Cosmological Inference from Galaxy-Clustering Power Spectrum

Gaussianisation and Covariance Decomposition

Mike Wang (王圣博) | ICG Portsmouth | Mike.Wang@port.ac.uk

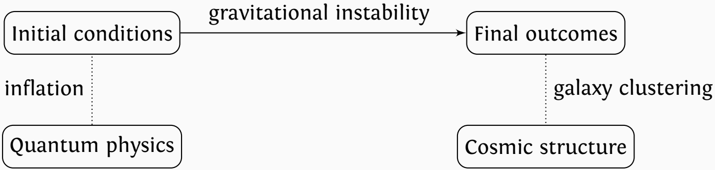
In collaboration with W Percival, R Crittenden, S Avila & D Bianchi

4th CINY Joint Workshop on Frontiers of Cosmology (*Beijing, 30 October 2018*)

Large-Scale Structure

An initial-value and inverse problem

Predictions are made by modelling the non-linear evolution.



Inference is performed by statistical analysis of the measurements.

Bayes' theorem

$$\mathcal{P}(\theta \,|\, \mathbf{X}) \propto \Pi(\theta)\, \mathcal{L}(\theta; \mathbf{X}),$$

where $(\Pi, \mathcal{L}, \mathcal{P})$ are the (prior, likelihood, posterior) of model parameters θ and measurements $\mathbf{X}$.

Necessity of a Statistical Approach

The statistical nature is manifold—

1) inherently probabilistic quantum theories;
2) emergent phenomena from complex microscopic processes;
3) intrinsically random measurements.

> *Cosmologists are often in error but seldom in doubt.*
>
> — Lev Landau (1908–1968)

Q. *How could one possibly make fair comparison between predictions and observations?*

A. *Forward modelling*!

Likelihood Analysis

Gaussian Likelihood

A 'universality' law

Gaussian assumption: $\mathbf{Z} \mid \theta \sim \mathrm{N}_p\left(\boldsymbol{\mu}(\theta), \Sigma(\theta)\right)$ with

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

where $\chi^2 \equiv (\mathbf{Z} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{Z} - \boldsymbol{\mu})$ and

- $\boldsymbol{\mu}(\theta)$ is the mean vector given by a parametric model;
- $\Sigma(\theta)$ is the covariance matrix.

Analytically simple and usually justified by…

Central limit theorem

Let $\{\mathbf{Z}_i\}_{i=1}^{N}$ be independent random variables, then (under regularity conditions) their average $\bar{\mathbf{Z}}$ converges in distribution to a multivariate normal variable in the asymptotic limit $N \to \infty$.

Covariance Estimation

From analytic formulae to mock catalogues

Three considerations—**accuracy**, **precision** and **computational costs**.

Three complications—

1) Precision matrix **bias**: $\mathbb{E}\left[\widehat{\Sigma}\right] = \Sigma \;\not\Rightarrow\; \mathbb{E}\left[\widehat{\Sigma}^{-1}\right] = \Sigma^{-1}$;
2) Random **scatter**: $\mathrm{Var}\left[\widehat{\Sigma}\right] \neq 0$;
3) Parameter **dependence**: $\Sigma \equiv \Sigma(\theta)$.

Numerous (statistical) efforts—

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

Breakdown of Asymptotic Normality

Two factors—
1) non-normal distribution
and 2) limited sample size.

Two consequences* –
1) biased estimate and
2) distorted error bounds.

This is the case for galaxy-clustering power spectrum on the largest survey scales!

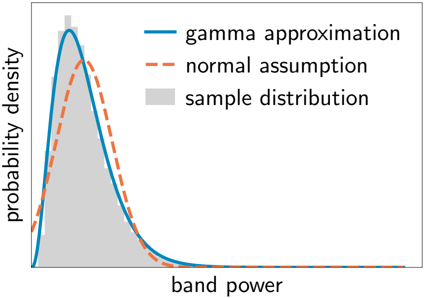


Figure 1. Comparison of distributions with the same mean and variance for the windowed power spectrum monopole.

* Cf. Kalus+ (2016) and Sellentin & Heavens (2018).

Power Spectrum Analysis

Homogeneous Cosmic Fields

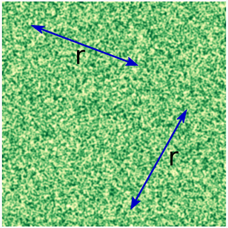


Figure 2. Correlators at separation r in a Gaussian random field $\delta(\mathbf{r})$.

2-point correlators:

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

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

Gaussian random field (GRF) $\delta(\mathbf{r})$:

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

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

A wealth of information:

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

Linear evolution of density perturbations on large scales *preserves Gaussian statistics* of initial conditions.

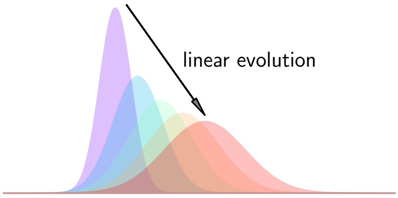


Figure 3. Linear evolution of a Gaussian random variable.

Feldman–Kaiser–Peacock Framework

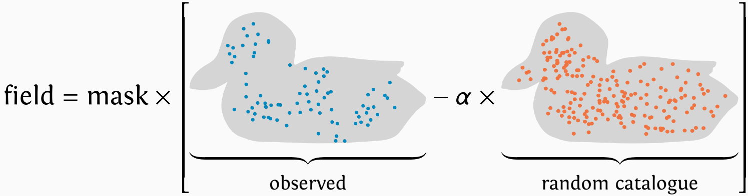


Figure 4. A mask $w(\mathbf{r})$, with $w(\mathbf{r}) = 0$ outside the survey volume, is applied to the de-meaned galaxy number density field.

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 α matches the observed and random catalogue mean number densities $\bar{n}(\mathbf{r}) = \langle n_{\mathrm{g}}(\mathbf{r})\rangle = \alpha \langle n_{\mathrm{s}}(\mathbf{r})\rangle$ (Feldman+, 1994).

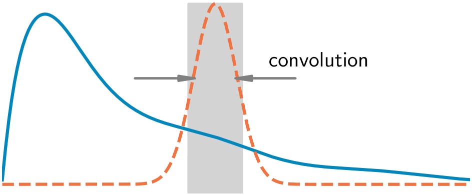


Figure 5. A broad window function (- -) brings power spectrum (—) from outside the band range into the band ( ).

Window function processes the underlying power spectrum:

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

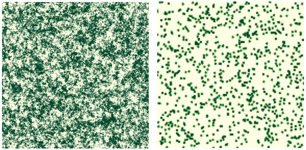


Figure 6. Discrete and 'rare' galaxy formation gives rise to a 'white' background that overlays the underlying structure.

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:

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

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

Band Power and Its Non-normal 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; Kalus+, 2016) 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 Transformation

Back to normal

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

Univariate Gaussianisation: an old friend... rekindled; see e.g. Joachimi & Taylor (2011) and Wilking & Schneider (2013).

Power transformation

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

- Inspired by the Box–Cox transformation;
- Motivated by suppression of higher-order moments;
- Justified by the approximate equivalence to multivariate Gaussianisation for weak correlation.

Covariance Treatment

Variance–correlation decomposition

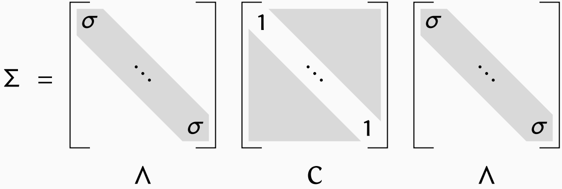


Figure 7. A covariance matrix Σ consists of cosmology-dependent diagonal variances Λ^2 and cosmology-independent off-diagonal correlation coefficients C.

Variance–correlation decomposition

Given an estimated covariance $\widehat{\Sigma}_\mathrm{f}$ matrix at fiducial cosmology θ_f, rescale by

$$\widehat{\Sigma}(\theta) = \Lambda(\theta)\Lambda_\mathrm{f}^{-1}\widehat{\Sigma}_\mathrm{f}\Lambda_\mathrm{f}^{-1}\Lambda(\theta)$$

to account for cosmological dependence of the covariance matrix.

Covariance Treatment

Bayesian marginalisation

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

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

Modified t-likelihood

Marginalising the Gaussian likelihood over the inverse Wishart posterior and dropping a normalisation constant, we obtain

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

Results from Numerical Tests

Testing for Distributions

✓ Gaussianisation tested: improved p-value in a frequentists' test for multivariate normality.

✓ Covariance rescaling tested:

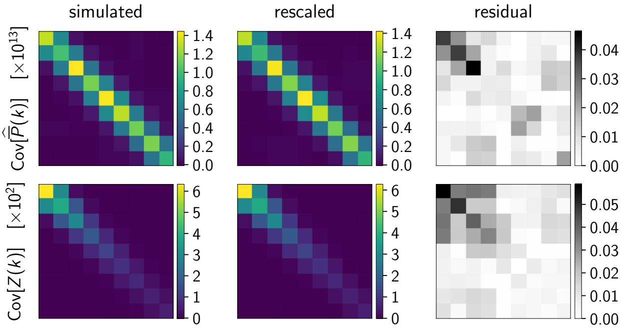


Figure 8. Comparison of covariance matrices estimated at 'true' cosmology (*left*) and rescaled covariance estimated at fiducial cosmology (*middle*) with their residuals (*right*).

Testing for Parameter Inference

We choose the local non-Gaussianity f_{NL} as a test parameter; it enters the power spectrum by modifying 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 large-scale measurements.

With Gaussianisation and covariance rescaling, we find for

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

Conclusions

A visual summary

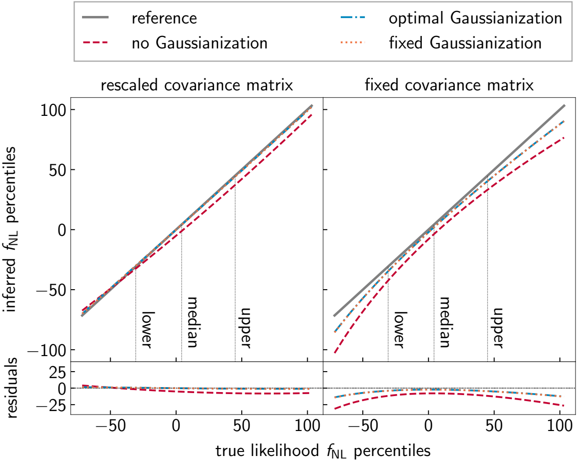


Figure 9. Inferred percentile–percentile plots (*top panel*) comparing three different posteriors with the true posterior (*unit-slope reference line*), for either the rescaled covariance estimate or the correct fixed fiducial covariance estimate. The dotted *vertical lines* show the median estimate and its 1-σ bounds inferred from the true posterior.

Acknowledgements. The presenter 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

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

B. Kalus, W. J. Percival & L. Samushia. *Mon. Notices Royal Astron. Soc.*, 455(3):2573–2581, 2016. [astro-ph.CO/1504.03979].

E. Sellentin & A. F. Heavens. *Mon. Notices Royal Astron. Soc.*, 473(2):2355–2363, 2018. [astro-ph.CO/1707.04488].

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.

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].

Backup Slides

Backup materials.