{b_M}(\chi)}{(1 + \alpha) + \bar{n}_h(\chi) P_i} \right]^2 \right\}^2 . \quad (\text{C24})$$

Appendix D: Proof that $\overline{b_N^2} \leq \overline{b_N^2}$

To begin, the Cauchy-Schwarz inequality for two functions f and g states that:

$$|\langle f|g \rangle|^2 \leq \langle f|f \rangle \langle g|g \rangle , \quad (\text{D1})$$

where in the above we are using the following notation:

$$\langle f|g \rangle \equiv \int dx q(x) f(x) g(x) , \quad (\text{D2})$$

and with $q(x)$ an arbitrary positive definite weighting function. Thus, if we take $f = b(M)$, $g = 1$ and $q = \bar{n}(M)$, then we have the following inequality:

$$\left[\int_{M_{\text{Lim}}}^\infty dM \bar{n}(M) b(M) \right]^2 \leq \left[\int_{M_{\text{Lim}}}^\infty dM \bar{n}(M) b^2(M) \right] \left[\int_{M_{\text{Lim}}}^\infty dM \bar{n}(M) \right] . \quad (\text{D3})$$

On dividing both sides of this inequality through by \bar{n}_h^2 we arrive at the stated result.