Cosmology beyond the power spectrum

Cosmological constraint with the galaxy power spectrum, bispectrum and parity-even trispectrum measured from differentiable DISCO-DJ simulations

Phu-Huy Nguyen[1,2] · Jiamin Hou[1]

[1] University Observatory, Ludwig-Maximilians-Universität München [2] Dept. of Physics, University of Rome Tor Vergata · Erasmus Mundus MASS

01 Why go beyond two-point statistics

Gravity drives an initially near-Gaussian density field into a strongly non-Gaussian one. On mildly non-linear scales the power spectrum $P(k)$ no longer carries the full information content of the galaxy field: the clustering amplitude stays degenerate with galaxy bias, and the expansion history is only weakly constrained.

Higher-order statistics respond to those degeneracies differently. The bispectrum and the parity-even trispectrum break the b_1–σ_8 degeneracy and add non-Gaussian information — at the cost of a data vector that grows from tens to hundreds of configurations.

That growth is exactly where the standard Gaussian-likelihood analysis fails. The likelihood of higher-order statistics is measurably non-Gaussian, and estimating an $n_d \times n_d$ covariance from simulations demands $N_{sim} \gg n_d$ realisations, paying the Dodelson–Schneider penalty on every error bar.

> We drop the assumed likelihood entirely: simulation-based inference on *compressed* higher-order summaries, trained on a differentiable forward model.

02 The summary statistics

POWER SPECTRUM

$\langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\rangle = (2\pi)^3\delta_D(\mathbf{k}_{12})\, P(k_1)$

BISPECTRUM — TRIANGLES

$\langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle = (2\pi)^3\delta_D(\mathbf{k}_{123})\, B(k_1,k_2,k_3)$

PARITY-EVEN TRISPECTRUM — QUADRILATERALS

$\langle\delta(\mathbf{k}_1)\ldots\delta(\mathbf{k}_4)\rangle_c = (2\pi)^3\delta_D(\mathbf{k}_{1234})\, T(k_1,k_2,k_3,k_4;\, K)$

Triangle and quadrilateral configurations (the latter labelled by an internal diagonal K) are binned to a joint data vector of `n_d = 903` elements (27 + 378 + 498) out to `k_max = [0.4] h/Mpc for P and B, and to k_max = [0.2] h/Mpc for T`, estimated with FFT-based separable estimators.

03 DISCO-DJ forward model

Training sets are generated with **DISCO-DJ**, a JAX-based differentiable cosmology code: a differentiable Einstein–Boltzmann solver supplies the linear input, and a differentiable particle-mesh N-body solver evolves it to the output redshift.

Because runs are batched on GPU, the thousands of realisations that SBI needs are affordable; because the model is end-to-end differentiable, gradients $\partial\mathbf{d}/\partial\boldsymbol{\theta}$ come for free and feed the score compression directly — no finite differences, no extra derivative simulations.

Box size `1000 Mpc/h`	Mesh / particles `256³`
Output redshift `z = 0.5`	Latin hypercube `[N_LH] cosmologies`
Fiducial set `2000 realisations`	Training set `48000 realisations`
Validation set `2000 realisations`	Sampled parameters $\Omega_c,\ \Omega_b,\ h,\ n_s,\ \sigma_8$

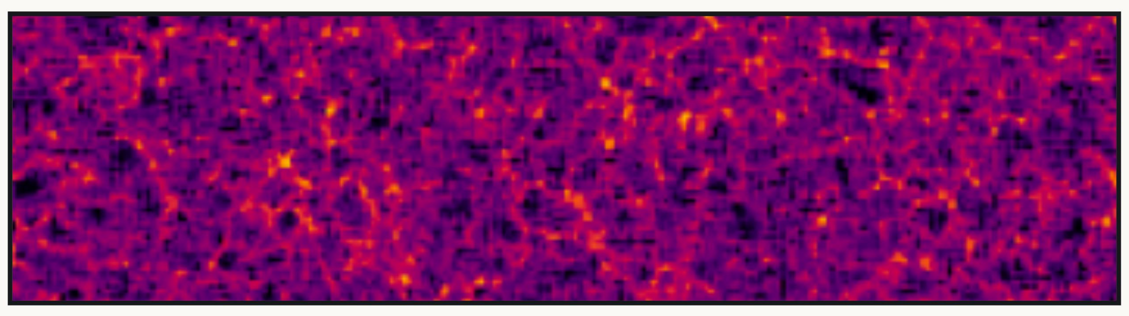
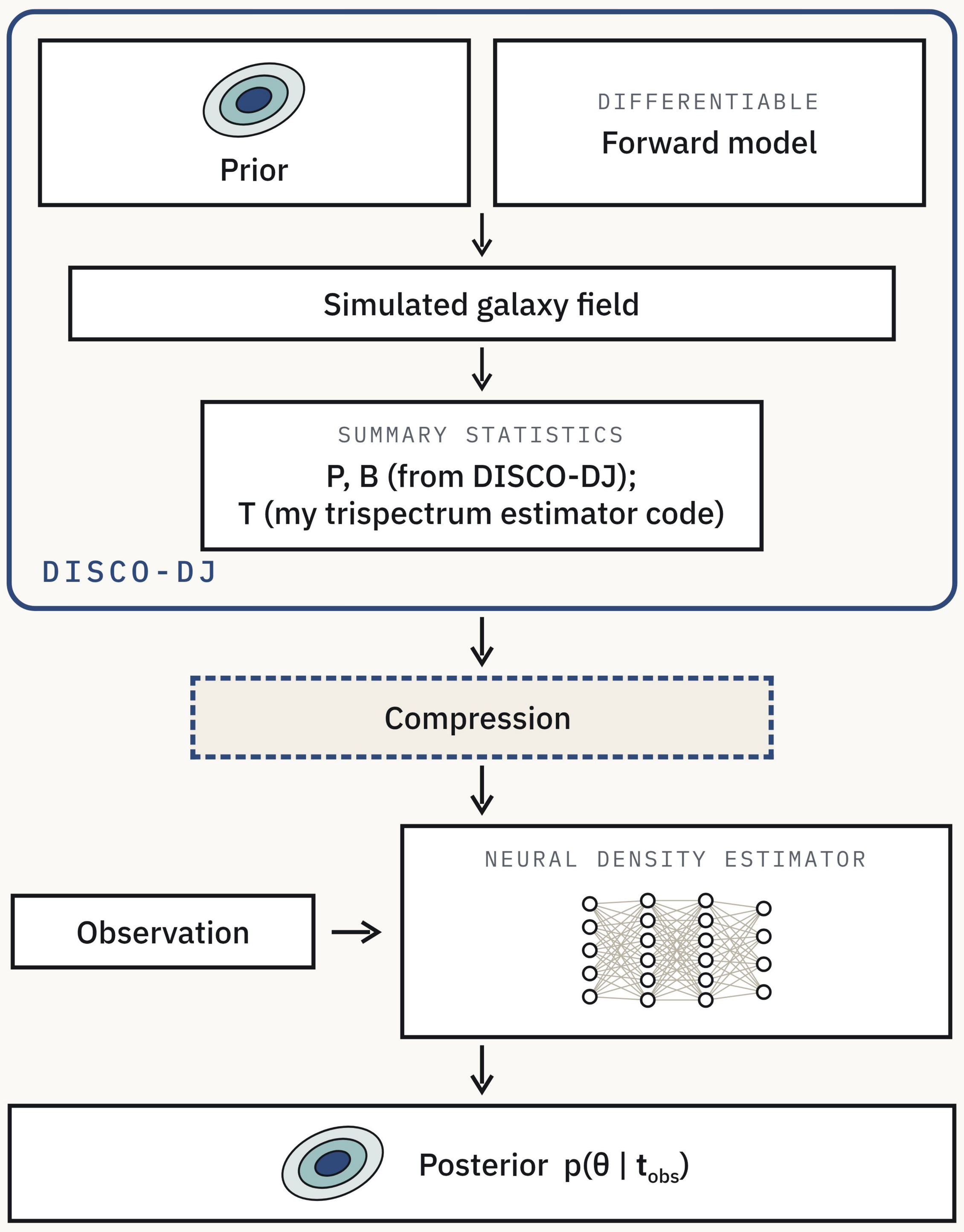


Fig. 1. Projected density field of one DISCO-DJ realisation at the output redshift.

04 The inference pipeline



05 Three ways to compress B

Appended raw, the bispectrum monopole multiplies the length of the data vector — and the covariance noise with it. We compare three compressions of B against two baselines: P(k) alone, and P(k) with B raw.

A **PCA.** Keeps the leading principal components of B_0 — two of them here. Parameter-agnostic: it needs the covariance, not the derivatives.

B **MOPED (score) compression.** One linear summary per parameter — five here — weighted by $\partial B/\partial\theta$ and $\mathbf{C}^{-1}$; optimal for a Gaussian likelihood with a known covariance.

C **Geometrical compression.** Groups triangle configurations by their geometry and keeps a few summaries per group — analytic and parameter-agnostic, but the grouping is chosen rather than learned.

06 Merit and bias together

Constraining power alone is no verdict: a lossy summary can look sharp and sit off-centre. We read both from the same runs.

FOM	FOB
$= 1 / (\det \mathbf{C}_\theta)^{1/2}$ how tight the posterior is	$\langle FoB^2\rangle/5 = \Delta\theta^T\mathbf{C}_\theta^{-1}\Delta\theta\, /\, 5$ how far off-centre it sits — 1 if unbiased

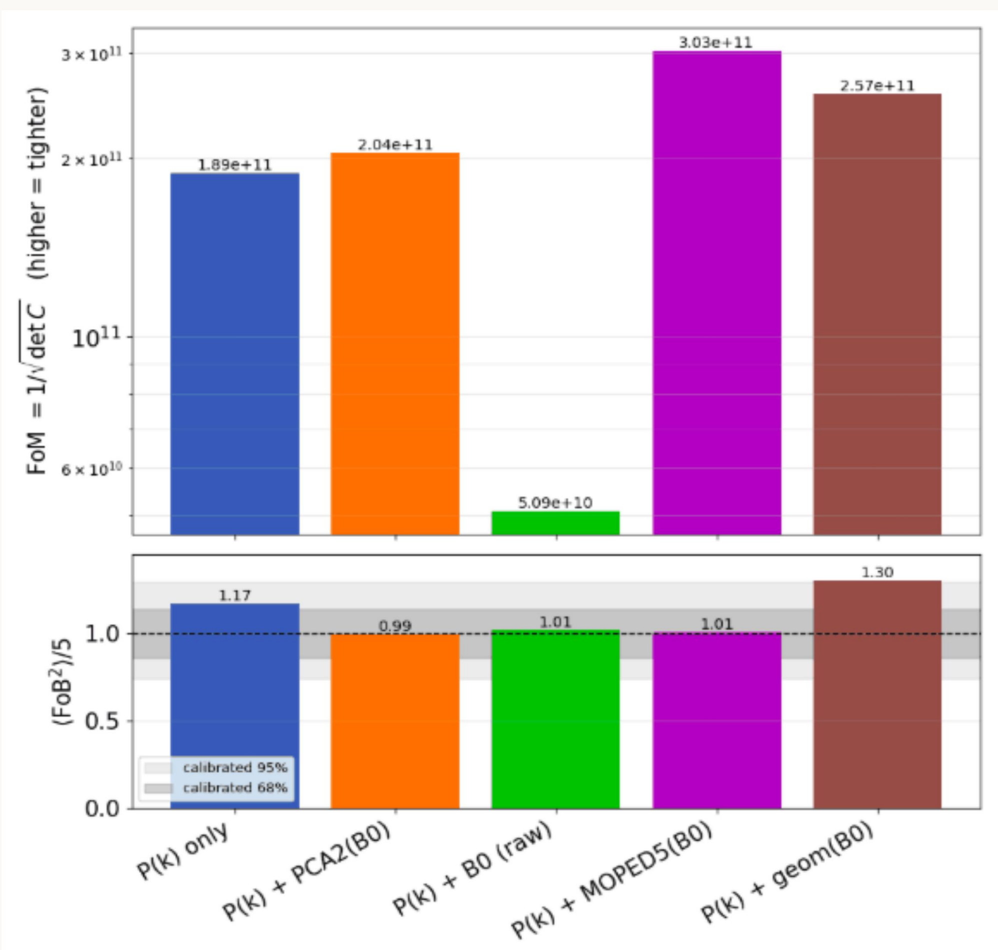


Fig. 2. Figure of merit (top) and figure of bias (bottom) for each compression of B_0. Bands mark the calibrated 68% and 95% ranges; a scheme earns its place only where the FoM rises while $\langle FoB^2\rangle/5$ stays inside them.

07 Fisher forecast for P + B + T

A debiased Fisher forecast on the same measurements bounds what the higher-order statistics can deliver, before any density estimator is trained.

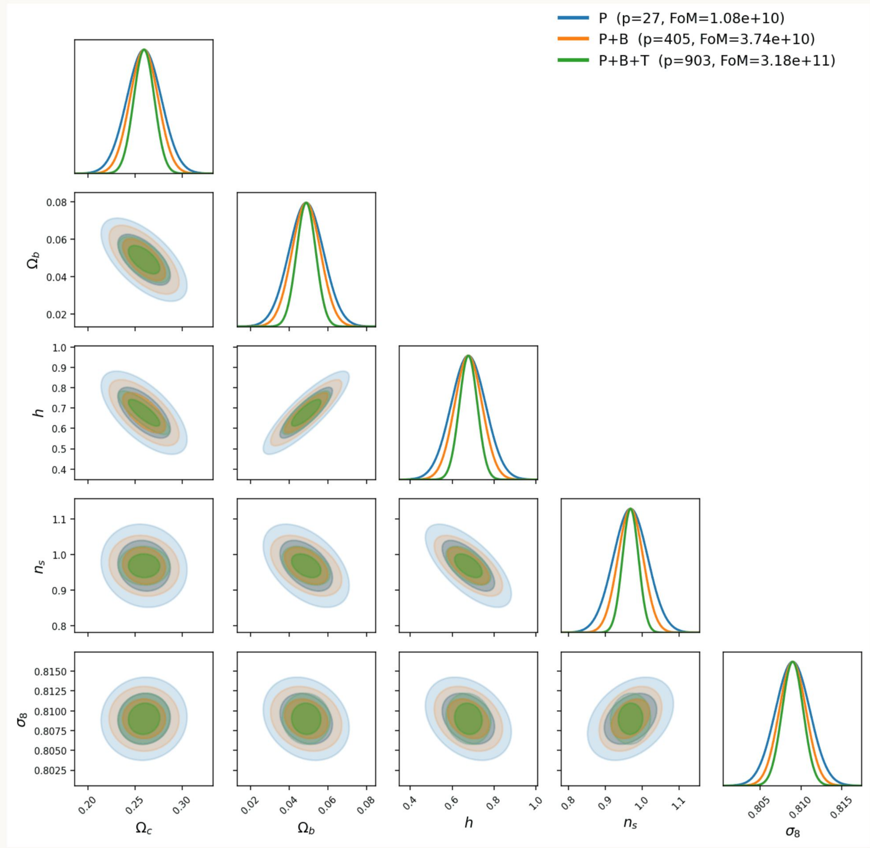


Fig. 3. Debiased Fisher forecast. Inner and outer contours are the 1σ and 2σ two-dimensional regions; *p* is the length of the data vector.

DATA VECTOR	P	FOM	GAIN
P	27	1.08×10^{10}	–
P + B	405	3.74×10^{10}	×3.5
P + B + T	903	3.18×10^{11}	×29

08 Is the posterior calibrated?

A tight, unbiased posterior is worthless if its credible intervals miss the truth at the stated rate. On held-out simulations we count how often the truth falls inside the α credible region.

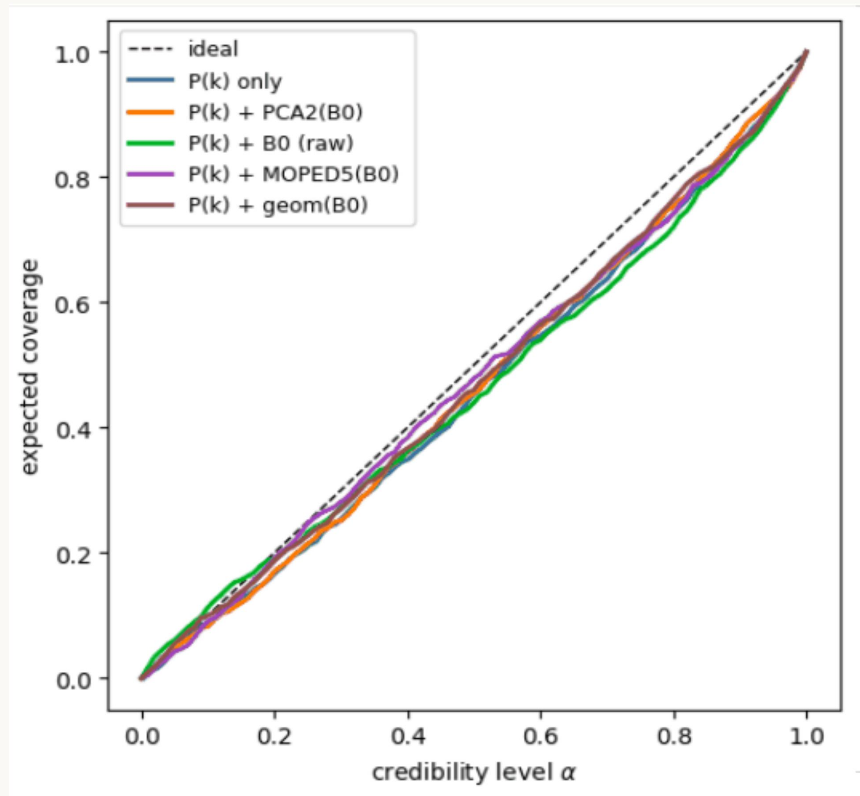


Fig. 4. Expected coverage against credibility level α; the dashed diagonal is perfect calibration. All five configurations track it closely, sitting marginally below it — a mild over-confidence shared by every data vector, not an artefact of one compression.

09 Takeaways

- In the Fisher forecast the bispectrum raises the figure of merit by ×3.5 over P alone, and the parity-even trispectrum takes it to ×29.
- Adding raw B make the inference worse, maybe due to the noisy data vector of B. Compressing B makes it better
- **Next:**
  - GPU computing for trispectrum estimator to later doing inference with T

REFERENCES

[1] List, Hahn & Rampf, *DISCO-DJ I: a differentiable Einstein–Boltzmann solver for cosmology*, JCAP 06 (2024) 063, arXiv:2311.03291

[2] List et al., *DISCO-DJ II: a differentiable particle-mesh code for cosmology*, arXiv:2510.05206

[3] Alsing & Wandelt, *Generalized massive optimal data compression*, MNRAS 476 (2018) L60

[4] Gualdi, Manera, Joachimi & Lahav, *Geometrical compression: a new method to enhance the BOSS galaxy bispectrum monopole constraints*, MNRAS 484 (2019) L29, arXiv:1901.00987

[5] Dodelson & Schneider, *The effect of covariance estimator error on cosmological parameter constraints*, PRD 88 (2013) 063537

Work carried out at the University Observatory, LMU Munich. Computing resources: Excellence Cluster ORIGINS.

huyphunguyen2808@gmail.com