Thinking Outside the Box: Effects of Modes Larger than the Survey on Matter Power Spectrum Covariance

Roland de Putter[1,2], Christian Wagner[1], Olga Mena[2], Licia Verde[1], Will J. Percival[3]

[1] *ICC, University of Barcelona (IEEC-UB), Marti i Franques 1, Barcelona 08028, Spain*
[2] *Instituto de Fisica Corpuscular, Universidad de Valencia-CSIC, Spain*
[3] *Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Bldg., Portsmouth, PO1 3FX, UK*

(Dated: March 28, 2012)

Accurate power spectrum (or correlation function) covariance matrices are a crucial requirement for cosmological parameter estimation from large scale structure surveys. In order to minimize reliance on computationally expensive mock catalogs, it is important to have a solid analytic understanding of the different components that make up a covariance matrix. Considering the matter power spectrum covariance matrix, it has recently been found that there is a potentially dominant effect on mildly non-linear scales due to power in modes of size equal to and larger than the survey volume. This *beat coupling* effect has been derived analytically in perturbation theory and while it has been tested with simulations, some questions remain unanswered. Moreover, there is an additional effect of these large modes, which has so far not been included in analytic studies, namely the effect on the estimated *average* density which enters the power spectrum estimate. In this article, we work out analytic, perturbation theory based expressions including both the beat coupling and this *local average effect* and we show that while, when isolated, beat coupling indeed causes large excess covariance in agreement with the literature, in a realistic scenario this is compensated almost entirely by the local average effect, leaving only $\sim 10\%$ of the excess. We test our analytic expressions by comparison to a suite of large N-body simulations, using both full simulation boxes and subboxes thereof to study cases without beat coupling, with beat coupling and with both beat coupling *and* the local average effect. For the variances, we find excellent agreement with the analytic expressions for $k < 0.2h\mathrm{Mpc}^{-1}$ at $z = 0.5$, while the correlation coefficients agree to beyond $k = 0.4h\mathrm{Mpc}^{-1}$. As expected, the range of agreement increases towards higher redshift and decreases slightly towards $z = 0$. We finish by including the large-mode effects in a full covariance matrix description for arbitrary survey geometry and confirming its validity using simulations. This may be useful as a stepping stone towards building an actual galaxy (or other tracer's) power spectrum covariance matrix.

I. INTRODUCTION

Galaxy surveys (and surveys of other dark matter tracers) are an important tool for constraining cosmological parameters and the tracer's power spectrum (or its Fourier transform, the correlation function) is the most valuable observable. It can be used to measure the scale of baryon acoustic oscillations (BAO, see e.g. [1–4]), which is a particularly robust feature and a strong probe of dark energy through the measurement's dependence on the expansion history, but it is also common to use the full shape of the power spectrum to constrain cosmology (e.g. [5, 6]). In order for these measurements to be useful for cosmology however, it is crucial to have an accurate estimate of the observable's covariance matrix. As surveys get larger, calculating such a matrix directly from simulations becomes ever more challenging. In fact, large numbers of mock catalogs are required to estimate covariance matrices directly from simulations and it is computationally very costly to create such numbers with the required volume and resolution. Moreover, the covariance matrix is cosmology dependent so ideally one would either be able to generate a large number of matrices for different cosmologies, or show that this cosmology dependence can be safely neglected. This requirement increases the number of mocks needed in the brute force method even further. It is therefore important to develop a good analytical understanding in order to eventually rely less on simulations.

While the density of tracers may have a non-trivial relation with the dark matter density, a good first step towards understanding the true covariance matrix is to study that of the dark matter power spectrum. The two main complications in understanding the dark matter matrix are the mode mixing due to the finite survey volume and the non-Gaussian nature of the matter overdensity field on small scales ($k \gtrsim 0.1h\mathrm{Mpc}^{-1}$) due to non-linear evolution. While the former is relatively straightforward to quantify ([7]), the latter effects are the subject of a large number of studies [8–20], involving N-body simulations, perturbation theory and the halo model. For covariances on non-linear scales, an important (but often overlooked) role is played by modes of wavelength comparable to and larger than the survey size. In particular, [21, 22] found that the interplay of mode mixing by the survey window with non-linear correlations between a pair of small-scale modes and a large mode, causes a covariance contribution proportional to the power in super-survey modes, and that, in fact, this is the dominant contribution to the covariance matrix on small scales. The added covariance affects both the variances and the off-diagonal covariances. This effect is usually

called *beat coupling* (see also [13, 23–25]). A second consequence of the presence of modes larger than the survey is that the average density (which is needed to construct the overdensity) estimated from the survey has a variance itself, leading to an additional covariance contribution proportional to the large scale power. We will refer to this as the *local average* effect.

In the standard approach to estimating covariances from N-body simulations, the power spectrum is calculated using the full, periodic simulation volume. However, this completely misses the large effects discussed above because there are no modes larger than the "survey" volume. In order to capture these effects, one thus needs to consider either a subvolume of a (much) larger simulation, or implement a varying zero mode ([12]) in the simulations. While some such studies have been performed (see e.g. [24, 25]), a systematic study, leading to a complete and well tested analytic description has until now been lacking.

In this article, we aim to present such a study. We use a set of large ($L = 2400h^{-1}$Mpc) N-body simulations and study the aforementioned effects of large modes by considering significantly smaller subboxes ($L = 600h^{-1}$Mpc), thus imitating the real world scenario of a finite survey embedded in an infinite universe. We build an analytic model for the covariance matrix, properly including the effect of the window function and describing non-linear effects using perturbation theory. This description includes the beat coupling model from [21], but adds to this an analytic estimate of the local average effect. The latter effect needs to be included when building a covariance matrix for a realistically estimated power spectrum and has to our knowledge not been calculated analytically previously.

We compare our model to simulations at five redshifts in the range $z = 0 - 2$ and generally find excellent agreement until well into the non-linear regime. Since we test several combinations of covariance matrix contributions individually by employing different approaches to estimating spectra from simulations, we are confident that our theoretical understanding of the dark matter covariance is correct and that we have captured the main effects.

One of our main results is that, while beat coupling indeed causes excess covariance which soon dominates, the contribution from the local average effect cancels most of this out, leaving only a reduced beat coupling contribution of about 10% of the original.

In addition to the effects on non-linear scales discussed above, the mode mixing due to the window function also has a well known effect on linear scales, correlating neighboring power spectrum estimators and reducing the variances. This effect needs to be taken into account for a realistic survey. We calculate these covariances directly using analytic expressions from [7] and show good agreement with simulations. Thus, while the main goal of this paper is to study the beat coupling and local average effects related to super-survey modes, we also provide a complete expression that can be used to calculate a (dark matter) covariance matrix for an arbitrary survey geometry and can serve as a stepping stone for calculating a galaxy (or other dark matter tracer) spectrum covariance matrix. We also study the cosmology dependence of the covariance matrix and find that for cosmological parameter variations relevant to current cosmological constraints, the covariance matrix undergoes changes at the $\sim 30\%$ level.

The outline of this article is as follows. We explain the analytic description of the covariance matrix in section II, using a simple formalism apt for the description of simulation results. This will lead to three expressions that can be tested against simulations (without beat coupling, with beat coupling, and with beat coupling *and* local average effect). We then describe our N-body simulations in section III and compare the results to theory in section IV. While for the main results in this paper, we stick to a simple description in terms of a discrete set of modes, we discuss a more rigorous and complete treatment of the window function in section V, leading to expressions that can be applied to arbitrary survey geometry. We conclude and summarize our work in section VI. Finally, we use our analytic expressions to briefly discuss the cosmology dependence of the covariance matrix in the Appendix.

As this article is rather long and technical, we provide some recommendations for a reader on a tight schedule. On a first reading, this reader may want to skip the mathematical discussion in section II and jump straight to section II E, where the main equations are summarized. Moreover, in section V, one could jump to Eq. (48), which presents the final covariance matrix expression for general geometry, and then focus on the results in Figs 6-9, where the green points and lines show the covariance matrix corresponding to said equation.

II. COVARIANCE MATRIX FORMALISM

In this section, we work out the formalism for calculating a covariance matrix for a power spectrum measured from a cubic volume with uniform selection function, i.e. the geometry of N-body simulations. We describe the density field in terms of a discrete set of modes, which significantly simplifies notation. We will provide a more general description, which can be applied to an arbitrary survey geometry (following [7]), in section V. The more technical discussion there also serves to motivate the simpler approach taken in this section.

A. N-body Simulation Geometry (Cubic Box)

We consider a cubic subvolume $V = L^3$ of the universe (which later in this article will be modeled by taking a subvolume of a much larger periodic simulation box). The matter overdensity field $\delta(\mathbf{x}) \equiv (\rho(\mathbf{x}) - \bar{\rho})/\bar{\rho}$, where $\rho(\mathbf{x})$ is the matter density and $\bar{\rho}$ its mean, can be expanded as

$$\delta(\mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\, e^{-i\mathbf{k}\cdot\mathbf{x}}\, \delta(\mathbf{k}). \tag{1}$$

To significantly simplify notation, we choose to describe the density field in the volume V in terms of discrete Fourier modes in analogy to the usual description of the density field in a periodic box, i.e. we effectively describe the field in the subvolume by

$$\delta(\mathbf{x}) \sim \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}}\, \delta_{\mathbf{k}}, \quad k_i = \frac{2\pi}{L}\, n_i \quad \text{for} \quad n_i = 0, \pm 1, \ldots. \tag{2}$$

To be more exact, we define the discrete modes as

$$\delta_{\mathbf{k}} \equiv \frac{1}{V} \int_V d^3\mathbf{x}\, e^{i\mathbf{k}\cdot\mathbf{x}}\, \delta(\mathbf{x}). \tag{3}$$

In other words, $\delta_{\mathbf{k}}$ is equal to the FKP estimator ([7]) $F(\mathbf{k})$ (up to a factor $V^{-\frac{1}{2}}$) applied to the volume V, and is therefore a weighted average over a range of continuum Fourier modes. In particular, while there is no true $\mathbf{k} = \mathbf{0}$ mode, the subvolume sees an effective zero mode (see also [12]) $\delta_{\mathbf{0}} = V^{-1} \int_V d^3\mathbf{x}\, \delta(\mathbf{x})$ which gets its main contribution from continuum Fourier modes $|\mathbf{k}| < 2\pi/L$. Note that Eq. (2) is not to be taken literally, as the field is not actually periodic with respect to the volume V. We discuss the effect of taking a subvolume more rigorously in section V.

The statistics of the discrete overdensity modes can be characterized by a sequence of connected $n-$point functions,

$$\begin{aligned}
\langle \delta_{\mathbf{k}}\, \delta_{\mathbf{k}'} \rangle &= \frac{P_k}{V}\, \delta^{\mathrm{K}}_{\mathbf{k}+\mathbf{k}'} \\
\langle \delta_{\mathbf{k}}\, \delta_{\mathbf{k}'}\, \delta_{\mathbf{k}''} \rangle_c &= \frac{B_{\mathbf{k},\mathbf{k}',\mathbf{k}''}}{V^2}\, \delta^{\mathrm{K}}_{\mathbf{k}+\mathbf{k}'+\mathbf{k}''} \\
\langle \delta_{\mathbf{k}}\, \delta_{\mathbf{k}'}\, \delta_{\mathbf{k}''}\, \delta_{\mathbf{k}'''} \rangle_c &= \frac{T_{\mathbf{k},\mathbf{k}',\mathbf{k}'',\mathbf{k}'''}}{V^3}\, \delta^{\mathrm{K}}_{\mathbf{k}+\mathbf{k}'+\mathbf{k}''+\mathbf{k}'''} \\
&\cdots\ ,
\end{aligned}$$

where we have defined the power spectrum P_k, bispectrum $B_{\mathbf{k},\mathbf{k}',\mathbf{k}''}$ an the trispectrum $T_{\mathbf{k},\mathbf{k}',\mathbf{k}'',\mathbf{k}'''}$. Note that these are technically weighted averages of the true, continuum power, bi- and trispectrum, for example (see also section V)

$$P_0 \equiv V\, \langle |\delta_{\mathbf{0}}|^2 \rangle = V \int \frac{d^3\mathbf{k}}{(2\pi)^3}\, P(k) \prod_{i=x,y,x} j_0^2(Lk_i/2), \tag{4}$$

where $j_0(x) = \sin x/x$ is the zeroth spherical Bessel function.

The power of an individual mode can be estimated as

$$\hat{P}_{\mathbf{k}} \equiv V\, |\delta_{\mathbf{k}}|^2, \tag{5}$$

such that $\langle \hat{P}_{\mathbf{k}} \rangle = P_k$ as desired, from which one can define a bin averaged estimator in order to maximize signal to noise,

$$\hat{P}_i \equiv \frac{1}{N_i} \sum_{\mathbf{k}\in i} \hat{P}_{\mathbf{k}}. \tag{6}$$

Here, the sum is over all $\mathbf{k} = \frac{2\pi}{L}\, \mathbf{n}$, with integer components n_i, such that k lies in some small range defining the i-th bin. N_i is the number of modes in i. The expectation value $P_i \equiv \langle \hat{P}_i \rangle \approx P_{k_i}$, with k_i a typical mode inside the bin.

We are interested in the covariance matrix of this bin averaged estimator,

$$\mathbf{C}_{ij} \equiv \langle \delta\hat{P}_i\, \delta\hat{P}_j \rangle. \tag{7}$$

Starting from the covariances in the individual mode estimators,

$$\langle \delta\hat{P}_{\mathbf{k}}\, \delta\hat{P}_{\mathbf{k}'}\rangle = P_k^2\, \left(\delta^{\rm K}_{\mathbf{k}+\mathbf{k}'} + \delta^{\rm K}_{\mathbf{k}-\mathbf{k}'}\right) + \frac{1}{V}\, T_{\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}'}, \tag{8}$$

we obtain

$$\mathbf{C}_{ij} = \frac{1}{N_i}\sum_{\mathbf{k}\in i}\frac{1}{N_j}\sum_{\mathbf{k}'\in j}\langle \delta\hat{P}_{\mathbf{k}}\, \delta\hat{P}_{\mathbf{k}'}\rangle = \frac{2P_i^2}{N_i}\,\delta^{\rm K}_{ij} + \frac{\bar{T}_{ij}}{V}, \tag{9}$$

where we have introduced the shorthand notation $\bar{T}_{ij}$ for the bin averaged trispectrum. The term proportional to the power spectrum squared is diagonal and given by simply counting modes because the binning in k-space only includes pairs of estimators that have zero covariance. As we will discuss in detail in section V, this is not in general the case as the window function may correlate different bins.

B. N-body Simulation Geometry: Perturbation Theory

The power spectrum appearing in the first term on the right hand side of Eq. (9) is the full non-linear power spectrum, which can be estimated directly from simulations. The trispectrum is harder to get from simulations (but see [19]), but can be modeled using a variety of perturbation theory schemes or using the halo model. In this work, we will use the simple framework of (Eulerian) standard perturbation theory (SPT; [8, 26–28]). Anticipating angle averaging, i.e. modulo transformations $\mathbf{k}\leftrightarrow-\mathbf{k}$ and $\mathbf{k}'\leftrightarrow-\mathbf{k}'$, the leading order trispectrum for the configuration of interest is given as (see[1] [8, 27])

$$\begin{aligned} T_{\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}'} \quad &= \quad 12\, P_k^{\rm lin}\, P_{k'}^{\rm lin}\, \left[F_3(\mathbf{k},-\mathbf{k},\mathbf{k}')P_k^{\rm lin} + (\mathbf{k}\leftrightarrow\mathbf{k}')\right] \\ +8 \quad P^{\rm lin}_{|\mathbf{k}-\mathbf{k}'|} \quad &\left[F_2(\mathbf{k}-\mathbf{k}',\mathbf{k}')\, P_{k'}^{\rm lin} + (\mathbf{k}\leftrightarrow\mathbf{k}')\right]^2 \\ +16 \quad P_k^{\rm lin} \quad &P_{k'}^{\rm lin}\, P_0^{\rm lin}\, F_2(-\mathbf{0},\mathbf{k})\, F_2(\mathbf{k}',\mathbf{0}), \end{aligned} \tag{10}$$

where F_2 and F_3 are the kernels for the second and third order contributions to the density field, as given in the appendix of [28]. We can now distinguish two types of terms. The first two lines represent the "standard" trispectrum contributions, that have always been included in PT studies of the trispectrum (see e.g. [8]). When no modes larger than the survey are present, like in the artificial case of a power spectrum estimated from the full volume of a periodic (N-body simulation) box, these are the only terms entering the trispectrum, giving

$$\bar{T}^0_{ij} = \frac{1}{N_i}\sum_{\mathbf{k}\in i}\frac{1}{N_j}\sum_{\mathbf{k}'\in j}[P_k^{\rm lin}\, P_{k'}^{\rm lin}\, \left[F_3(\mathbf{k},-\mathbf{k},\mathbf{k}')P_k^{\rm lin} + (\mathbf{k}\leftrightarrow\mathbf{k}')\right] + 8\, P^{\rm lin}_{|\mathbf{k}-\mathbf{k}'|}\, \left[F_2(\mathbf{k}-\mathbf{k}',\mathbf{k}')\, P_{k'}^{\rm lin} + (\mathbf{k}\leftrightarrow\mathbf{k}')\right]^2]. \tag{11}$$

We will evaluate $\bar{T}^0_{ij}$ numerically as no further analytic simplifications are possible.

C. The Beat Coupling Effect

The third line in Eq. (10) is the beat coupling contribution, the importance of which has been realized only more recently ([13, 21–25]). Our notation here requires some explanation (see also section V). Consistent with our discussion in the previous section, quantities evaluated at $\mathbf{k}=\mathbf{0}$ should really be interpreted in terms of an *effective* zero mode arising from contributions at $k<2\pi/L$, i.e. $P_0^{\rm lin} = V\,\langle|\delta_{\mathbf{0}}^{\rm lin}|^2\rangle \approx P^{\rm lin}(k\sim\pi/L)$, generated by continuum contributions $T(\mathbf{k}+\epsilon,-\mathbf{k}+\epsilon',\mathbf{k}'+\epsilon'',-\mathbf{k}'+\epsilon''')$ with $\epsilon+\epsilon' = -(\epsilon''+\epsilon''')\sim\pi/L$. Moreover, F_2, given by

$$F_2(\mathbf{k_1},\mathbf{k_2}) = \frac{5}{7} + \frac{\hat{\mathbf{k}}_{\mathbf{1}}\cdot\hat{\mathbf{k}}_{\mathbf{2}}}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{2}{7}\,(\hat{\mathbf{k}}_{\mathbf{1}}\cdot\hat{\mathbf{k}}_{\mathbf{2}})^2, \tag{12}$$

is at first sight not well defined if one of the arguments equals $\mathbf{0}$. $F_2(-\mathbf{0},\mathbf{k})$ (and the other affected term) should thus be interpreted as a limit

$$F_2(-\mathbf{0},\mathbf{k}) = \int\frac{d\Omega_{\hat{\epsilon}}}{4\pi}F_2(-\epsilon,\mathbf{k}), \tag{13}$$

[1] Note that the factor 16 in Eq. (7) of [8] should be a factor 8.

with $|\epsilon| \sim \pi/L$ (see section V for a more rigorous treatment of the beat coupling term with the same result). The integral over the direction of ϵ makes this quantity well defined.

Using the fact that

$$\int \frac{d\Omega_{\mathbf{k_1}}}{4\pi} F_2(\mathbf{k_1}, \mathbf{k_2}) = \frac{17}{21}, \tag{14}$$

the beat coupling term can be angle (or bin) averaged analytically so that we end up with

$$\bar{T}_{ij} = \bar{T}^0_{ij} + 16 \left(\frac{17}{21}\right)^2 P^{\rm lin}_{k_i} P^{\rm lin}_{k_j} P^{\rm lin}_0. \tag{15}$$

Physically, the presence of the beat coupling term is an interesting interplay of mode mixing due to the window functions with correlations between pairs of non-linear modes and one larger mode (see [21]). Due to the finite volume from which the power spectrum is measured, the estimator $\hat{P}(\mathbf{k})$ really consists of a weighted average of pairs of density modes $\delta(\mathbf{k}+\epsilon)\,\delta(-(\mathbf{k}+\epsilon'))$, with $|\epsilon|$ of order of the fundamental mode $2\pi/L$. For k in the non-linear regime, such pairs correlate with the large scale perturbation $\delta(\epsilon'-\epsilon)$, which in turn causes the covariance between the power spectrum estimators to be proportional to the power in these large modes. In our description, this is captured by the power in the effective zero mode. Note that the beat coupling creates an excess both in the variance and in the off-diagonal elements of the covariance matrix.

D. The Local Average Effect

While the beat coupling term derived in the previous section adds significant covariance on non-linear scales, there is an additional effect coupling small scale covariance to power in the zero-mode that plays a role in the power spectrum estimated from an actual survey. This second effect is caused by the fact that to obtain the overdensity δ, one needs an estimate of the average density $\bar{\rho}$. In a realistic survey, one does not know the true average number density of, say, galaxies, but instead has to rely on an estimate of the average density within the survey volume, which is modulated by the zero mode $\delta_{\mathbf{0}}$. This results in a decrease in covariance, partially canceling the beat coupling effect[2]. We will refer to this contribution as the *local average effect*.

The local average effect causes the true overdensity estimator to be given by

$$\tilde{\delta}_{\mathbf{k}} \equiv \frac{\delta_{\mathbf{k}}}{1+\delta_{\mathbf{0}}}. \tag{16}$$

Since $\bar{\rho}$ also appears in the numerator of δ, there technically is also a $\delta_{\mathbf{0}}$ contribution there. However, the Fourier transform of the zero mode for non-zero wave vector $\mathbf{k}$ vanishes so this term can be omitted. Eq. (16) can be expanded in powers of $\delta_{\mathbf{0}}$ to derive expressions for the n-point functions to the desired order in perturbation theory. For example, to next-to-leading order, the expectation value of the power spectrum estimator becomes

$$\begin{aligned}
\langle \hat{\tilde{P}}_{\mathbf{k}} \rangle &\equiv \langle \tilde{\delta}_{\mathbf{k}}\, \tilde{\delta}_{-\mathbf{k}} \rangle \\
&= V \left\langle \delta_{\mathbf{k}}\, \delta_{-\mathbf{k}} \left(1 - 2\delta_{\mathbf{0}} + 3\delta^2_{\mathbf{0}} + \mathcal{O}\left(\delta^3_{\mathbf{0}}\right)\right)\right\rangle \\
&= P_k - 2\frac{B_{\mathbf{k},-\mathbf{k},\mathbf{0}}}{V} + 3\frac{P_{\mathbf{0}}}{V} P_k + \mathcal{O}\left(P^3/V^2\right) \ .
\end{aligned} \tag{17}$$

The expectation value of the angle-averaged estimator is then

$$\begin{aligned}
\langle \hat{\tilde{P}}_i \rangle &= P_{k_i} - \frac{2}{V}\frac{68}{21} P_{k_i} P_0 + \frac{3}{V} P_{k_i} P_0 + \mathcal{O}\left(P^3/V^2\right) \\
&= P_{k_i} \left(1 - \frac{73}{21}\frac{P_0}{V}\right) + \mathcal{O}\left(P^3/V^2\right) \ .
\end{aligned} \tag{18}$$

Hence, the power spectrum receives a small bias due to the local average effect. The relative correction is $\lesssim 10^{-4}$ for a $1h^{-3}\mathrm{Gpc}^3$ survey so it can be safely ignored for a realistic survey. Here, we have used that the bispectrum ([13])

$$B_{\mathbf{k},\mathbf{k}',\mathbf{k}''} = 2\, P_k\, P_{k'}\, F_2(\mathbf{k}, \mathbf{k}') \ + \ (\text{cyclic}) \tag{19}$$

[2] The difference between using the true and the local average was also commented on briefly in [29]

and we have applied the identity (14) to carry out the angle averaging.

Using the same approach for the covariance matrix, one gets to next to leading order

$$\tilde{\mathbf{C}}_{ij} = \langle \hat{\tilde{P}}_i \, \hat{\tilde{P}}_j \rangle - \langle \hat{\tilde{P}}_i \rangle \, \langle \hat{\tilde{P}}_j \rangle = \frac{2P_{k_i}^2}{N_i} \, \delta_{ij}^{\mathrm{K}} + \frac{\bar{T}_{ij}^0}{V} + 16 \left(\frac{17}{21}\right)^2 P_{k_i}^{\mathrm{lin}} \, P_{k_j}^{\mathrm{lin}} \, \frac{P_0^{\mathrm{lin}}}{V} - \frac{188}{21} \, P_{k_i}^{\mathrm{lin}} \, P_{k_j}^{\mathrm{lin}} \, \frac{P_0^{\mathrm{lin}}}{V}, \tag{20}$$

where we have ignored a small relative correction to the diagonal term of order P_0^{lin}/V. The local average effect thus introduces a term of the same form as the beat coupling term from the previous section, proportional to $P_{k_i} \, P_{k_j} \, P_0/V$, but with opposite sign. Comparing the coefficients, one finds that the two effects almost entirely cancel out, leaving only a small positive coefficient, $16 \left(\frac{17}{21}\right)^2 - \frac{188}{21} \approx 1.5 \sim 10\%$ of the original beat coupling coefficient. Note that the local average effect to this order consists of contributions from the bispectrum and trispectrum, but also from terms that would even be there had the field remained completely Gaussian (but that happen to be of the same order as the leading non-Gaussian corrections).

Eq. (20) thus gives us an expression for the most realistic case we will consider, where modes larger than the survey are present *and* the spectrum is estimated using the local average in the survey volume.

E. Theory Summary and Outlook

In the remainder of this article, we will use N-body simulations to test the expressions derived above. We will consider three different cases which allow us to separately constrain different combinations of contributions to the total covariance:

- **Case 1:** *periodic box*
  The spectrum is estimated from the full, periodic simulation volume. In this case, there is no beat coupling nor a local average effect and the prediction for the covariance is to leading order

  $$\mathbf{C}_{ij} = \frac{2P_{k_i}^2}{N_i} \, \delta_{ij}^{\mathrm{K}} + \frac{\bar{T}_{ij}^0}{V}. \tag{21}$$

  We obtain the (non-linear) power spectrum appearing above by applying the Halofit[3] prescription [30] to the linear power spectrum calculated using CAMB [31]. The bin-averaged trispectrum $\bar{T}_{ij}^0$ is given by Eq. (11), where we make one modification. Instead of using the linear power spectra, we use the non-linear spectra. This is consistent to the desired order in perturbation theory and turns out to slightly improve the accuracy of the model on strongly non-linear scales.

- **Case 2:** *subbox of periodic box*
  Since modes larger than the "survey volume" are now present, there is a beat coupling effect. The leading order covariance prediction is

  $$\mathbf{C}_{ij} = \frac{2P_{k_i}^2}{N_i} \, \delta_{ij}^{\mathrm{K}} + \frac{\bar{T}_{ij}^0}{V} + 16 \left(\frac{17}{21}\right)^2 P_{k_i} \, P_{k_j} \, \frac{P_0}{V}. \tag{22}$$

  Note that also for the additional trispectrum terms, we choose to use the non-linear power spectrum as opposed to the linear one.

- **Case 3:** *subbox of periodic box, using subbox mean*
  This is the most realistic case, where not only there are modes larger than the survey, but the overdensity $\delta = (\rho(\mathbf{x}) - \bar{\rho})/\bar{\rho}$ is calculated in terms of the average density $\bar{\rho}$ of the subbox, as opposed to the "true" average density of the full box. The leading order part of the covariance matrix is given by

  $$\tilde{\mathbf{C}}_{ij} = \frac{2P_{k_i}^2}{N_i} \, \delta_{ij}^{\mathrm{K}} + \frac{\bar{T}_{ij}^0}{V} + \frac{676}{441} \, P_{k_i} \, P_{k_j} \, \frac{P_0}{V}. \tag{23}$$

In the next section, we first describe the details of our simulations. We will discuss the results of our comparison in section IV.

---

[3] We have checked that the difference between using the Halofit spectrum and the average simulated spectrum is small for $k < 0.4h/\mathrm{Mpc}$.

III. N-BODY SIMULATIONS

In order to test our analytic predictions we use a large suite of N-body simulations. We have two sets of simulations consisting of 160 runs of a $2400h^{-1}$Mpc box with 768^3 particles and 1024 runs of a $600h^{-1}$Mpc box with 192^3 particles. The initial conditions of the 160 and 1024 simulations were set up using different realizations of a Gaussian random field with the power spectrum given by CAMB. We adopted a flat ΛCDM cosmology consistent with the current observational constraints [32]. The cosmological parameters are the present-day matter fraction $\Omega_m = 0.27$, Hubble constant $h = 0.7$, baryon fraction $\Omega_b h^2 = 0.023$, spectral index $n_s = 0.95$, and present-day normalization $\sigma_8 = 0.7913$.

The particles were displaced from their initial grid points according to second-order Lagrangian perturbation theory using an initial redshift $z_i = 19$.

The simulations were performed with the Tree-PM code Gadget-2 [33] taking only the gravitational force into account. We applied a force softening of $70h^{-1}$kpc and used a particle mesh of 2048^3 and 512^3 for the $2400h^{-1}$Mpc and $600h^{-1}$Mpc runs, respectively. Using a much higher resolution simulation, we checked that with these simulation settings the power spectrum derived from the simulation data is accurate at the 1% level up to $k < 0.2h\text{Mpc}^{-1}$ and remains accurate within 4% up to $k < 0.4h\text{Mpc}^{-1}$ at all redshifts.

To compute the power spectrum from the simulation data, we assign the particles using the cloud-in-cell (CIC) scheme to a regular grid with a fixed grid spacing of $1.5625h^{-1}$Mpc in all cases (full box, subbox, subbox with zero-padding). Hence, the Nyquist frequency of the grid is the same in all cases, $k_{\rm Ny} \approx 2h^{-1}$Mpc, which is 5 times larger than the scales we consider in this paper $k < 0.4h\text{Mpc}^{-1}$. Therefore we expect the effects due to finite grid size (smearing and aliasing effects) to be very small. Nevertheless, we do correct for the smearing due to the CIC assignment [34].

The simulations and the power spectrum computation needed about 300,000 CPU hours on our in-house cluster Hipatia at ICCUB.

IV. RESULTS

We now compare the simulation results to the analytic predictions summarized in section II E. For the covariances of the full periodic box (Case 1), we compare to the 1024 simulations of the $600h^{-1}$Mpc cubed volume and to the 160 simulations of the $2400h^{-1}$Mpc cubed volume. To study the *subbox* Cases 2 & 3, we divide the $L = 2400h^{-1}$Mpc volume into 64 smaller boxes with side $L/4 = 600h^{-1}$Mpc each, which provides us with $160 \times 64 = 10,240$ different subbox realizations.

For each volume, we estimate the power spectrum $\hat{P}_i$ in isotropic bins of width $\Delta k = 0.01h\text{Mpc}^{-1}$ in the range $k = 0 - 0.4h\text{Mpc}^{-1}$. We then estimate the covariance matrix for the full box, or for a particular subbox of it, by

$$\hat{\mathbf{C}}_{ij} = \frac{1}{N_r - 1} \sum_r (\hat{P}_i - \langle \hat{P}_i \rangle)(\hat{P}_j - \langle \hat{P}_j \rangle), \tag{24}$$

where the sum runs over all N_r simulation realizations and $\langle \hat{P}_i \rangle$ is the average power spectrum over simulations. In the subbox case, we then improve the accuracy by averaging the subbox covariance estimate in Eq. (24) over all 64 subboxes. Note that, since the subboxes living in the same simulation volume are not independent, it would be wrong to directly apply Eq. (24) to all $N_r = 10,240$ subboxes.

Since the non-linear effects discussed in section II become stronger with time, it is interesting to consider the covariance matrix for a range of redshifts from $z = 2 - 0$. We expect our analytic predictions to have the largest range of validity at $z = 2$, as non-linear effects there are smallest.

In Fig. 1, we show results for the variances in the power spectrum, normalized by the variance based on mode counting and the linear power spectrum. Case 1 is tested by both the full $L = 2400h^{-1}$Mpc box case (black dots - N-body results, black dashed - analytic) and the full $L = 600h^{-1}$Mpc box case (red). The blue dots and dashed lines test Case 2 and the green ones test Case 3. For all cases and at all redshifts, we find good agreement between theory and simulation for bins up to a maximum value $k_{\rm max}$ that lies in the non-linear regime. The values for $k_{\rm max}$ are reasonable given the range of scales over which perturbation theory is expected to be applicable, $k_{\rm max} \sim 0.4h\text{Mpc}^{-1}$ at $z = 2$ to $k_{\rm max} \sim 0.15 - 0.2h\text{Mpc}^{-1}$ at $z = 0$.

We now comment on the three scenarios of interest individually.

- **Case 1:** *periodic box*
  The $L = 600h^{-1}$Mpc and $L = 2400h^{-1}$Mpc full-box simulation results agree very well with each other, as expected because they are both described by Case 1 and have no beat coupling. They also agree well with the analytic prediction. This confirms that excess covariance is not simply caused by the mere presence of large

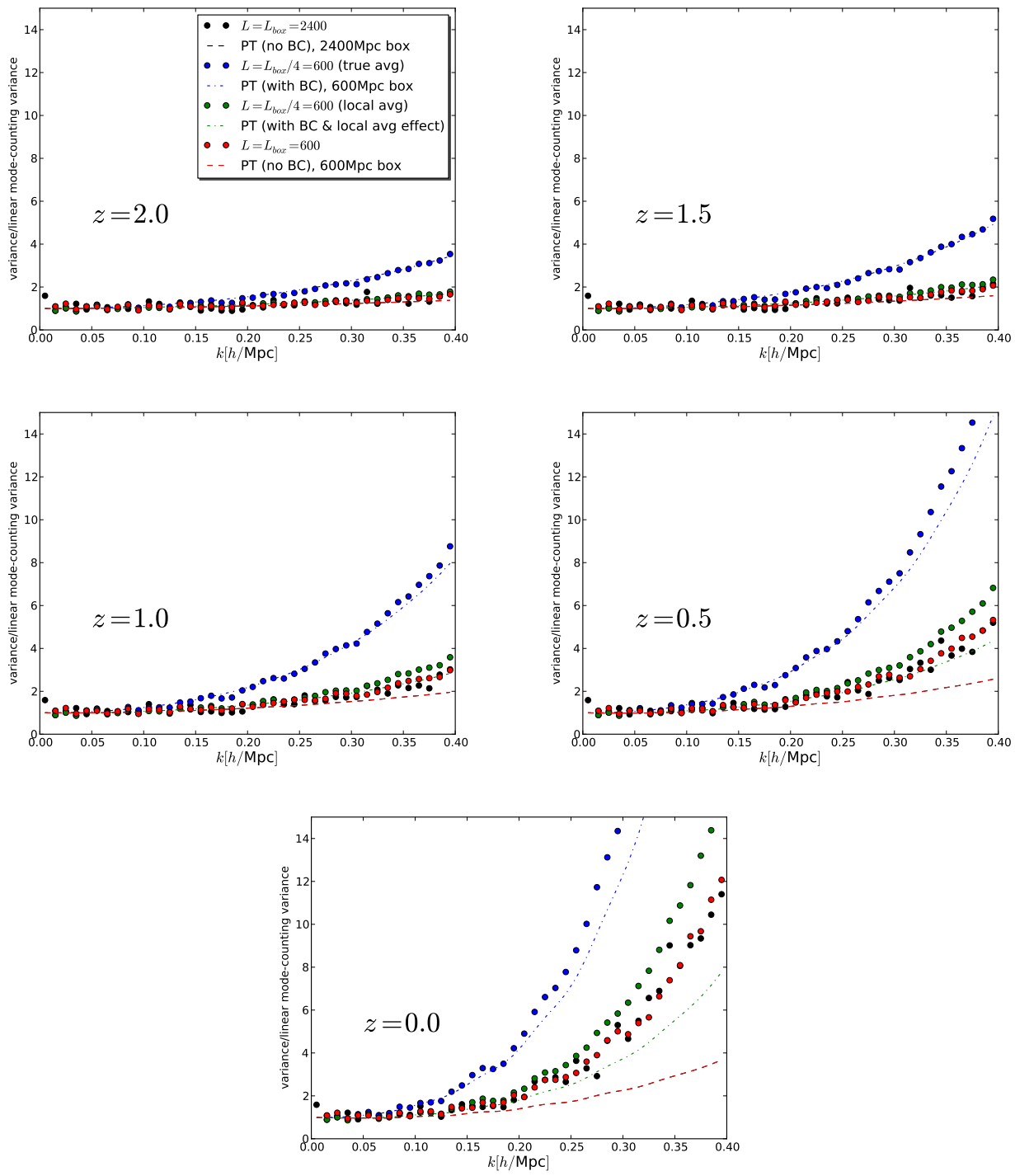

FIG. 1: Power spectrum variance relative to variance based on linear spectrum and mode counting ($2P_{\rm lin}^2/N_k$ with N_k the number of modes per bin) for redshifts $z = 0 - 2$. The dots are results from N-body simulations and the (dashed) lines theory predictions from Eqs (21)-(23), using perturbation theory (PT). The results in black and red represent the case where no modes larger than the survey are present (but other non-linear contributions are included). The blue results represent the case where these modes *are* present and thus display large excess covariance due to beat coupling. In the most realistic case (green), the large modes also affect the power spectrum through the estimated average density and this *local average effect* reduces the excess covariance by $\sim 90\%$.

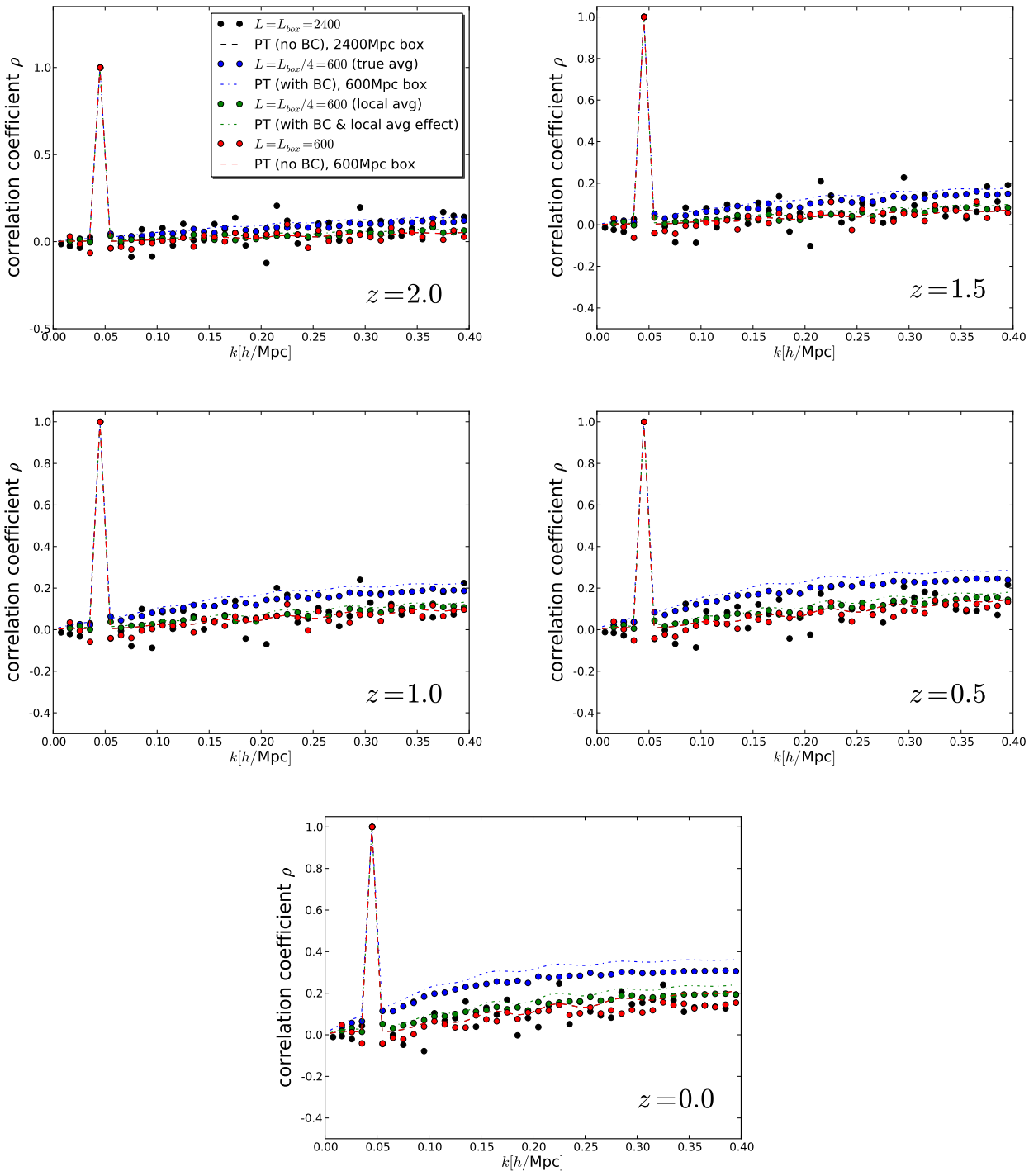

FIG. 2: Correlation coefficients $\rho_{ij} \equiv \mathbf{C}_{ij}/\sqrt{\mathbf{C}_{ii}\,\mathbf{C}_{jj}}$ of i-th bin at k on x-axis, relative to j-th bin at $k = 0.04 - 0.05h\mathrm{Mpc}^{-1}$. Same comparisons and color coding as in Fig. 1. The beat coupling causes a significant increase in the off diagonal correlations (blue), but this is again undone by the local average effect (green). The effect is described well by Eqs (21)-(23).

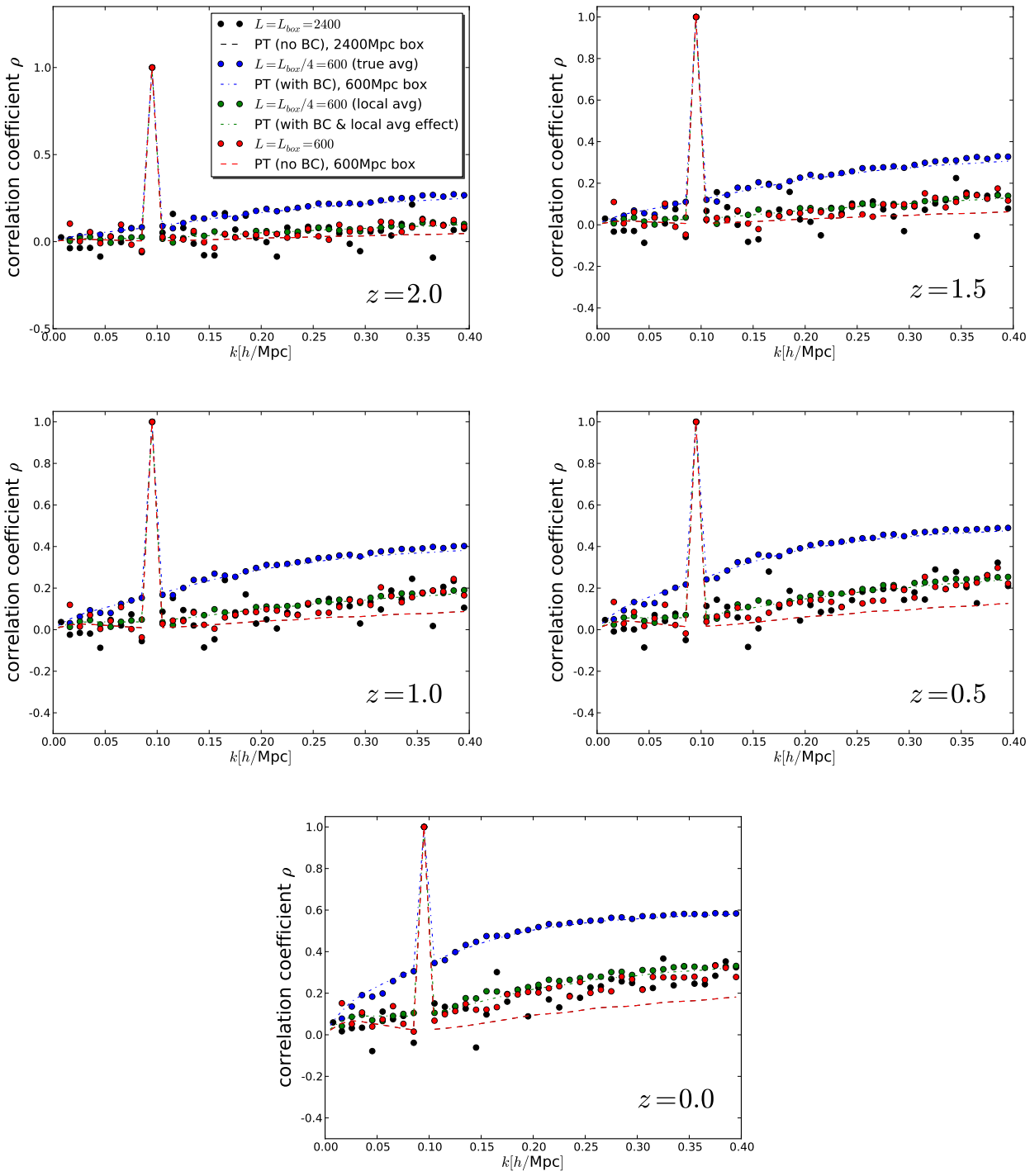

FIG. 3: As in Fig. 2, but here correlation coefficients relative to bin at $k = 0.09 - 0.1h\mathrm{Mpc}^{-1}$ are shown.

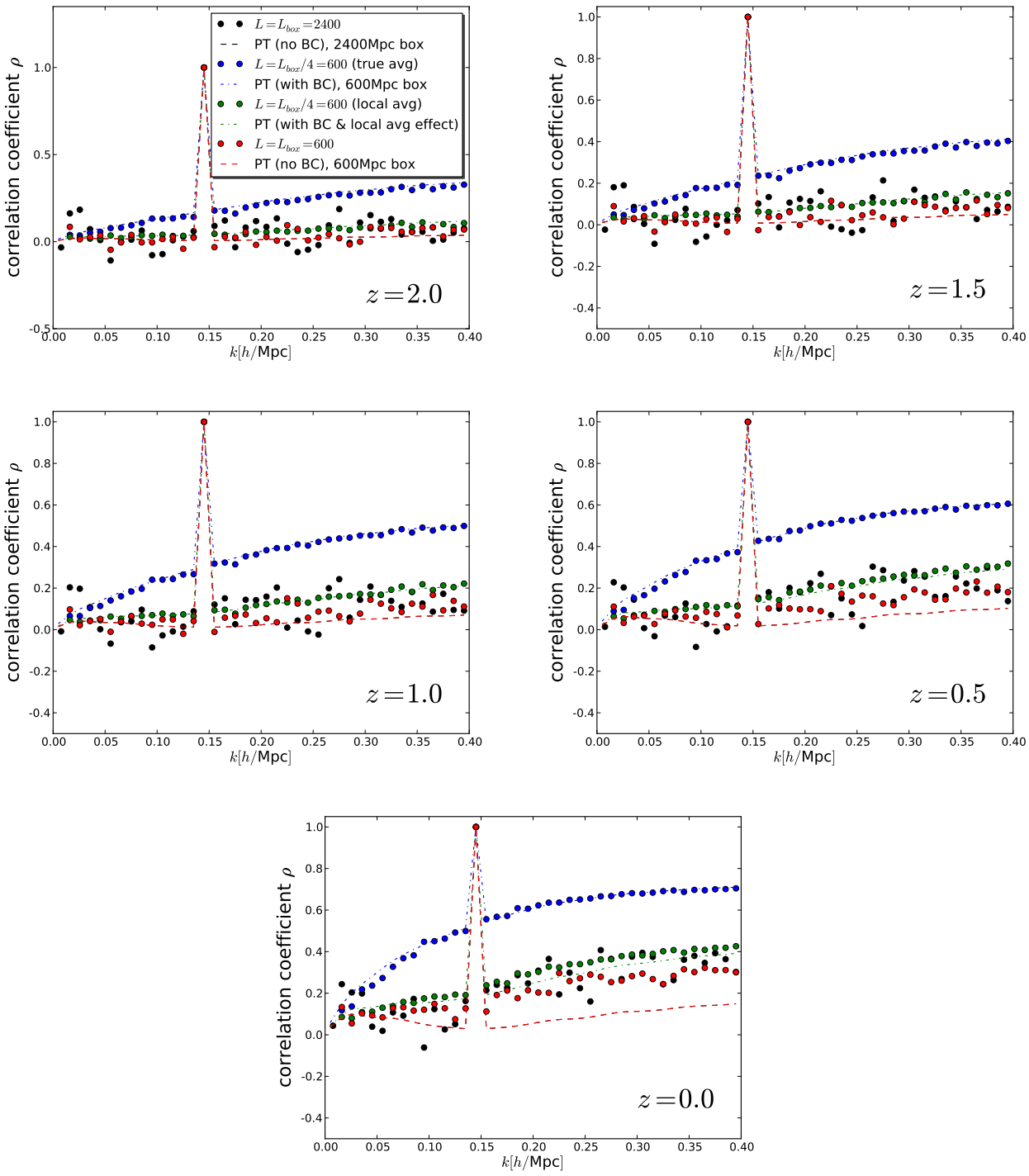

FIG. 4: As in Fig. 2, but here correlation coefficients relative to bin at $k = 0.14 - 0.15h\mathrm{Mpc}^{-1}$ are shown.

modes. The following two cases show that what matters is the presence of modes larger than the "survey" volume.

- **Case 2:** *subbox of periodic box*
  Beat coupling introduces a large amount of excess variance and is well described by perturbation theory (for the range of scales discussed above).

- **Case 3:** *subbox of periodic box, using subbox mean*
  As expected from theory, the local average effect undoes most of the excess variance due to beat coupling. Given the noise in the simulation results, the Case 3 simulation results do not distinguish with much significance the Case 3 theory prediction from the Case 1 prediction (but certainly from the Case 2 prediction). However, the Case 3 simulation results do clearly display a larger variance than the Case 1 results at large k, looking consistent with having $\sim 10\%$ of the beat coupling excess variance.

In Figs. 2 - 4, we quantify the off diagonal correlations by plotting the correlation coefficients between bins,

$$\rho_{ij} = \frac{\mathbf{C}_{ij}}{\sqrt{\mathbf{C}_{ii}\,\mathbf{C}_{jj}}}, \tag{25}$$

where we keep one bin fixed and let the value of k corresponding to the other one vary on the horizontal axis. The fixed bins are at $k \approx 0.05h\mathrm{Mpc}^{-1}$ (Fig. 2), $k \approx 0.1h\mathrm{Mpc}^{-1}$ (Fig. 3) and $k \approx 0.15h\mathrm{Mpc}^{-1}$ (Fig. 4).

The conclusions for the correlation coefficients are similar to those drawn about the variances and again the analytic and numerical results agree well. In fact, for Cases 2 & 3, there now is good agreement up to $k = 0.4h\mathrm{Mpc}^{-1}$ for all redshifts. We do note that the correlations relative to the $k \approx 0.05h\mathrm{Mpc}^{-1}$ bin appear to be overestimated somewhat by the analytic expressions. However, this is because the variance from simulations in the $k = 0.04 - 0.05h\mathrm{Mpc}^{-1}$ bin is somewhat high due to noise. Since this variance appears in the denominator of Eq. (25), it brings the simulation correlation coefficients down, thus explaining the slight disagreement with the analytic result. The other deviation is that the correlation coefficients relative to the bins at wave number ($k \approx 0.1h\mathrm{Mpc}^{-1}$ and $k \approx 0.15h\mathrm{Mpc}^{-1}$) are higher than expected for Case 1 at low redshift. Whereas they should theoretically not have a beat coupling contribution, these correlation coefficients behave as if they do get such a contribution similar to the one in the scenario with beat coupling and the local average effect (Case 3). This may simply be due to non-linear effects beyond the order included in our theoretical expressions. Other than this, the agreement is very good.

To conclude this section, we use our analytic expressions to quantify the importance of non-linear corrections to the covariance matrix on a range of scales relevant to large scale structure surveys. In the left panel of Fig. 5, we show the theory based variance, at redshift zero, taking into account all non-linear effects (green), and taking into account everything except for the beat coupling and local average effects (red). The figure is the same as in Fig. 1, except that the focus is on the theory curves in the range $k = 0 - 0.2h/\mathrm{Mpc}$. We see that the full non-linear corrections change the variance at the 25% level at $k = 0.15h/\mathrm{Mpc}$ and by $\sim 50\%$ at $k = 0.2h/\mathrm{Mpc}$. For comparison, we also show in dashed blue the variance if we do not take into account the reduction in excess variance due to the local average effect. Neglecting to take the latter effect into account would lead to a gross overestimate of the non-linear variance.

To quantify the effect of not only the non-linear corrections to the diagonal of the covariance matrix, but also of the non-diagonal elements, we next calculate the squared signal-to-noise ratio,

$$(S/N)^2(k_{\mathrm{max}}) \equiv \sum_{k_i,k_j<k_{\mathrm{max}}} \mathbf{C}^{-1}_{k_i,k_j}\,P(k_i)\,P(k_j). \tag{26}$$

The sum here is over all pairs of power spectrum bins with both central k values below k_{max}. For $P(k_i)$, we use the (non-linear) power spectrum in the fiducial model, so that S/N can be thought of as a *detection* (or amplitude) signal-to-noise. The top figure of the right panel of Fig. 5 shows this statistic using a (diagonal) linear covariance matrix (based on mode counting; black curve), the full non-linear matrix (green), and the non-linear matrix with the off-diagonal elements set to zero (dashed green). The bottom figure shows the relative difference with the linear signal-to-noise squared. Compared to the linear case, the non-linear, but diagonal covariance matrix indeed decreases the signal-to-noise, as expected from the increase in variance seen in the right panel. However, the off-diagonal elements have a significantly stronger effect and decrease the signal-to-noise even further. The signal-to-noise gives an idea of how much Fisher matrix elements in a parameter forecast, or χ^2 values in a Monte Carlo chain, are affected by the non-linear corrections to the covariance matrix. This means that when a power spectrum study, whether with real data or in a Fisher forecast, includes k modes well into the non-linear regime, it is not enough to just include non-linear corrections into the variance of the power spectrum. Instead, the full non-linear covariance matrix needs to be taken into account.

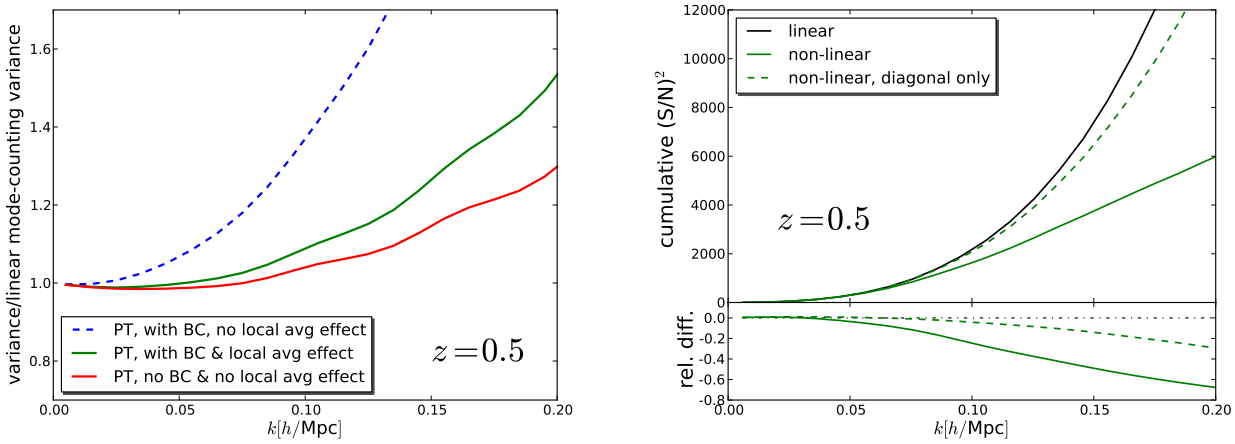

FIG. 5: *Left Panel:* Power spectrum variance relative to variance based on linear power spectrum and mode counting ($2P_{\rm lin}^2/N_k$ with N_k the number of modes per bin) for redshift $z = 0.5$. The red curve represents the case where no modes larger than the survey are present (but other non-linear contributions are included). The most realistic case is depicted in green and includes the beat coupling and local average effects. Non-linear effects cause an increase of variance of up to 50% on scales $k < 0.2h/\mathrm{Mpc}$. The dashed blue curve shows the variance when the local average effect is not taken into account. *Right Panel:* Cumulative signal-to-noise squared in the amplitude of the non-linear power spectrum (see text for details) as a function of largest k-bin included (top; bottom shows fractional difference relative to the linear matrix). The non-linear, off-diagonal terms in the covariance matrix give the strongest contribution to the decrease in signal-to-noise at large k.

V. GENERAL SURVEY GEOMETRY

So far, we have found it convenient to describe the statistics of the density field in terms of a discrete set of overdensity modes. Whereas this is exact when considering the modes of a full periodic box, this should be considered an effective description in the case of subboxes, justified by their simple geometry. In this section, we work out the formalism for a general survey geometry in terms of the continuum of Fourier modes that exist in an infinite universe. We closely follow the notation and results of [7], but will add to this the trispectrum contributions arising in perturbation theory. The discussion in this section will provide justification for the discrete description of the previous sections. One issue that we will pay particular attention to is the motivation for using the diagonal mode-counting expression for the disconnected part of the covariance matrix in Eq. (9). As we will see, in the FKP formalism, the window function correlates power spectrum estimators in different bins (even in the Gaussian case), and the mode counting expression is not accurate. We will explain why mode counting was justified in the previous sections, argue that for a realistic survey the FKP description including cross correlations is the relevant one, and finally test the latter description against simulations.

We first define our Fourier convention as

$$\delta(\mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\, e^{i\mathbf{k}\cdot\mathbf{x}}\, \delta(\mathbf{k}). \tag{27}$$

The statistics of the continuum of Fourier modes are given by

$$\begin{aligned}
\langle \delta(\mathbf{k})\, \delta(\mathbf{k}')\rangle &= (2\pi)^3\, P(k)\, \delta^D(\mathbf{k}+\mathbf{k}') \\
\langle \delta(\mathbf{k})\, \delta(\mathbf{k}')\, \delta(\mathbf{k}'')\rangle &= (2\pi)^3\, B(\mathbf{k},\mathbf{k}',\mathbf{k}'')\, \delta^D(\mathbf{k}+\mathbf{k}'+\mathbf{k}'') \\
\langle \delta(\mathbf{k})\, \delta(\mathbf{k}')\, \delta(\mathbf{k}'')\, \delta(\mathbf{k}''')\rangle_c &= (2\pi)^3\, T(\mathbf{k},\mathbf{k}',\mathbf{k}'',\mathbf{k}''')\, \delta^D(\mathbf{k}+\mathbf{k}'+\mathbf{k}''+\mathbf{k}''') \\
&\cdots
\end{aligned} \tag{28}$$

For a survey with background number density $\bar{n}(\mathbf{x})$, consider then the weighted density field

$$F(\mathbf{x}) = \frac{\bar{n}(\mathbf{x})\, w(\mathbf{x})\, \delta(\mathbf{x})}{\left[\int d^3\mathbf{x}\, \bar{n}^2(\mathbf{x})\, w^2(\mathbf{x})\right]^{\frac{1}{2}}} \equiv G(\mathbf{x})\, \delta(\mathbf{x}), \tag{29}$$

where $w(\mathbf{x})$ is a weight function that can be chosen to maximize signal to noise. We will ignore shot noise in the following, but it is straightforward to include it in the Gaussian approximation (see FKP). The Fourier transform is given by $F(\mathbf{k}) = (G * \delta)(\mathbf{k})$, where "$*$" indicates a convolution, and the power spectrum can be estimated by $\hat{P}(\mathbf{k}) = |F(\mathbf{k})|^2$, so that

$$\langle \hat{P}(\mathbf{k}) \rangle = \int \frac{d^3\mathbf{k}'}{(2\pi)^3} P(k') \, |G(\mathbf{k} - \mathbf{k}')|^2. \tag{30}$$

The measured power spectrum is thus a weighted average of the true spectrum, with a weight function of width of order the fundamental mode of the survey, i.e. $\Delta k \sim \pi/L$ if L is the typical scale of the survey. As a relevant example, the window function for a cubic box of side L ($\bar{n} \equiv$ const, $w \equiv 1$ inside the box and zero outside) is given by

$$G(\mathbf{k}) = L^{3/2} \, j_0 \left(Lk_x/2\right) \, j_0 \left(Lk_y/2\right) \, j_0 \left(Lk_z/2\right), \tag{31}$$

with $j_0(x) = \sin x/x$ the zeroth spherical Bessel function. If the true spectrum varies little across this range of scales, we get

$$\langle \hat{P}(\mathbf{k}) \rangle \approx P(k). \tag{32}$$

A. Covariances - disconnected (Gaussian) contribution

We first consider the contribution to the covariance between FKP power spectrum estimators arising from products of two-point functions. This is the only contribution for a Gaussian density field and we can follow FKP for its description.

We start from the two-point function

$$\langle F(\mathbf{k}) \, F^*(\mathbf{k}') \rangle = \int \frac{\mathbf{k}''}{(2\pi)^3} P(k'') \, G(\mathbf{k} - \mathbf{k}'') \, G^*(\mathbf{k}' - \mathbf{k}'') \approx P(k) \, Q(\mathbf{k}' - \mathbf{k}'), \tag{33}$$

where the second equality is true in the same limit where the (expectation value of the) power spectrum estimator equals the true spectrum, and $Q(\mathbf{k})$ is the Fourier transform of the (normalized) squared window function,

$$Q(\mathbf{x}) = \frac{\bar{n}^2(\mathbf{x}) \, w^2(\mathbf{x})}{\int d^3\mathbf{x} \, \bar{n}^2(\mathbf{x}) \, w^2(\mathbf{x})}. \tag{34}$$

Hence, the covariance

$$\langle \delta\hat{P}(\mathbf{k}) \, \delta\hat{P}(\mathbf{k}') \rangle \approx |P(k) \, Q(\mathbf{k} - \mathbf{k}')|^2 + (\mathbf{k}' \to -\mathbf{k}'). \tag{35}$$

In FKP, the power spectrum in a bin i is given by

$$\hat{P}_i = \int_i \frac{d^3\mathbf{k}}{V_{k,i}} \, \hat{P}(\mathbf{k}), \tag{36}$$

where $V_{k,i}$ is the k-volume of bin i, so that the covariance between (isotropic) bins is

$$\mathbf{C}_{ij} = 2 \int_i \frac{d^3\mathbf{k}}{V_{k,i}} \int_j \frac{d^3\mathbf{k}'}{V_{k,j}} \, P^2(k) \, |Q(\mathbf{k} - \mathbf{k}')|^2. \tag{37}$$

Because of the extended nature of the window function $|G|^2$, there will thus be correlations between different bins. In the limit that the bin width is much larger than the width of this window function, these correlations are negligible and one can make a further simplification by integrating out $|G|^2$. Using

$$\int \frac{d^3\mathbf{k}}{(2\pi)^3} \, |Q(\mathbf{k})|^2 = \frac{\int d^3\mathbf{x} \, \bar{n}^4(\mathbf{x}) \, w^4(\mathbf{x})}{\left[\int d^3\mathbf{x} \, \bar{n}^2(\mathbf{x}) \, w^2(\mathbf{x})\right]^2} \equiv V_{\rm eff}^{-1} \tag{38}$$

one ends up with

$$\mathbf{C}_{ij} = \frac{2P^2(k_i)}{N_k} \, \delta^K_{ij}, \qquad N_k = \frac{V_{k,i}}{(2\pi)^3/V_{\rm eff}}, \tag{39}$$

where k_i is a typical mode in the i-th bin and we have further assumed that the power spectrum varies little across a bin. This is the well know mode-counting result. The expression for the effective volume is further simplified in FKP by choosing an optimal weight function, but we will stick to general $w(\mathbf{x})$. Note that for our cubic box (or any other geometry with constant background number density and weighting), $V_{\rm eff} = V$.

Assuming that embedding our $L_{\rm sub} = 600h^{-1}$Mpc cubed box in a $L = 2400h^{-1}$Mpc cubed periodic box is a good approximation to embedding it in an infinite universe with a continuum of Fourier modes, we can apply the above to our simulated scenario and find that for our bin width $\Delta k = 0.01h{\rm Mpc}^{-1}$, there should be significant correlations between bins and the mode-counting argument is not valid, in fact overpredicting the variances. We show this in figures 6 - 9, dashed lines, which we will discuss in more detail below.

Why then was it correct to use the diagonal, mode-counting expression in the previous sections? The reason is that, in keeping with our effective periodic box description, our bin average only includes modes which are multiples of the *subbox*'s fundamental mode, i.e. $\mathbf{k} = 2\pi/L_{\rm sub}\,\mathbf{n}$ with the components of $\mathbf{n}$ integers,

$$\hat{P}_i = \frac{1}{N_i}\sum_{\mathbf{k}\in i}\hat{P}(\mathbf{k}). \tag{40}$$

This decorrelates the binned power spectrum estimators as the mode mixing kernel

$$Q(\mathbf{k}) = j_0\left(L_{\rm sub}k_x/2\right)\, j_0\left(L_{\rm sub}k_y/2\right)\, j_0\left(L_{\rm sub}k_z/2\right) \tag{41}$$

vanishes for separations that are a non-zero multiple of $2\pi/L_{\rm sub}$. Since the kernel equals unity for zero separation, the covariance in this averaging scheme is given by the mode-mixing result,

$$\mathbf{C}_{ij} = 2\frac{1}{N_i}\frac{1}{N_j}\sum_{\mathbf{k}\in i}\sum_{\mathbf{k}'\in j} P^2(k)\,|Q(\mathbf{k}-\mathbf{k}')|^2 = \delta^K_{ij}\,2\frac{1}{N_i}\frac{1}{N_j}\sum_{\mathbf{k}\in i} P^2(k) = \frac{2P^2(k_i)}{N_i}\,\delta^K_{ij}. \tag{42}$$

This begs the question if a similar binning scheme can be applied to an actual survey, in order to decorrelate power spectrum estimates in different bins. Unfortunately, for a realistic survey, the geometry will be much more complicated than a simple cube, making this rather difficult. In practice, therefore, one would typically embed the survey in a much larger cubic volume (zero-padding the part not covered by the survey), apply a Fourier transform, estimate the power spectrum for each $\mathbf{k}$ on the grid by the embedding box (i.e. modulo $2\pi/L$ with L the size of the large box), and finally average to obtain the binned spectrum. This is thus a much denser sampling than the one we applied previously and approaches the infinitely dense FKP bin average of Eq. (36). In a more realistic scenario therefore, mode counting would not be sufficient and the off diagonal covariances are significant.

For this reason, it is useful to numerically test the mode mixing in the covariance matrix due to the window function as given in Eq. (37). We do this by again calculating the covariance matrix for our $L_{\rm sub} = 600h^{-1}$Mpc subbox, but this time using a binned spectrum averaged over all multiples of the fundamental mode of the large $L = 2400h^{-1}$Mpc simulation box, thus increasing the sampling density by a factor of $4^3 = 64$ and approaching the FKP bin average. In practice, we zero-pad the exterior of the subbox and then apply the Fourier transform to the full box. The results are shown in figures 6 - 9 and compared to the analytic expression (37). To model the inevitable non-linear effects, we add trispectrum contributions to the theory prediction, which we will describe in the next subsection. Even without these contributions however, we can already see from the linear regime that the FKP expression works very well and that indeed with the more realistic averaging scheme, the off diagonal covariances are considerable and the variances are reduced accordingly.

To conclude this subsection, we briefly compare our beat coupling investigation to that of [25]. They too study the beat coupling effect by considering a subbox of a larger simulation volume (although in their case the subbox is only a factor $2^3 = 8$ smaller than that of the simulation, $L_{\rm sub} = 500h^{-3}{\rm Mpc}^3$). They use two approaches for estimating the power spectrum. Their "zero-padding" treatment corresponds to the FKP-like method described in the preceding paragraphs, while the other approach is equivalent to the one we used consistently before this section. While they do not comment on the origin of the observed differences, it is reassuring that the simulation results in their Fig.(10) agree with our explanation and with our own simulation results. Interestingly, in [25], the excess covariance due to large modes is significantly smaller than expected based on beat coupling only. However, they use the local (subbox) average to calculate the overdensity. The lower variance can thus be explained very well by the local average affect discussed in this article.

B. General Survey Geometry: Perturbation Theory

The discussion in the previous section covers the linear regime where the underlying density field is to a good approximation Gaussian. Including non-linear effects, it still describes the disconnected contribution to the covariance

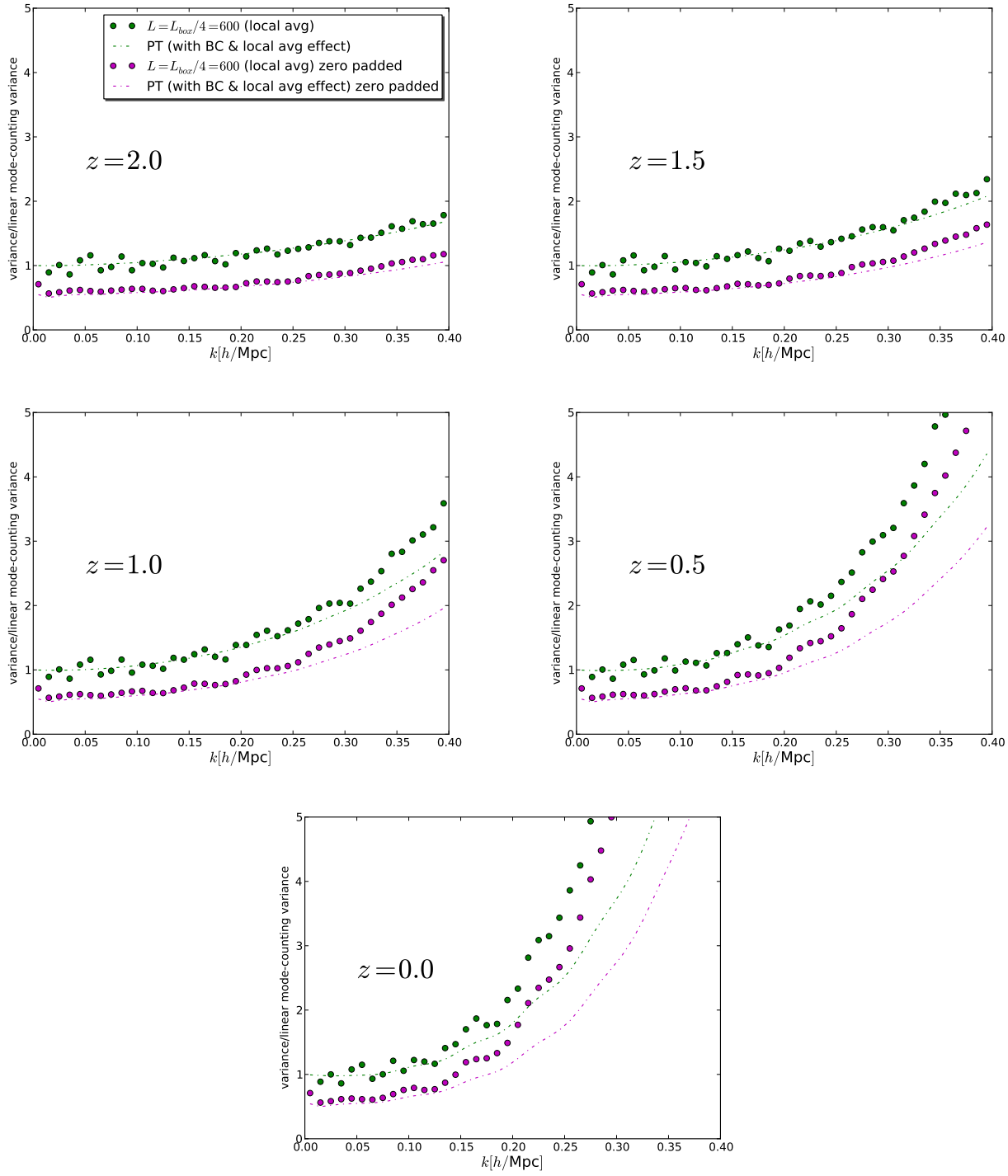

FIG. 6: Power spectrum variance relative to (linear) mode-counting variance at redshifts $z = 0 - 2$ for Case 3, i.e. with both beat coupling *and* the local average effect included. Comparison is between variances in spectrum averaged over bins (of width $\Delta k = 0.01h\mathrm{Mpc}^{-1}$) with fine sampling (as in FKP, Eq. (40)), which has mode mixing due to the window function (magenta dots), and average over modes that are multiples of fundamental mode of subbox (Eq. (36)), which decorrelates the power spectrum estimator in bins (green dots). We also refer to the former case as the "zero padding" covariance matrix, as the power is estimated from simulations by using the full $2400h^{-1}$Mpc simulation box to do the Fourier transform, but zero padding the region outside the $600h^{-1}$Mpc subbox. Dashed lines show theory predictions from Eq. (23) (green) and Eq. (48) (magenta). The theory covariance matrix thus successfully includes window function effects and non-linear effects. On large scales (small k), the green curve and dots approach unity because the bin averaging scheme causes the variance to be given by mode counting and because non-linear effects vanish. The more realistic zero padding case however has significantly lower variances, but strong correlations between neighboring bins (even on linear scales) so that the total information content is the same.

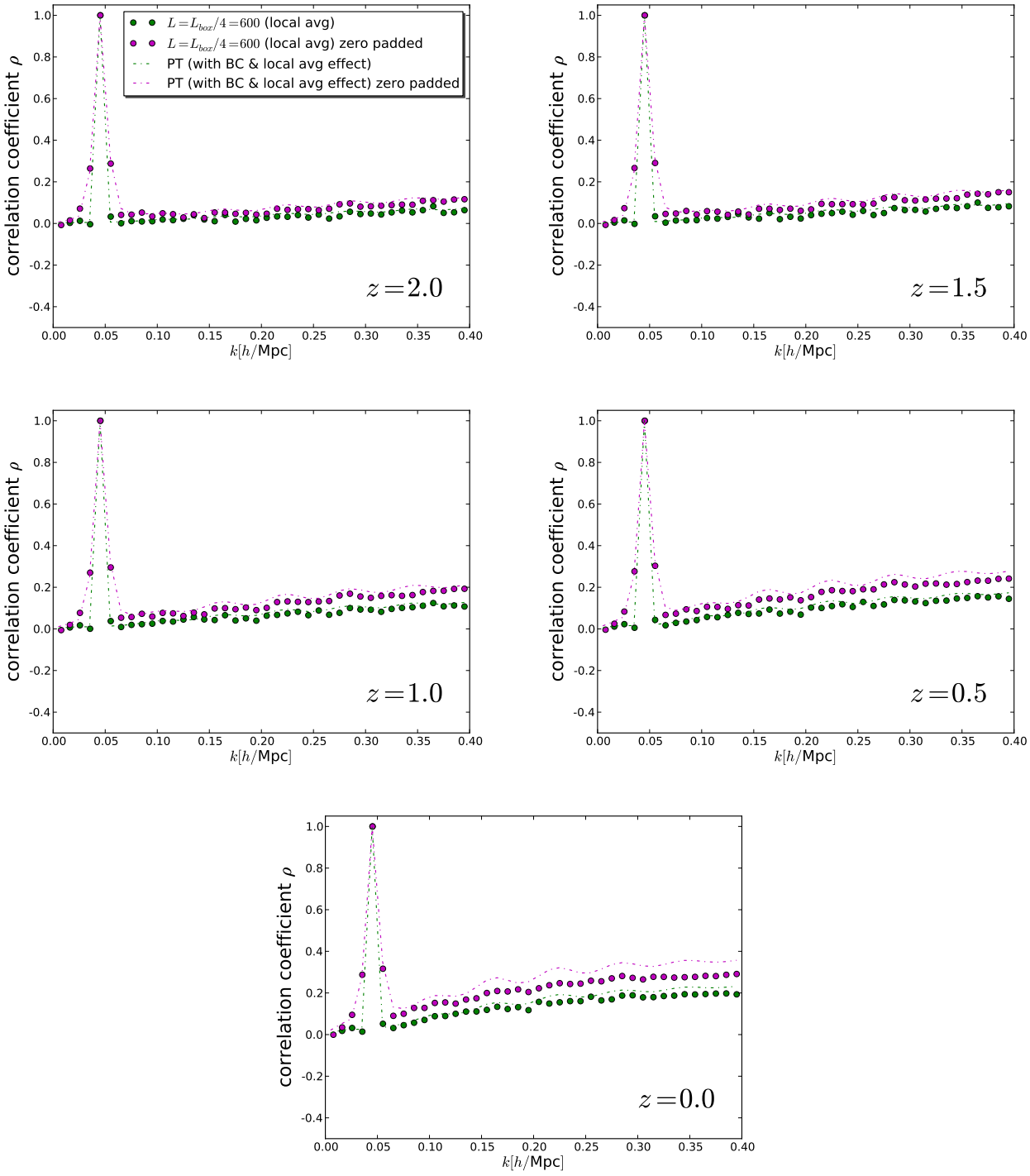

FIG. 7: As Fig. 6, but here showing the correlation coefficients $\rho_{ij} \equiv \mathbf{C}_{ij}/\sqrt{\mathbf{C}_{ii}\,\mathbf{C}_{jj}}$ of i-th bin at k on x-axis, relative to j-th bin at $k = 0.04 - 0.05h\mathrm{Mpc}^{-1}$. The theory predictions again agree quite well with the simulation results although more so for correlations relative to bins at larger k, as shown in Figs 8 and 9. Note in particular the large correlation between neighboring bins in the case where the bin average is based on fine sampling/zero padding (Eq. (36), magenta). This is due to mode mixing by the window function and is separate from the cross correlations due to non-linear evolution. It is absent when estimators are averaged using the sparse bin sampling (green), Eq. (40).

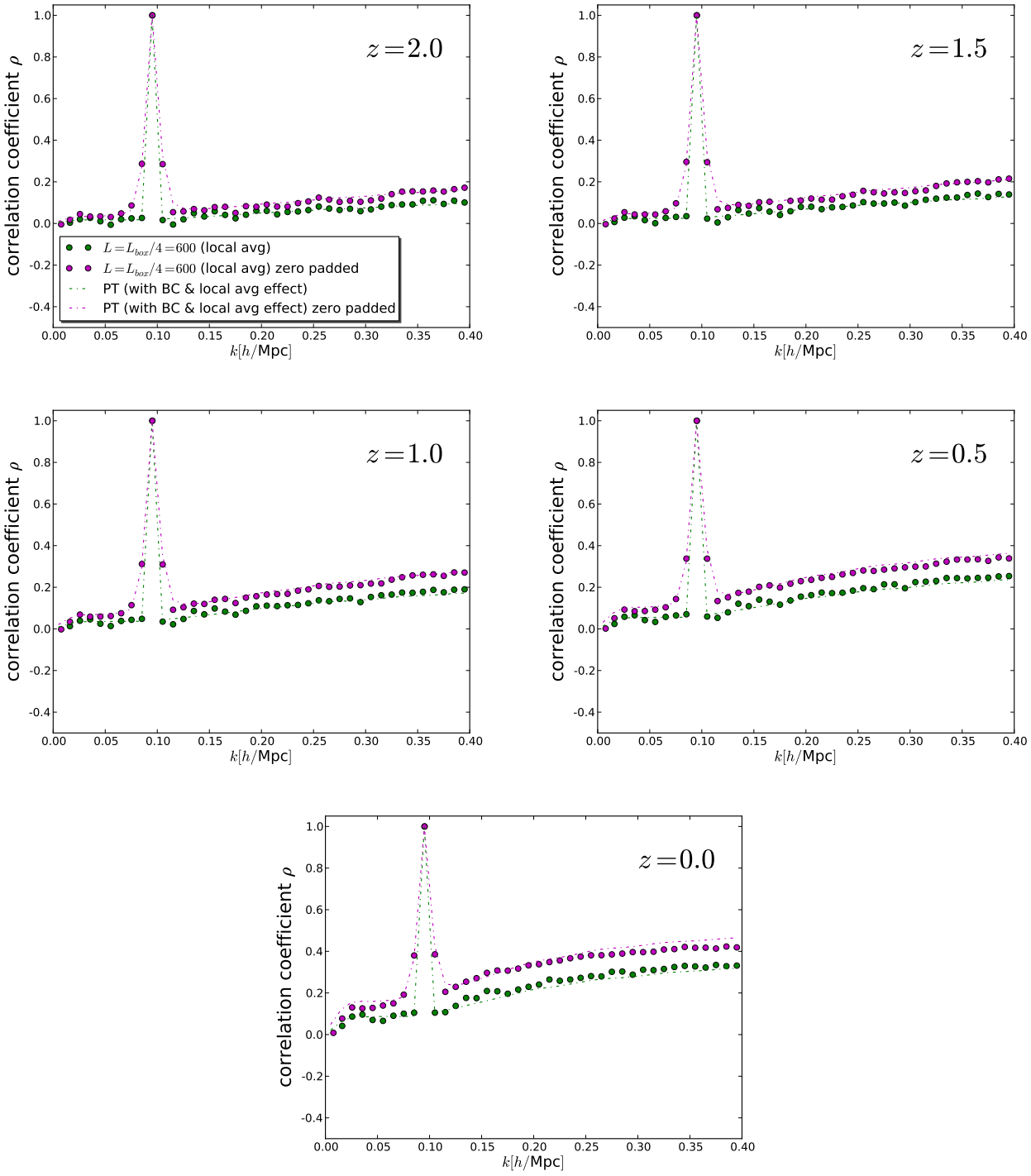

FIG. 8: As Fig. 7, but correlations relative to bin at $k = 0.09 - 0.1h\mathrm{Mpc}^{-1}$.

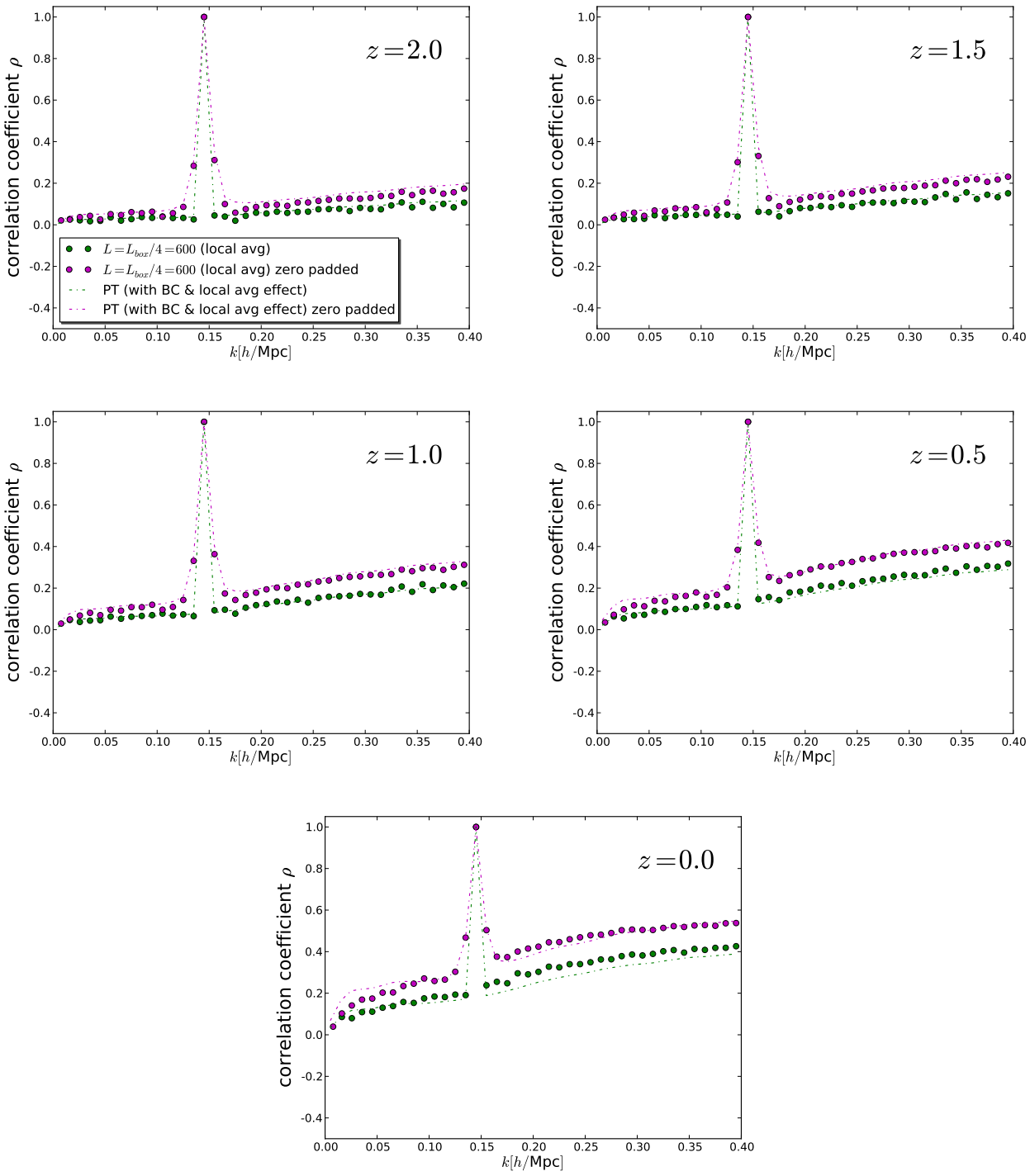

FIG. 9: As Fig. 7, but correlations relative to bin at $k = 0.14 - 0.15h\mathrm{Mpc}^{-1}$.

matrix provided that the non-linear power spectrum is used. However, additional terms are now needed as the covariance matrix receives contributions from the connected part of the four-point function. These terms were already included in the figures 6 - 9 (and were discussed in a simpler framework in section II). Here we discuss the details of the inclusion of the trispectrum terms for arbitrary window function.

Including these terms, the covariance between single-mode estimators now becomes

$$\begin{aligned}\langle\delta\hat{P}(\mathbf{k})\,\delta\hat{P}(\mathbf{k}')\rangle &= |P(k)\,Q(\mathbf{k}-\mathbf{k}')|^2 + (\mathbf{k}'\to-\mathbf{k}') \\ &+ \frac{1}{(2\pi)^9}\int d^3\mathbf{q}_1 d^3\mathbf{q}_2 d^3\mathbf{q}_3 d^3\mathbf{q}_4\,\delta^D(\mathbf{q}_1-\mathbf{q}_2+\mathbf{q}_3-\mathbf{q}_4)\,G(\mathbf{q}_1)G^*(\mathbf{q}_2)G(\mathbf{q}_3)G^*(\mathbf{q}_4) \\ &\times\ T(\mathbf{k}-\mathbf{q}_1,-(\mathbf{k}-\mathbf{q}_2),\mathbf{k}'-\mathbf{q}_3,-(\mathbf{k}'-\mathbf{q}_4)). \end{aligned} \tag{43}$$

Note that this is the generalization of Eq. (8), which we used to describe the covariance in a cubic box. To leading order in perturbation theory, the trispectrum is given by[4] (see [8])

$$T(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) = 4\left[F_2(\mathbf{k}_{12},-\mathbf{k}_1)\,F_2(\mathbf{k}_{12},-\mathbf{k}_3)\,P(k_1)\,P(k_{12})\,P(k_3)+\text{perm.}\right]+6\left[F_3(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)\,P(k_1)\,P(k_2)\,P(k_3)+\text{cyc.}\right], \tag{44}$$

with 12 distinct permutations for the first term, 3 cyclic permutations for the second, and where $\mathbf{k}_{ij}\equiv\mathbf{k}_i+\mathbf{k}_j$. Inserting the arguments from Eq. (43) into this perturbation theory expression, and omitting the $\mathbf{q}_i$ dependence whenever $\mathbf{q}_i$ appears in an argument as a correction to a larger vector of order k, k', $|\mathbf{k}+\mathbf{k}'|$, etc, the trispectrum contribution becomes

$$\begin{aligned} T\ \ &(\mathbf{k}-\mathbf{q}_1,-(\mathbf{k}-\mathbf{q}_2),\mathbf{k}'-\mathbf{q}_3,-(\mathbf{k}'-\mathbf{q}_4)) = \Big[4\,P(|\mathbf{k}+\mathbf{k}'|)\ (F_2(\mathbf{k}+\mathbf{k}',-\mathbf{k})\,P(k)+(\mathbf{k}\to\mathbf{k}'))^2+(\mathbf{k}'\to-\mathbf{k}')\Big] \\ &+\ 4\,P(k)\,P(k')\,P(|\mathbf{q}_2-\mathbf{q}_1|)\ (F_2(\mathbf{q}_2-\mathbf{q}_1,\mathbf{k})+F_2(\mathbf{q}_2-\mathbf{q}_1,-\mathbf{k}))\ (F_2(\mathbf{q}_2-\mathbf{q}_1,\mathbf{k}')+F_2(\mathbf{q}_2-\mathbf{q}_1,-\mathbf{k}')) \\ &+\ 12\,P(k)\,P(k')\ [F_3(\mathbf{k},-\mathbf{k},\mathbf{k}')\,P(k)+(\mathbf{k}\leftrightarrow\mathbf{k}')] \end{aligned} \tag{45}$$

where we have used that the trispectrum arguments sum up to zero. The terms that are independent of the $\mathbf{q}_i$'s can be pulled out of the integral in Eq. (43) so that the integral over mode mixing kernels simply gives a factor of $V_{\rm eff}^{-1}$. The other terms only depend on the combination $\mathbf{q}_1-\mathbf{q}_2$ so that three of the four Fourier integrals can be carried out analytically and one remains, giving the result

$$\begin{aligned}\langle\delta\hat{P}(\mathbf{k})\,\delta\hat{P}(\mathbf{k}')\rangle &= |P(k)\,Q(\mathbf{k}-\mathbf{k}')|^2 + (\mathbf{k}'\to-\mathbf{k}') \\ &+ \frac{1}{V_{\rm eff}}\Big[4\,P(|\mathbf{k}+\mathbf{k}'|)\ (F_2(\mathbf{k}+\mathbf{k}',-\mathbf{k})\,P(k)+(\mathbf{k}\leftrightarrow\mathbf{k}'))^2+(\mathbf{k}'\to-\mathbf{k}')\Big] \\ &+ \frac{1}{V_{\rm eff}}12\,P(k)\,P(k')\ [F_3(\mathbf{k},-\mathbf{k},\mathbf{k}')\,P(k)+(\mathbf{k}\leftrightarrow\mathbf{k}')] \\ &+ \frac{1}{V_{\rm eff}}4\,P(k)\,P(k')\ \left(\int\frac{d^3\mathbf{u}}{(2\pi)^3}\,|Q|^2(\mathbf{u})\right)^{-1} \\ &\times\ \int\frac{d^3\mathbf{u}}{(2\pi)^3}\,|Q|^2(\mathbf{u})P(u)\ (F_2(\mathbf{u},\mathbf{k})+F_2(\mathbf{u},-\mathbf{k}))\ (F_2(\mathbf{u},\mathbf{k}')+F_2(\mathbf{u},-\mathbf{k}'))\,). \end{aligned} \tag{46}$$

The final step towards a covariance matrix is to apply the bin average to Eq. (46), but for the trispectrum terms the subtle, averaging related effects described in the previous subsection are not important because the correlations from the trispectrum are not as narrow as those for the disconnected (or Gaussian) terms. The angle averaging can be done analytically for the last term (the beat coupling term) and has to be carried out numerically for the remaining trispectrum contributions. The final result is

$$\begin{aligned}\mathbf{C}_{ij} &= 2\int\frac{d^3\mathbf{k}}{V_{k,i}}\int\frac{d^3\mathbf{k}'}{V_{k,j}}\,P^2(k)\,|Q(\mathbf{k}-\mathbf{k}')|^2 \\ &+ \frac{1}{V_{\rm eff}}\int\frac{d^3\mathbf{k}}{V_{k,i}}\int\frac{d^3\mathbf{k}}{V_{k,i}}\,T^0(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}') \\ &+ \frac{1}{V_{\rm eff}}\,16\left(\frac{17}{21}\right)^2\left(\int\frac{d^3\mathbf{u}}{(2\pi)^3}\,|Q|^2(\mathbf{u})\right)^{-1}\int\frac{d^3\mathbf{u}}{(2\pi)^3}\,|Q|^2(\mathbf{u})P(u), \end{aligned} \tag{47}$$

[4] Unlike in section II, but consistent with perturbation theory, we now immediately express perturbation theory quantities in terms of the non-linear power spectrum P and not $P^{\rm lin}$.

where $T^0(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}')$ represents the second and third lines of Eq. (46). The last term is the beat coupling term, which can now explicitly be seen to be proportional to a weighted average of the power spectrum over large modes. The integral describing this average can be compared to that for the expectation value of the zero-mode power spectrum, Eq. (30). The only difference is that the latter is given in terms of G while the former is given in terms of Q (note that the normalization integral appearing in the beat coupling term is equal to one when Q is replaced by G). However, for a cubic box (with $\bar{n} \equiv \text{const}, w \equiv 1$ inside the box), the two quantities are exactly equal and the average appearing in the beat coupling expression above can be replaced by the zero-mode power spectrum, thus justifying our use of $P_0^{\rm lin}$ in Eq. (15) in section II B.

We will not rederive the correction due to the local average effect for the case of arbitrary geometry, but in analogy with the cubic subbox case, we will use the following expression for Case 3 in Figs 6-9:

$$\begin{aligned}
\mathbf{C}_{ij} &= 2\int \frac{d^3\mathbf{k}}{V_{k,i}} \int \frac{d^3\mathbf{k}'}{V_{k,j}} P^2(k)\,|Q(\mathbf{k}-\mathbf{k}')|^2 \\
&+ \frac{1}{V_{\rm eff}} \int \frac{d^3\mathbf{k}}{V_{k,i}} \int \frac{d^3\mathbf{k}}{V_{k,i}} T^0(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}') \\
&+ \frac{1}{V_{\rm eff}} \frac{676}{441} \left(\int \frac{d^3\mathbf{u}}{(2\pi)^3}\,|Q|^2(\mathbf{u})\right)^{-1} \int \frac{d^3\mathbf{u}}{(2\pi)^3}\,|Q|^2(\mathbf{u}) P(u). \qquad (48)
\end{aligned}$$

Note that for a varying background density, the estimate of the background density $\bar{\rho}(\mathbf{x})$ used in $\delta(\mathbf{x}) = (\rho(\mathbf{x}) - \bar{\rho}(\mathbf{x}))/\bar{\rho}(\mathbf{x})$ is not only affected by the effective zero mode, but also by smaller modes. These will also affect the covariance matrix, but we will not go into this effect here, as it is beyond the topic we set out to study.

VI. SUMMARY AND DISCUSSION

In this work, we have studied the effects of modes larger than the survey on the (dark matter) power spectrum covariance matrix. We have built an analytic description that includes the beat coupling effect of [21], but also the previously overlooked (at least in analytic studies) effect of large modes on the estimated average density $\bar{\rho}$ which enters the overdensity through $\delta = (\rho - \bar{\rho})/\bar{\rho}$ (the local average effect). We have confirmed this model and the role of individual contributions by comparing to covariance matrices obtained from N-body simulations. To study the role of super-survey modes in the simulations, we estimated the power spectrum from a subvolume embedded in a considerably larger simulation volume.

We summarize our main results below:

- We build a model based on perturbation theory for the matter covariance matrix that includes the effects of modes larger than the survey. For the variances, we find excellent agreement with simulations for $k < 0.4h\text{Mpc}^{-1}$ (or larger) at $z = 2$ and for $k < 0.2(0.15)h\text{Mpc}^{-1}$ at $z = 0.5(0)$. Agreement is even better for the correlation coefficients, with our model predicting the correct coefficients for at least $k < 0.4h\text{Mpc}^{-1}$ at all redshifts.

- When isolated, the beat coupling effect from [21, 22], can indeed be described by the last term in Eq. (22), as shown by the blue points and curves of Figs 1-4.

- In a more realistic approach, the *local average effect* needs to be taken into account as well. This has previously been overlooked in analytic studies and we derive its effect using perturbation theory, leading to Eq. (23). It reduces the covariance, leaving only 10% of the original beat coupling excess covariance, as shown by the green lines and points in Figs 1-4. This also explains the disagreement found between the beat coupling-only expression and simulations in [25]. We conclude that the beat coupling excess covariance is not as important as previously thought.

- Eq. (48) gives the final result for the matter covariance matrix for arbitrary survey geometry, including not only the above mentioned effects, but also the correlations between neighboring power spectrum bins and the related reduced variance due to the survey's window function. It is depicted in Figs 6-9 and again agrees well with simulations. It can be used as a first step towards a covariance matrix for the galaxy (or other tracer's) power spectrum.

In a real survey of large scale structure, the effects discussed in this paper are relevant because there are always modes larger than the survey volume. Our results are thus important for large scale structure surveys and in particular galaxy surveys, as we quantify the expected excess covariance due to these modes, as well as the more standard covariance

contributions. With a complete description of the covariance matrix now available, it is also possible to study the cosmology dependence of the covariance matrix in an efficient manner. We refer to the Appendix for a first look into this issue.

To build a complete covariance matrix for the galaxy power spectrum, one needs to include the effects of shot noise, galaxy bias and redshift space distortions. While the effect of shot noise is easy to incorporate within the FKP formalism (at least in the limit where the shot noise can be treated as Gaussian), the other effects are more complicated and clearly beyond the scope of our paper. However, a very rough first approximation to a galaxy covariance matrix could be obtained by multiplying the dark matter matrix by the galaxy bias to the fourth power, including an angle averaged Kaiser factor (as in [25]) to account for redshift space distortions, and using FKP to include the effect of shot noise.

In addition to the complications arising from observing galaxies (or other tracers) instead of dark matter, one also needs to be able to describe the covariance for a realistic survey geometry (not many surveys have a cubic footprint). For this reason, we presented in section V a full description of the model for arbitrary survey geometry, assuming the FKP estimator is used. By comparing to simulations, we showed that the expressions from [7] (with trispectrum terms added) can be used to accurately describe the mode mixing due to the survey window function and the resulting correlations between neighboring power spectrum bins. While section V does not present any completely new results relative to the previous sections and what is in the literature, its main purpose is to present a complete set of equations to describe the matter power spectrum for arbitrary survey geometry. This can then serve as a stepping stone towards a full galaxy covariance matrix.

Acknowledgments

We thank Ryuichi Takahashi, Masahiro Takada and Shun Saito for useful discussion and for sharing their simulation outputs, which helped motivate our investigation. We also thank Beth Reid for insightful discussions. RdP and CW are supported by FP7-IDEAS-Phys.LSS 240117. OM is supported by AYA2008-03531 and the Consolider Ingenio project CSD2007-00060. LV acknowledges support of FP7-IDEAS-Phys.LSS 240117. N-body simulations and calculations on N-body outputs were done on the cluster of computers *Hipatia* (UB computing facilities and ERC grant FP7- IDEAS Phys.LSS 240117).

Appendix A: Cosmology Dependence of Covariance Matrix

One major advantage of the analytic expressions presented in this work, is that they allow for a quick estimate of the covariance matrix, especially compared to methods using N-body simulations. The ease with which covariance matrices can be calculated makes the analytic method perfectly suited for studying the cosmology dependence of the covariance matrix. A full study of this cosmology dependence would consider how constraints on cosmological parameters are affected and would quantify the error induced by ignoring the cosmology dependence. Such an investigation is beyond the scope of this article and we will leave it for future work. In this appendix, we will simply quantify how much the covariance matrix changes as we vary individual cosmological parameters.

For simplicity, we again consider the covariance in the matter power spectrum, as estimated from a cubic volume with $V = L^3 = (600h^{-1}\mathrm{Mpc})^3$. We imagine the spectrum is "measured" at redshift $z = 0.5$ and include the smearing effect due to the window function that arises when the FKP estimator is used. In other words, we will consider the covariance matrix given by Eq. (48), with the power spectrum evaluated at $z = 0.5$.

In practice, galaxy surveys typically measure the power spectrum relative to a fixed, fiducial background cosmology, see e.g. [35]. Therefore, when a Monte Carlo chain is run, the theoretical power spectrum at each point in parameter space needs to be rescaled to account for the effect of using the fiducial background cosmology as opposed to the actual background cosmology at that point in parameter space. Only after this rescaling can it be compared to the observed spectrum. For consistency, we therefore also rescale the covariance to the fiducial background cosmology,

$$\mathbf{C}_f(\Delta k_i, \Delta k_j) = A^{-6}\, \mathbf{C}(\Delta k_i/A, \Delta k_j/A)|_{V=V_f\, A^3}, \tag{A1}$$

where $\mathbf{C}_f$ is the covariance matrix of the power spectrum estimator relative to the fiducial background cosmology, $\mathbf{C}$ the true covariance matrix given by Eq. (48) (evaluated at the actual volume $V = V_f\, A^3$, with V_f the survey volume calculated in the fiducial cosmology), Δk_i the bin widths, and $A \equiv d_V(z)/d_V^f(z)$ the dilation factor (see [35]).

Since the different terms in the covariance matrix scale either like the second or third power of the power spectrum, we expect a significant dependence on σ_8 and other parameters changing the overall normalization of the spectrum. The beat coupling/local average term is particularly sensitive to the power at $k \lesssim 2\pi/L$ so that it should depend on

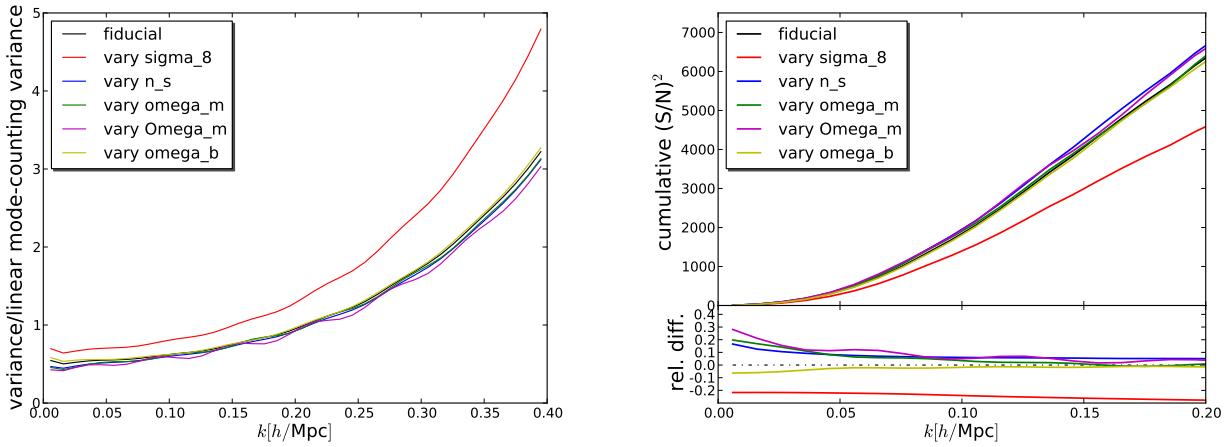

FIG. 10: *Left Panel:* Variance relative to linear, mode counting based variance as a function of cosmology. The black curve shows the fiducial cosmology, and the other curves depict the effect of varying each ΛCDM parameter by approximately twice its error bar expected from a large scale structure plus CMB measurement (see text, and Table I). *Right Panel:* The signal-to-noise squared in the detection/amplitude of the power spectrum as a function of the largest included mode k (main figure). The bottom inset shows the relative difference with the fiducial cosmology. Both panels show that the strongest parameter dependence is on σ_8 and Ω_m.

any parameter that affects the power on very large scales. To make this more quantitative, we consider the ΛCDM fiducial cosmology described in section III and change each parameter p by a step Δp, one by one, keeping the remaining parameters fixed. For the step sizes, we choose $\Delta p = 2\,\sigma_p$, where σ_p is the parameter uncertainty from [6] (Table 3), obtained from combining the SDSS-II Data Release 7 [36] halo power spectrum with WMAP5 [37] cosmic microwave background (CMB) data. These step sizes are indicative of the relevant parameter range in a Markov Chain Monte Carlo (MCMC) analysis with power spectrum data. We show the fiducial parameters and their step sizes in Table I.

	ω_b	ω_m	Ω_m	σ_8	n_s
fiducial	0.023	0.1323	0.27	0.79	0.96
step size	0.00116	0.008	0.038	0.05	0.026

TABLE I: Fiducial values and step sizes used to test cosmology dependence of matter covariance matrix.

In the left panel of figure 10, we show the effect on the variance of each parameter. As before, we normalize the variance by the variance based on mode counting and the linear power spectrum (as calculated in the fiducial cosmology). The units on the horizontal axis are $[h_f/\mathrm{Mpc}]$, where $h_f = 0.7$ is the dimensionless Hubble parameter in the fiducial model. The variance in the fiducial model is given by the black curve. As expected, the largest variation is obtained when varying σ_8, causing a significant increase in variance. The effect of the other parameters is much smaller.

The effect of the parameter variations on the correlation coefficients is always below $|\Delta\rho_{ij}| \lesssim 0.05$, with ω_m decreasing the coefficients most, and σ_8 increasing them by the largest amount. Finally, to incorporate the properties of the full covariance matrix in a single statistic, we calculate the squared signal-to-noise ratio in the (non-linear) power spectrum amplitude as a function of $k_{\rm max}$, see Eq. (26). In a more thorough analysis, one could replace the power spectrum in the expression for $(S/N)^2$ by derivatives with respect to cosmological parameters in order to create a full Fisher matrix.

The right panel of figure 10 depicts this signal-to-noise squared as a function of $k_{\rm max}$ for the different cosmologies, with the bottom inset showing the relative difference with respect to $(S/N)^2$ in the fiducial model. We focus on the range $k = 0 - 0.2\,h_f/\mathrm{Mpc}$, as this is a more realistic range for a galaxy survey (due to strong non-linearities and shot noise on smaller scales), and because we have seen that our approach loses accuracy on smaller scales. Consistent with the picture arising from the left panel, we see again that σ_8 has by far the largest effect, decreasing the signal-to-noise

by $\sim 30\%$ (with only a weak dependence of the relative change on $k_{\rm max}$). The other parameters have a much more modest effect ($\lesssim 10\%$ for $k_{\rm max} \sim 0.1h/{\rm Mpc}$). We conclude that reasonable variations in cosmic parameters can cause $\sim 30\%$ changes in the covariance matrix and thus in $\Delta\chi^2$ values in an MCMC chain.

Finally, we note that, for a galaxy survey, the effect of galaxy bias on the covariance matrix will be very important as well, and will look similar to the effect of σ_8 discussed above.

[1] H. Seo and D. J. Eisenstein, Astrophys. J. **598**, 720 (2003), arXiv:astro-ph/0307460.
[2] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol, R. Scranton, H.-J. Seo, M. Tegmark, Z. Zheng, et al., Astrophys. J. **633**, 560 (2005), arXiv:astro-ph/0501171.
[3] S. Cole, W. J. Percival, J. A. Peacock, P. Norberg, C. M. Baugh, C. S. Frenk, I. Baldry, J. Bland-Hawthorn, T. Bridges, R. Cannon, et al., Mon. Not. R. Astron. Soc. **362**, 505 (2005), arXiv:astro-ph/0501174.
[4] C. Blake, E. A. Kazin, F. Beutler, T. M. Davis, D. Parkinson, S. Brough, M. Colless, C. Contreras, W. Couch, S. Croom, et al., Mon. Not. R. Astron. Soc. p. 1598 (2011), 1108.2635.
[5] G. Efstathiou, S. Moody, J. A. Peacock, W. J. Percival, C. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, S. Cole, M. Colless, et al., Mon. Not. R. Astron. Soc. **330**, L29 (2002), arXiv:astro-ph/0109152.
[6] B. A. Reid, W. J. Percival, D. J. Eisenstein, L. Verde, D. N. Spergel, R. A. Skibba, N. A. Bahcall, T. Budavari, J. A. Frieman, M. Fukugita, et al., Mon. Not. R. Astron. Soc. **404**, 60 (2010), 0907.1659.
[7] H. A. Feldman, N. Kaiser, and J. A. Peacock, Astrophys. J. **426**, 23 (1994), arXiv:astro-ph/9304022.
[8] R. Scoccimarro, M. Zaldarriaga, and L. Hui, Astrophys. J. **527**, 1 (1999), arXiv:astro-ph/9901099.
[9] A. Meiksin and M. White, Mon. Not. R. Astron. Soc. **308**, 1179 (1999), arXiv:astro-ph/9812129.
[10] A. Cooray and W. Hu, Astrophys. J. **554**, 56 (2001), arXiv:astro-ph/0012087.
[11] R. Scoccimarro and R. K. Sheth, Mon. Not. R. Astron. Soc. **329**, 629 (2002), arXiv:astro-ph/0106120.
[12] E. Sirko, Astrophys. J. **634**, 728 (2005), arXiv:astro-ph/0503106.
[13] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, Phys. Rev. D **74**, 023522 (2006), arXiv:astro-ph/0604505.
[14] M. C. Neyrinck, I. Szapudi, and C. D. Rimes, Mon. Not. R. Astron. Soc. **370**, L66 (2006), arXiv:astro-ph/0604282.
[15] M. C. Neyrinck and I. Szapudi, Mon. Not. R. Astron. Soc. **375**, L51 (2007), arXiv:astro-ph/0610211.
[16] R. E. Smith, Mon. Not. R. Astron. Soc. **400**, 851 (2009), 0810.1960.
[17] M. Sato, T. Hamana, R. Takahashi, M. Takada, N. Yoshida, T. Matsubara, and N. Sugiyama, Astrophys. J. **701**, 945 (2009), 0906.2237.
[18] M. C. Neyrinck, Astrophys. J. **736**, 8 (2011), 1103.5476.
[19] J. Harnois-Deraps and U.-L. Pen, ArXiv e-prints (2011), 1109.5746.
[20] W.-H. W. Ngan, J. Harnois-Déraps, U.-L. Pen, P. McDonald, and I. MacDonald, ArXiv e-prints (2011), 1106.5548.
[21] A. J. S. Hamilton, C. D. Rimes, and R. Scoccimarro, Mon. Not. R. Astron. Soc. **371**, 1188 (2006), arXiv:astro-ph/0511416.
[22] C. D. Rimes and A. J. S. Hamilton, Mon. Not. R. Astron. Soc. **371**, 1205 (2006), arXiv:astro-ph/0511418.
[23] M. Takada and B. Jain, Mon. Not. R. Astron. Soc. **395**, 2065 (2009), 0810.4170.
[24] B. A. Reid, D. N. Spergel, and P. Bode, Astrophys. J. **702**, 249 (2009), 0811.1025.
[25] R. Takahashi, N. Yoshida, M. Takada, T. Matsubara, N. Sugiyama, I. Kayo, A. J. Nishizawa, T. Nishimichi, S. Saito, and A. Taruya, Astrophys. J. **700**, 479 (2009), 0902.0371.
[26] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Physics Reports **367**, 1 (2002), arXiv:astro-ph/0112551.
[27] J. N. Fry, Astrophys. J. **279**, 499 (1984).
[28] M. H. Goroff, B. Grinstein, S.-J. Rey, and M. B. Wise, Astrophys. J. **311**, 6 (1986).
[29] M. C. Neyrinck, I. Szapudi, and A. S. Szalay, Astrophys. J. Lett. **698**, L90 (2009), 0903.4693.
[30] R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce, P. A. Thomas, G. Efstathiou, and H. M. P. Couchman, Mon. Not. R. Astron. Soc. **341**, 1311 (2003), arXiv:astro-ph/0207664.
[31] A. Lewis, A. Challinor, and A. Lasenby, Astrophys.J. **538**, 473 (2000), arXiv:astro-ph/9911177.
[32] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolta, L. Page, et al., ArXiv e-prints (2010), arXiv:1001.4538.
[33] V. Springel, Mon. Not. R. Astron. Soc. **364**, 1105 (2005), arXiv:astro-ph/0505010.
[34] Y. P. Jing, Astrophys. J. **620**, 559 (2005), arXiv:astro-ph/0409240.
[35] M. Tegmark, D. J. Eisenstein, M. A. Strauss, D. H. Weinberg, M. R. Blanton, J. A. Frieman, M. Fukugita, J. E. Gunn, A. J. S. Hamilton, G. R. Knapp, et al., Phys.Rev.D **74**, 123507 (2006), arXiv:astro-ph/0608632.
[36] K. N. Abazajian, J. K. Adelman-McCarthy, M. A. Agüeros, S. S. Allam, C. Allende Prieto, D. An, K. S. J. Anderson, S. F. Anderson, J. Annis, N. A. Bahcall, et al., Astrophys. J. Suppl. **182**, 543 (2009), 0812.0649.
[37] J. Dunkley, E. Komatsu, M. R. Nolta, D. N. Spergel, D. Larson, G. Hinshaw, L. Page, C. L. Bennett, B. Gold, N. Jarosik, et al., Astrophys. J. Suppl. **180**, 306 (2009), 0803.0586.