SF2A 2018
P. Di Matteo, F. Billebaud, F. Herpin, N. Lagarde, J.-B. Marquette, A. Robin, O. Venot (eds)

BEYOND THE POWER SPECTRUM WITH LARGE DEVIATION THEORY

S. Codis[1]

Abstract. A large-deviation principle is used to model the time-evolution of the large-scale structure of the Universe. This approach allows for analytical predictions in the mildly non-linear regime, beyond what is commonly achievable via other statistics such as correlation functions. The idea is to measure the mean cosmic densities within concentric spheres and study their joint statistics.The spherical symmetry then leads to surprisingly accurate predictions where standard calculations of perturbation theory usually break down. Results for the one-point statistics of the cosmic density field are shown and implications for future large galaxy surveys are discussed.

Keywords: cosmology: theory — large-scale structure of Universe — methods: analytical, numerical

1 Introduction

In the first stages of structure formation, cosmic fields evolve linearly and therefore remain Gaussian. They are thus completely described by their two-point correlation function (or power spectrum in Fourier space). Those primordial Gaussian inhomogeneities then grow under the laws of gravity in our expanding Universe and give rise to the large-scale cosmic web. This subsequent evolution on mildly non-linear scales generates gravitational non-Gaussianities and information then leaks from the power spectrum to higher order statistics inducing non-zero contributions to the whole hierarchy of N-point correlation functions. The above mentioned time evolution of the large-scale structure of the Universe is governed by the so-called Vlasov-Poisson system, which for a self-gravitating collisionless fluid (our hypothesis here) simply expresses that the volume in phase space is conserved (Liouville theorem) and the gravitational potential is sourced by density fluctuations (Poisson equation). This highly non-linear set of equations can be solved using numerical simulations (with N-body methods or resolving the full phase space) or analytically in some specific regimes. Let us emphasize that efforts on the analytical side are not vain as they allow us to understand the details of structure formation and provide important tools to analyse observational datasets, in particular when computing power becomes a limitation. This is notably the case when covariance matrices need to be finely modelled, therefore requiring thousands of numerical realisations of the observable Universe, which is far beyond the reach of our existing facilities. Eventually, hybrid approaches combining simulations and first principles calculations will probably be the optimal way to extract cosmological information from the large-scale structure of the Universe. In the context of covariance matrices, the response function formalism is a nice illustration of how analytical methods can reduce the computing needs.

In this proceeding, we focus on the analytical modelling of the distribution of cosmic fields. Sect. 2 first introduces perturbative approaches while Sect. 3 presents a more successful non-perturbative method based on large-deviation theory. Finally, Sect. 4 wraps up.

2 Cosmological perturbation theories

2.1 N-point correlation functions

On large scales or at early times when the fields are weakly non-Gaussian, perturbation theory (hereafter PT) techniques can be implemented in order to derive predictions for any observable that depends on the density and velocity fields. This is achieved by first taking the first two moments of the Vlasov equation and closing the

[1] CNRS and Sorbonne Université, UMR 7095, Institut d'Astrophysique de Paris, 98 bis Boulevard Arago, F-75014 Paris, France

system assuming there is no shell crossing, therefore no pressure term. Then the fields can be expanded with respect to the initial fields and can eventually be computed hierarchically (order by order) using convolutions of the successive PT kernels (for a review, see Bernardeau et al. 2002). This approach is valid only where the density contrast is small $\delta \ll 1$. Once applied to predict the power spectrum, it allows to reach scales k of a few 0.1 h/Mpc at most. Going deeper into the non-linear regime is extremely difficult due to the slow convergence of the perturbative expansion on the one hand (coupling with small non-linear scales) and to the limitation of the no-shell crossing condition on the other hand. The price to pay is often to introduce a myriad of free parameters that encode our ignorance of the small-scale physics (as is consistently done for instance in the EFTofLSS approach). However, in the context of cosmological inference, it is still unclear whether there is any gain in going to smaller scales (hence having more statistics given the number of modes available in this regime) if many nuisance parameters have to be introduced and eventually marginalised over.

In particular, non-linearities make optimal extraction of information a difficult problem and therefore the question of which observables to use becomes crucial. The current most common approach is to start from the power spectrum which contains all the information in the Gaussian case and go beyond with higher order N-point correlation functions. In practice, the bispectrum is already quite challenging and any analysis is usually restricted to the power spectrum, possibly the bispectrum but no more. An alternative route could be to look for other observables which can be predicted from first principles and can probe the mildly non-linear regime (i.e beyond the weakly non-linear regime but before shell crossing). To achieve this goal, one may think that imposing some level of symmetry could potentially provide us with observables that are less prone to the effect of the unknown small scales. To date, one such configuration has been evidenced: count-in-cells.

2.2 *Count-in-cells : a perturbative approach*

The idea behind count-in-cells is to draw spheres of a given size and measure the mean cosmic matter densities within them. The distribution of these mean densities $\mathcal{P}(\delta)$ (i.e their one point statistics) is initially Gaussian in the first stages of structure formation and subsequently develops some level of non-Gaussianity being in particular increasingly skewed towards small densities as voids are occupying more and more volume and peaks are getting more and more concentrated. Using a perturbative approach, this probability distribution function

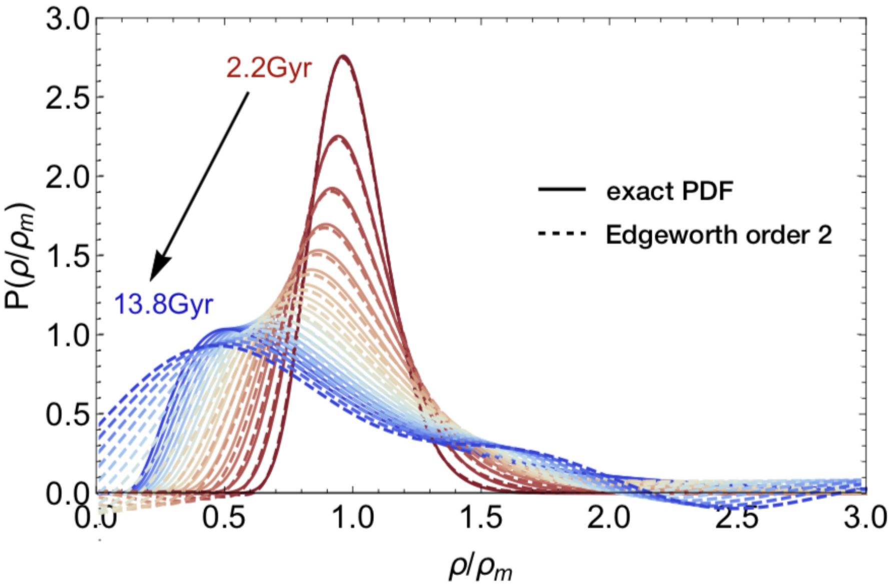

Fig. 1. Distribution of density contrast in spheres of 10Mpc/h at various redshifts from blue to red. The dashed lines are the corresponding second order Edgeworth expansion.

(PDF) of the density field can be written as an Edgeworth expansion (at second order here) around a Gaussian kernel

$$\mathcal{P}(\delta) = \mathcal{G}(\delta)\left[1 + \sigma\frac{S_3}{3!}H_3\left(\frac{\delta}{\sigma}\right) + \sigma^2\left(\frac{S_4}{4!}H_4\left(\frac{\delta}{\sigma}\right) + \frac{1}{2}\left(\frac{S_3}{3!}\right)^2 H_6\left(\frac{\delta}{\sigma}\right)\right) + \cdots\right] \quad (2.1)$$

where H_i are probabilists' Hermite polynomials. This Edgeworth expansion involves the successive field cumulants which in PT are known to scale as increasing power of the amplitude of fluctuation σ, $\langle \delta^i \rangle = S_i \sigma^{2(i-1)}$, where the reduced cumulants S_i are independent of σ at leading order. As an example, Fig. 1 displays the true PDF with solide lines while a second order Edgeworth expansion is superimposed with dashed lines. The various colours show different redshift. As expected the PDF is initially close to Gaussian and becomes more and more skewed with cosmic time. As non-linearities develop, the Edgeworth expansion performs increasingly badly as expected. An important difficulty here is that a truncated Edgeworth expansion never provides us with a well defined non-Gaussian PDF (normalised to one and positive) as soon as the non-Gaussianity is non zero and even if tiny (this is due to the fact that a cumulant generating function is a polynomial only in the case of a Gaussian random variable). However, we will show in the next section that large deviations theory can be used to predict from first principles this PDF without suffering from the issues of negativity and normalisation. In this framework, we will see that all cumulants are accounted for and are exact at tree order.

3 Count-in-cells as modelled by Large deviation theory

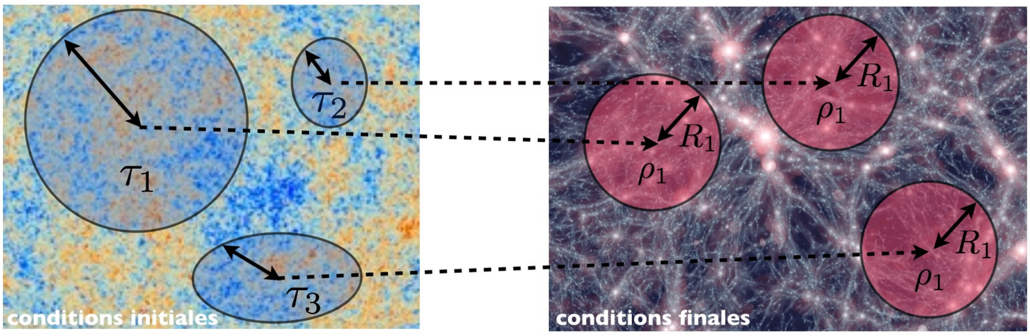

Fig. 2. Spheres of the same radius and density at late times (**right**) can originate from very different patches in the initial conditions (**left**), with various sizes, shapes and densities.

Large deviation theory is at the core of many developments in physics, in particular in statistical physics, in order to describe the tails of exponentially suppressed distributions with the idea that "an unlikely fluctuation is brought about by the least unlikely among all unlikely paths". As such it encompasses and improves upon the central limit theorem which only describes behaviours close to the maximum of the PDF (likely configurations). This theory has recently been applied in cosmology to study the time evolution of the cosmic density field on large scales (Bernardeau & Reimberg 2016). A precise formulation of large deviation theory is beyond the scope of this proceeding and we therefore chose to only describe the qualitative picture here. Further mathematical details can be found in e.g Uhlemann et al. (2016).

The goal here is to predict the statistics of densities in spheres of a given size. In the Gaussian initial conditions, this statistics is Gaussian and well known. However in the late time Universe and on intermediate scales, the field is not Gaussian anymore. As illustrated on Fig. 2, the Lagrangian patches (i.e where the matter is coming from in the initial conditions) corresponding to spheres of same size R and density ρ, show a large variety of shapes, densities and volumes (only the mass is conserved from the initial to the final conditions). This is expected since at a given macroscopic configuration (mean density ρ in the sphere) correspond many different microscopic states that originate from very different initial configurations. In other words, many different paths can connect the initial conditions to the same late time mean density in a sphere. Amongst all this possible paths, there is however one path that dominates the statistics even more so that the event is rare and therefore the variance σ^2 of the field is small. Given that we consider spheres that are maximally symmetric, one can conjecture that the most likely path should also respect this symmetry and may therefore be the so-called spherical collapse model. The spherical collapse mapping is a well known exact non-linear solution of the gravitational evolution when the initial condition is an isolated spherically symmetric fluctuation. Explicit solutions can be found for specific cosmologies (e.g Einstein de Sitter) while in general numerical solutions are necessary. However, it was shown that once expressed in terms of the linear density, the spherical collapse mapping is very weakly dependent on cosmology and is well reproduced by the simple analytic form $\zeta_{\rm SC}(\delta_{\rm l}) = (1 - \delta_{\rm l}/\nu)^{-\nu}$ with $\nu = 21/13$. Once the conjecture on the most likely path is made and in the limit of small variance, large deviation theory can be used to predict the cumulant generating function of the field and eventually get the resulting PDF which is nothing but the inverse Laplace transform of the cumulant generating function. Eventually, the explicit prediction for the density PDF can be obtained (see e.g equation of Codis et al. (2016b) and the corresponding public code `LSSFast` distributed freely at `http://cita.utoronto.ca/~codis/LSSFast.html`).

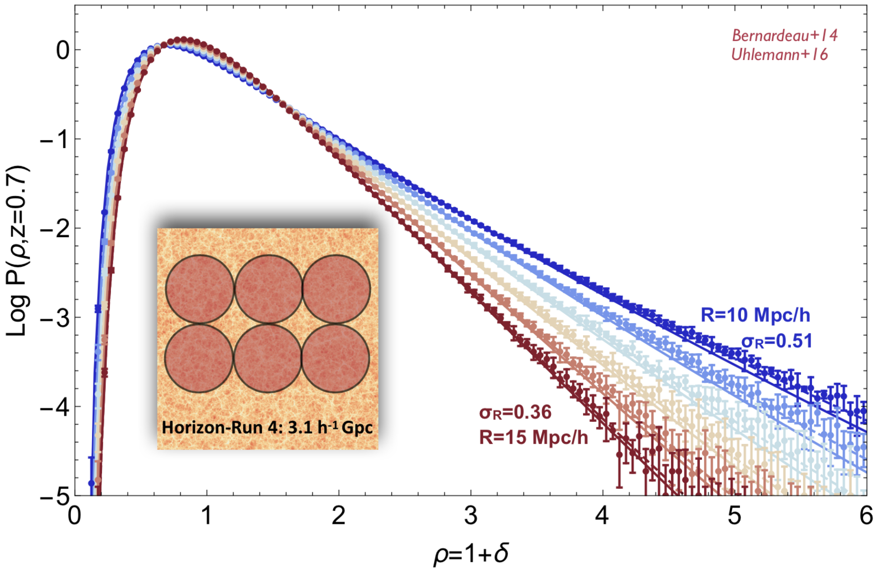

Fig. 3. Distribution of density measured in the Horizon-Run 4 simulation (with error bars) at $z = 0.7$ and for various radii as labelled and compared to the large deviations theory predictions (solid).

All cumulants are predicted at once by this formalism and are exact at tree order in PT. Thanks to this property, the predicted PDF matches very well the measurements in simulation for variance of order unity as can be seen in Fig. 3, even in the tails that are governed by rare (non-linear) events which are beyond the reach of PT techniques. In addition, all predictions are analytical with an explicit dependence on cosmology and no free parameters, a unique situation in this field of research. Note that this formalism can be implemented not only for one sphere but for an arbitrary number of concentric spheres. Large-deviation theory has been successfully applied to the density one-point distribution but also to the density slopes or profiles (Bernardeau et al. 2015), cosmic velocities, to two-point statistics that allowed us to model the error budget expected from finite volumes probed by galaxy surveys (Codis et al. 2016a). Primordial non-Gaussianities were also investigated in this framework. It was shown that using the full knowledge of the PDF, tighter constraints on cosmology (in particular the equation of state of dark energy) could be obtained (Codis et al. 2016b) and once tracer bias is accounted for (we observe galaxies not the underlying total matter density field), density PDFs could help breaking the degeneracies between biasing (nuisance) parameters and cosmological ones (Uhlemann et al. 2018).

4 Conclusions

The distribution of cosmic fields cannot be modelled in the tails using perturbative techniques. A large deviation principle can however be implemented and provides us with a detailed and accurate modelling even in the tails of the distributions and in the mildly non-linear regime beyond what is usually achievable via PT, reaching sub-percent precision for scales above 10 Mpc/h at redshift zero. These ideas are promising to extract cosmological information from the non-Gaussian large-scale structure of the Universe as observed by future galaxy surveys (Euclid among others).

SC thanks Fondation Merac for funding. This work and SC's attendance to the 2018 SF2A meeting were also partially supported by the Programme National Cosmologie et Galaxies (PNCG) of CNRS/INSU with INP and IN2P3, co-funded by CEA and CNES.

References

Bernardeau, F., Codis, S., & Pichon, C. 2015, MNRAS, 449, L105
Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, Phys. Rep., 367, 1
Bernardeau, F. & Reimberg, P. 2016, Phys. Rev. D, 94, 063520
Codis, S., Bernardeau, F., & Pichon, C. 2016a, MNRAS, 460, 1598
Codis, S., Pichon, C., Bernardeau, F., Uhlemann, C., & Prunet, S. 2016b, MNRAS, 460, 1549
Uhlemann, C., Codis, S., Pichon, C., Bernardeau, F., & Reimberg, P. 2016, MNRAS, 460, 1529
Uhlemann, C., Feix, M., Codis, S., et al. 2018, MNRAS, 473, 5098