Cosmological Inference from Galaxy-Clustering Power Spectrum

Gaussianisation and Covariance Decomposition

Mike Shengbo Wang | `Mike.Wang@port.ac.uk` | `arXiv:1811.08155`

In collaboration with W Percival, S Avila, R Crittenden & D Bianchi

Euclid Consortium UK Meeting 2018 (*Oxford, 17 December*)

Information in the Power Spectrum

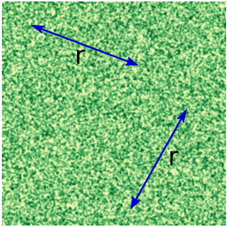

Figure 1. Correlators at separation r in a Gaussian random field $\delta(\mathbf{r})$.

2-point correlators:

$$\xi(\mathbf{r}) = \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) \rangle \underset{\mathscr{F}^{-1}}{\overset{\mathscr{F}}{\rightleftharpoons}} P(\mathbf{k}) = \left\langle \delta(\mathbf{k})\delta^*(\mathbf{k}) \right\rangle$$

$$\text{excess probability} \rightleftharpoons \text{amount of structure}$$

Gaussian random field (GRF):

$$\mathrm{Re}\left\{\delta(\mathbf{k})\right\}, \mathrm{Im}\left\{\delta(\mathbf{k})\right\} \sim \mathrm{N}\left(0, P(\mathbf{k})/2\right)$$

All information about a GRF is encoded in its 2-point statistics:

$$\text{observable} \longleftarrow \boxed{P_{\mathrm{g}}(k) \propto b^2 P_{\mathrm{m}}(k) \propto |T(k)|^2 P_{\mathcal{R}}(k)} \longleftarrow \text{cosmological parameters}$$

Likelihood Analysis

Gaussian Likelihood / Covariance Estimation

Central limit theorem: $\mathbf{Z} \mid \theta \sim \mathrm{N}(\boldsymbol{\mu}(\theta), \Sigma(\theta))$ with

$$\mathcal{L}(\theta; \mathbf{Z}) = |2\pi\Sigma|^{-1/2} \exp\left[-\frac{1}{2}\chi^2(\mathbf{Z}; \boldsymbol{\mu}, \Sigma)\right].$$

Covariance estimation:

- Accuracy? $\mathbb{E}\left[\widehat{\Sigma}\right] = \Sigma \not\Longrightarrow \mathbb{E}\left[\widehat{\Sigma}^{-1}\right] = \Sigma^{-1}$!
- Precision? $\mathrm{Var}\left[\widehat{\Sigma}\right] \neq 0$!
- Computation? $\widehat{\Sigma} \equiv \widehat{\Sigma}(\theta)$!

Statistical efforts:

- Debiasing (Hartlap+, 2007; Taylor+, 2013);
- Error propagation (Dodelson & Schneider, 2013; Percival+, 2014);
- Shrinkage (Pope & Szapudi, 2008) / tapering (Paz & Sánchez, 2015);
- Bayesian marginalisation (Sellentin & Heavens, 2016)...

Breakdown of Asymptotic Normality

Two factors—

1) non-normal distributions;
2) limited sample size.

Two consequences—

1) biased estimate;
2) distorted error bounds.

This is the case for the linear power spectrum near the survey scale.

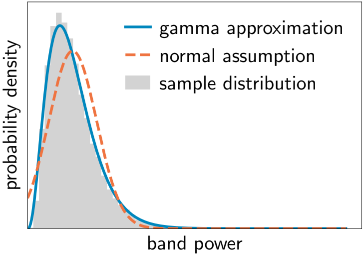

Figure 2. Comparison of distributions with the same mean and variance for the windowed power spectrum monopole.

Power Spectrum Analysis

Feldman–Kaiser–Peacock Framework

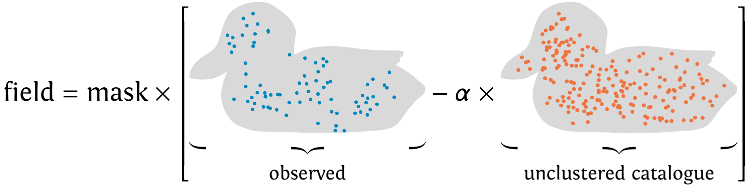

Figure 3. A weighted mask $w(\mathbf{r})$ is applied to the centred galaxy number density field.

Weighted density field: given a weight $w(\mathbf{r})$,

$$F(\mathbf{r}) \equiv w(\mathbf{r})\left[n_{\mathrm{g}}(\mathbf{r}) - \alpha n_{\mathrm{s}}(\mathbf{r})\right],$$

where $\bar{n}(\mathbf{r}) = \langle n_{\mathrm{g}}(\mathbf{r})\rangle = \alpha \langle n_{\mathrm{s}}(\mathbf{r})\rangle$ (Feldman+, 1994).

Window Effects

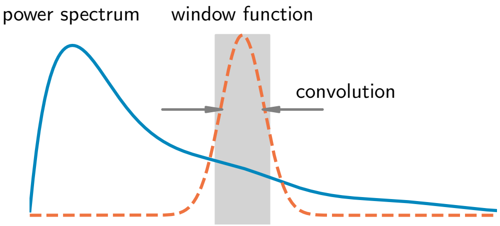

Figure 4. The window function (- -) brings power (—) from outside the band range into the band ( ).

Window function correlates the underlying power spectrum at different wave numbers:

$$\widetilde{P}(\mathbf{k}) = \int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} \mathcal{W}(\mathbf{k}-\mathbf{q}) P(\mathbf{q}) \,, \quad \mathcal{W} = \mathscr{F}\{w\bar{n}\}^2 \,.$$

Shot Noise

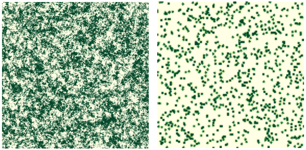

Figure 5. Discrete and 'rare' galaxy formation gives rise to a 'white' background that overlays the underlying structure.

Galaxy formation is a spatially independent Poisson point process*:

$$n_{\mathrm{g}}(\mathbf{r}) \sim \mathrm{Poisson}\left(\bar{n}[1 + \delta(\mathbf{r})]\right) .$$

Shot noise is an additional scale-independent* component:

$$\widetilde{P}(\mathbf{k}) \mapsto \widetilde{P}(\mathbf{k}) + P_{\mathrm{shot}} , \quad P_{\mathrm{shot}} \propto \bar{n}^{-1} .$$

* Non-Poissonian and scale-dependent on small scales due to non-linear clustering and exclusion effects; see e.g. Baldauf+ (2013).

3-D Power Distribution

Band power: spherical average[†] in a shell V_k

$$\bar{P}(k) = \int_{V_k} \frac{\mathrm{d}^3\mathbf{k}}{V_k} \widetilde{P}(\mathbf{k}) \quad \leftrightarrow \quad \bar{\mathbf{P}} = \mathsf{W}\big(\mathbf{P} + P_{\mathrm{shot}} \mathbf{1}_p\big) \,.$$

Exponential mixture: exponentially-distributed mode power and shot noise (Peacock & Nicholson, 1991) are mixed into an approximate gamma distribution

$$\widehat{\bar{P}} \sim \Gamma(R, \eta)$$

with shape–scale parameters (R, η).

[†] Identified with windowed power spectrum monopole by noting that the zeroth Legendre polynomial is $L_0 \equiv 1$.

Power-Spectrum Likelihood Analysis

Gaussianising Data Transformation

Curse of dimensionality: absence of an appropriate multivariate gamma distribution.

Univariate Gaussianisation: an old friend... See e.g. Smith+ (2006), Joachimi & Taylor (2011) and Wilking & Schneider (2013).

"Box–Cox" transformation

$$\widehat{\bar{P}} \mapsto Z \equiv \widehat{\bar{P}}^{\nu}, \quad \nu(R) \approx 1/3\,.$$

Covariance Treatment

Variance–correlation decomposition

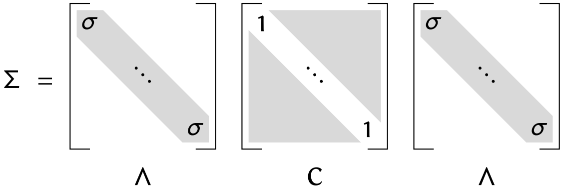

Figure 6. Schematic decomposition of a covariance matrix Σ into cosmology-dependent diagonal variances Λ^2 and cosmology-independent off-diagonal correlation C.

Variance–correlation decomposition

Given an estimated covariance $\widehat{\Sigma}_f$ at fiducial model θ_f, rescale by

$$\widehat{\Sigma}(\theta) = \Lambda(\theta)\Lambda_f^{-1}\widehat{\Sigma}_f\Lambda_f^{-1}\Lambda(\theta)\,.$$

Covariance marginalisation (Sellentin & Heavens, 2016): given an estimate $\widehat{\Sigma} \in \mathbb{R}^{p \times p}$ from $(m + 1)$ mock catalogue samples

1) recognise the Wishart distribution $\widehat{\Sigma} \big| \Sigma \sim \mathrm{W}_p\,(\Sigma/m, m)$;
2) choose the uninformative Jeffreys prior $\Pi(\Sigma) = |\Sigma|^{-(p+1)/2}$;
3) deduce the inverse Wishart posterior $\Sigma \big| \widehat{\Sigma} \sim \mathrm{W}_p^{-1}\left(m\widehat{\Sigma}, m\right)$.

Modified *t*-likelihood

Marginalising over the inverse Wishart posterior,

$$2 \log \mathcal{L}(\theta; \mathbf{Z}) = \log \left|\widehat{\Sigma}\right| - (m + 1) \log \left[1 + m^{-1} \chi^2 (\mathbf{Z}; \boldsymbol{\mu}, \widehat{\Sigma})\right] .$$

Results from Numerical Tests

Testing for Distributions

- ✓ Multivariate normality is improved in a frequentists' test.
- ✓ Rescaled covariance agrees with directly sampled covariance.

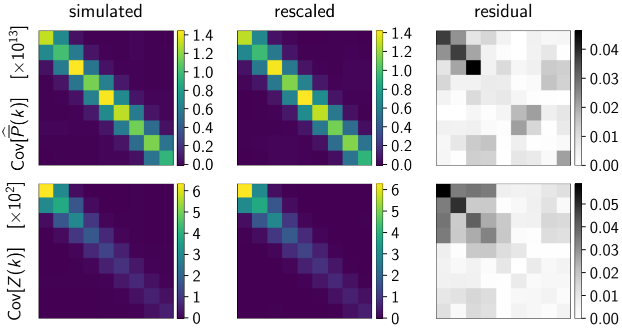

Figure 7. Comparison of covariance matrices estimated at 'true' cosmology (*left*) and rescaled covariance estimated at fiducial cosmology (*middle*) with their residuals (*right*).

Testing for Parameter Inference

The local non-Gaussianity parameter f_{NL} modifies the linear bias (Slosar+, 2008)

$$b \mapsto b + f_{\mathrm{NL}} A(k)(b-1)\,, \quad A(k) \propto k^{-2} T(k)^{-1}\,,$$

so is a sensitive parameter to the large-scale power spectrum.

With Gaussianisation and covariance rescaling, we find for

- ✓ **Bayesian point estimation**: improved parameter estimates and associated uncertainties;
- ✓ **posterior shape**: improved Kullback–Leibler divergence from the posterior of the exact likelihood.

Conclusions: Simply Gaussianise and Rescale!

A visual summary

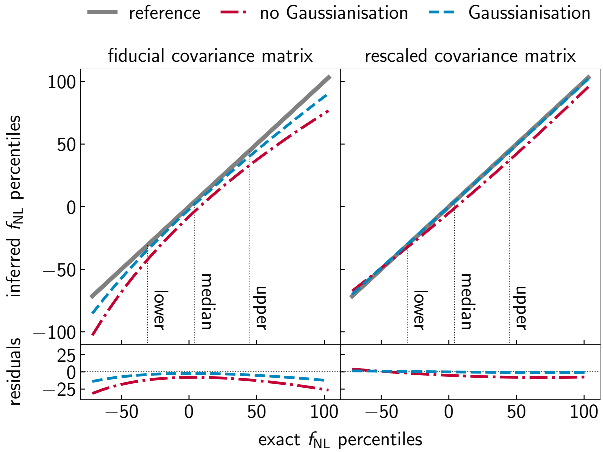

Figure 8. Inferred f_{NL} percentiles with or without Gaussianisation/covariance rescaling against percentiles of the exact posterior, with the median estimate and its uncertainty bounds marked.

Acknowledgements. The speaker thanks his supervisory team for their guidance. Numerical computations are performed on the Sciama High Performance Computing (HPC) cluster which is supported by the Institute of Cosmology and Gravitation (ICG), the South East Physics Network (SEPnet) and the University of Portsmouth.

Question Time

[1811.08155]

References i

J. Hartlap, P. Simon & P. Schneider. *Astron. Astrophys.*, 464(1):399–404, 2007. [astro-ph/0608064].

A. Taylor, B. Joachimi & T. Kitching. *Mon. Notices Royal Astron. Soc.*, 432(3):1928–1946, 2013. [astro-ph.CO/1212.4359].

S. Dodelson & M. D. Schneider. *Phys. Rev.*, D88:063537, 2013. [astro-ph.CO/1304.2593].

W. J. Percival et al. *Mon. Notices Royal Astron. Soc.*, 439(3):2531–2541, 2014. [astro-ph.CO/1312.4841].

A. C. Pope & I. Szapudi. *Mon. Notices Royal Astron. Soc.*, 389(2):766–774, 2008. [astro-ph/0711.2509].

D. J. Paz & A. G. Sánchez. *Mon. Notices Royal Astron. Soc.*, 454(4):4326–4334, 2015. [astro-ph.CO/1508.03162].

E. Sellentin & A. F. Heavens. *Mon. Notices Royal Astron. Soc.*, 456(1):L132–L136, 2016. [astro-ph.CO/1511.05969].

References ii

H. A. Feldman, N. Kaiser & J. A. Peacock. *Astrophys. J.*, 426(1):23–37, 1994. [astro-ph/9304022].

T. Baldauf, U. Seljak, Robert E. Smith, N. Hamaus & V. Desjacques. *Phys. Rev.*, D88(8): 083507, 2013. [astro-ph.CO/1305.2917].

J. A. Peacock & D. Nicholson. *Mon. Notices Royal Astron. Soc.*, 253:307–319, 1991.

S. Smith, A. Challinor & G. Rocha. *Phys. Rev.*, 73:023517, 2006. [astro-ph/0511703].

B. Joachimi & A. N. Taylor. *Mon. Notices Royal Astron. Soc.*, 416(2):1010–1022, 2011. [astro-ph.CO/1103.3370].

P. Wilking & P. Schneider. *Astron. Astrophys.*, 556:A70, 2013. [astro-ph.CO/1304.4781].

A. Slosar, C. Hirata, U. Seljak, S. Ho & N. Padmanabhan. *J. Cosmol. Astropart. Phys.*, 0808:031, 2008. [astro-ph/0805.3580].