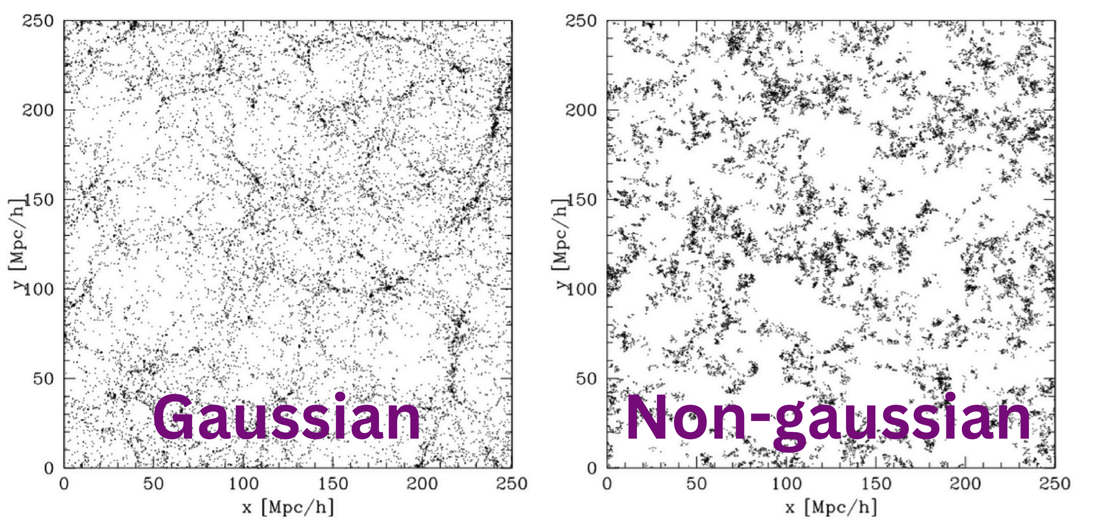
Machine learning solution for enabling cosmological analysis with the matter anisotropic 3PCF

Kristers Nagainis[1], Massimo Guidi[1,2], Sofia Contarini[3], Michele Moresco[1,2], Antonio Farina[4], Alfonso Veropalumbo[2,4]

1. INTRODUCTION

- Large scale structures of the universe contain a wealth of information about cosmological parameters, that can be extracted with **clustering statistics**.
- Two-point statistics: the power spectrum in Fourier space and the two-point correlation function (2PCF) in configuration space completely describe a Gaussian density field.
- Three-point statistics: the **three-point correlation function (3PCF)** and the bispectrum are the higher-order statistics with the highest signal-to-noise ratio being able to capture **non-Gaussianity** in the cosmological density field.[1]



- N-point correlation functions have an advantage over Fourier space of **not needing to deal with mode coupling** (e.g. holes in the survey) introduced by the window function.
- Due to the long computation[2], 3PCF has been mostly used only for template fitting.
- To obtain 3PCF model's prediction of the matter 3PCF for all selected triplets beyond monopole, it takes over 60 minutes on 48CPU.
- The aim of this work is to **enable cosmological parameter analysis** with 3PCF, which can be achieved by emulating the 3PCF.

2. FORMALISM

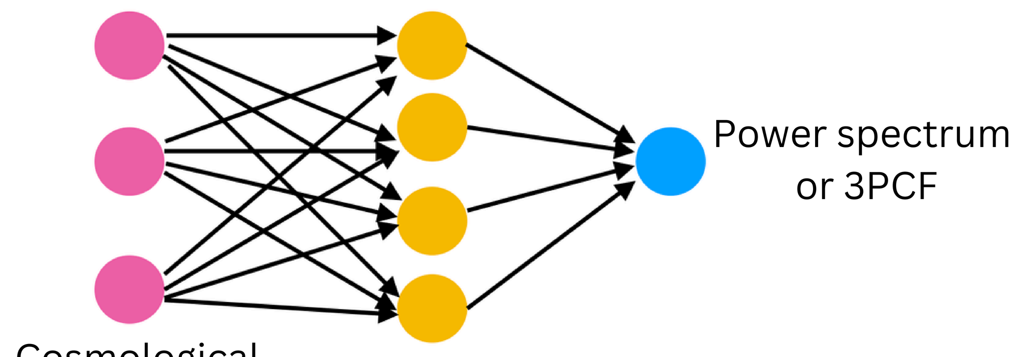
- A great way to deal with measurements - especially in redshift space - is to decompose the bispectrum and 3PCF in all triangle configurations isotropically and anisotropically.[3][4]
- 3PCF can be **decomposed into multipoles**, using the tri-polar spherical harmonics:

$$\zeta(\mathbf{r_1}, \mathbf{r_2}, \hat{n}) = \sum_{\ell_1+\ell_2+L=even} \zeta_{\ell_1,\ell_2,L}(r_1, r_2) S_{\ell_1,\ell_2,L}(\hat{r_1}, \hat{r_2}, \hat{n})$$

- Here ℓ_1 is the multipole for the corresponding $\mathbf{r_1}$, the ℓ_2 for $\mathbf{r_2}$ and L is for the line of sight vector $\hat{n}$, but the $S_{\ell_1,\ell_2,L}$ is the TripoSH base.
- For example, if L is zero, it is the monopole or the isotropic component.
- If L is non zero, that is the anisotropic case, which is only induced by the RSD and AP effects.
- To go from bispectrum multipoles to 3PCF multipoles, Hankel transformation must be used.
- The **model includes infrared re-summation** for BAO scales, that depends on the P_{lin} and P_{nw}.

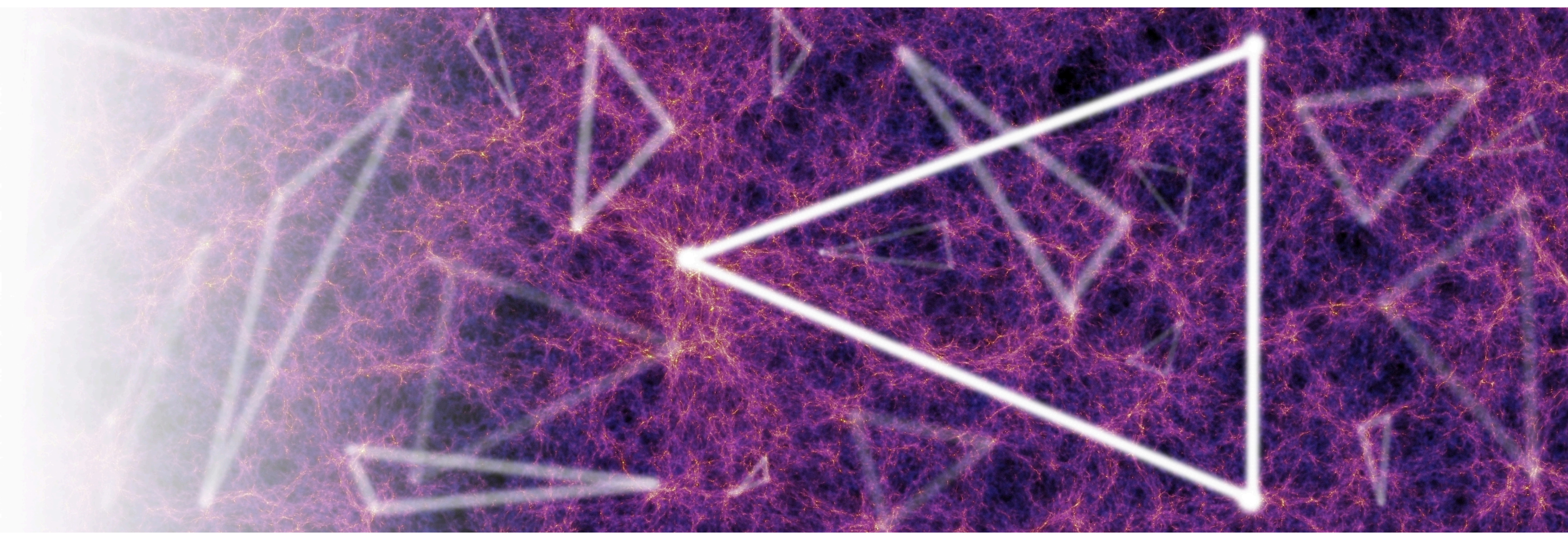
3. EMULATION

- Emulator - a fast prediction machine based on neural network.
- The emulator is constructed on the basis of **CosmoPower**.[6]



- Even parameter space distribution is achieved by **Symmetric Latin Hypercube**.
- The amount of training data and hyperparameters must be chosen carefully for each scientific case.
- Model accuracy can be **estimated by percentage difference** or by calculating the ratio between emulated and testing quantities in comparison with theoretical errors.

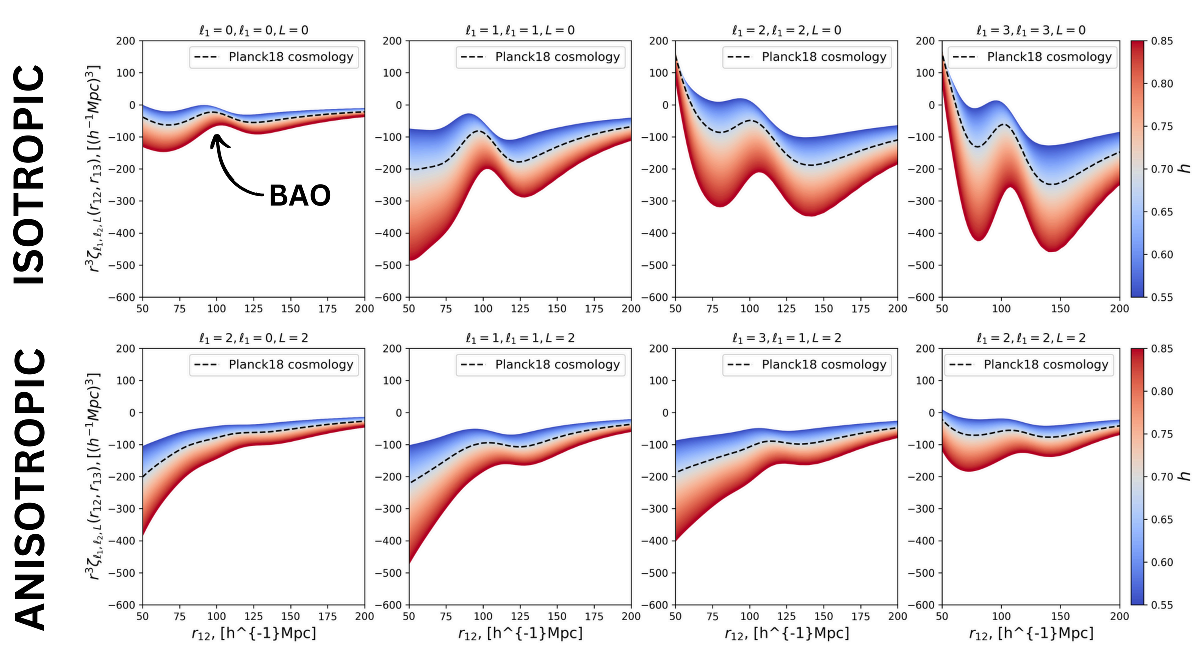
Faster than ever: three-point correlation function produced in less than a second!



4. 3PCF

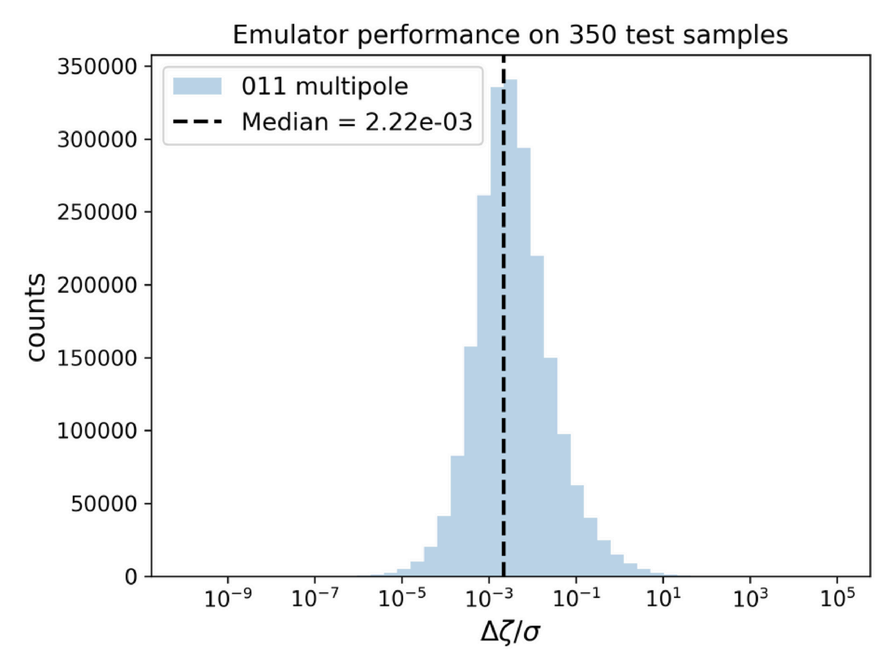
- **Three cosmological parameters** are varied: Ω_m, h, and A_s.
- Isotropic and anisotropic models were calculated with four multipole configurations each.[5]
- Nearly **4000 3PCF** were calculated, using 60000 CPU hours.

EQUILATERAL CONFIGURATION



5. EMULATOR PERFORMANCE

- 8 different emulators were created for each multipole.
- **Normalisation** is **implemented** in the Python package *CosmoPower* for 3PCF emulation.
- Theoretical covariance computed at **V = 8 (Gpc/h)^3**.
- Computational time for each emulator is **less than 0.1 second**.
- The accuracy of the emulator is well below the standard observational noise.

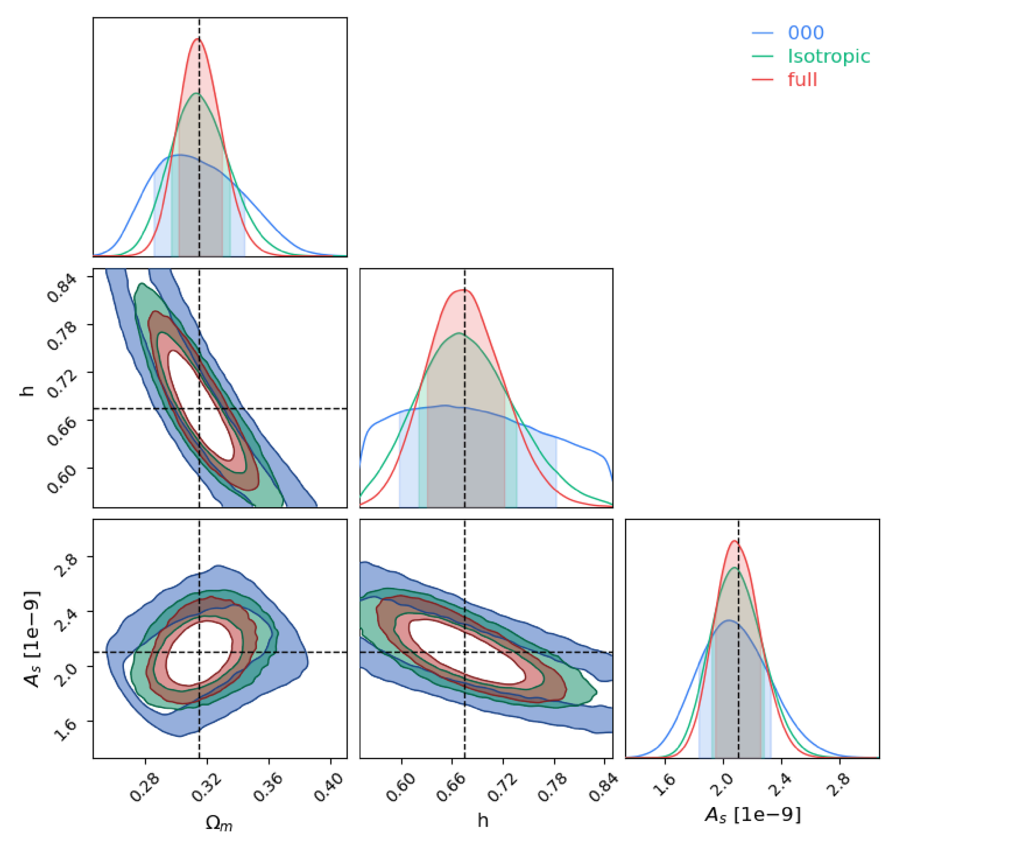


6. MCMC

- Markov chain Monte Carlo (MCMC) is used to **constrain cosmological parameters**.
- 30 walkers with 20'000 iterations each.
- Anisotropic multipole inclusion yields around **20% improvement** for the constraints.
- If **squeezed triangles** ($r_{min} \geq 50\ h^{-1}Mpc$) are **removed** from the analysis, then improvements are negligible.

7. TAKE-AWAY MESSAGE

- **First ever emulation of 3PCF in redshift space**.
- The emulation of 3PCF allows it to be acquired **10 million times faster**.
- MCMC analysis predicts significant **improvement** of parameter constraining by including **anisotropic multipoles**.



8. FUTURE PLANS

- Create a **galaxy** anisotropic 3PCF emulator at redshift one.
- Use **Pinocchio covariance** for analysis.
- Repeat the **MCMC analysis** for galaxy 3PCF to validate results.
- **Apply** the analysis to spectroscopic **survey data**, such as BOSS, DESI, and Euclid.
- Develop **2PCF + 3PCF** analysis.

REFERENCES:

[1] Sefusatti, E. and Scoccimarro, R., 2005. Galaxy bias and halo-occupation numbers from large-scale clustering. Physical Review D, 71(6), p.063001.
[2] Slepian, Z. and Eisenstein, D.J., 2015. Computing the three-point correlation function of galaxies in time. Monthly Notices of the Royal Astronomical Society, 454(4), pp.4142-4158.
[3] Sugiyama, Naonori S., Shun Saito, Florian Beutler, and Hee-Jong Seo. "A complete FFT-based decomposition formalism for the redshift-space bispectrum." Monthly Notices of the Royal Astronomical Society 484, no. 1 (2019): 364-384.
[4] Sugiyama, N.S., Saito, S., Beutler, F. and Seo, H.J., 2021. Towards a self-consistent analysis of the anisotropic galaxy two-and three-point correlation functions on large scales: application to mock galaxy catalogues. Monthly Notices of the Royal Astronomical Society, 501(2), pp.2862-2896.
[5] A. Farina, A. Veropalumbo, E. Branchini, M. Guidi - in preparation
[6] Spurio Mancini, A., Piras, D., Alsing, J., Joachimi, B. and Hobson, M.P., 2022. CosmoPower: emulating cosmological power spectra for accelerated Bayesian inference from next-generation surveys. Monthly Notices of the Royal Astronomical Society, 511(2), pp.1771-1788.

INSTITUTIONS:

1 Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, via Piero Gobetti 93/2, 40129 Bologna, Italy
2 INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy
3 Fakultät für Physik, Ludwig-Maximilians Universität München, Geschwister-Scholl-Platz 1, 80539 München, Germany
4 Dipartimento di Fisica, Università di Genova, Via Dodecaneso, 33, 16146 Genova GE, Italy

CONTACTS:

Kristers Nagainis: *kristers.nagainis@studio.unibo.it*
Massimo Guidi: *massimo.guidi6@unibo.it*
Michele Moresco: *michele.moresco@unibo.it*
Sofia Contarini: *sofia.contarini3@unibo.it*
Antonio Farina: antonio.farina@edu.unige.it
Alfonso Veropalumbo: alfonso.veropalumbo@edu.unige.it

ACKNOWLEDGMENTS:

- We acknowledge the use of computational resources from the parallel computing cluster of the Open Physics Hub (https://site.unibo.it/openphysicshub/en) at the Department of Physics and Astronomy in Bologna.
- We acknowledge ISCRA for awarding this project access to the LEONARDO supercomputer, owned by the EuroHPC Joint Undertaking, hosted by CINECA (Italy) .
- We acknowledge support from MIUR, PRIN 2022 (grant 2022NY2ZRS 001).