A Large-Deviation Principle at play in large-scale structure cosmology

Old ideas in a « new pot »

Francis Bernardeau
IAP Paris
and
IPhT Saclay

P. Reimberg, FB, '15, '17
FB, Pichon and Codis '13
FB, Pichon and Codis '15

With the advent of a new generation of wide field cosmological surveys aiming at characterizing the mass and energy content of the universe, it becomes important to develop tools for predicting and understanding non-linear field statistical properties, such as cosmic density spectra or bispectra, beyond the linear regime. To achieve such a goal, flexible theoretical approaches are needed to compute such quantities in a controlled way. Furthermore, these methods should in principle be extended to a variety of cosmological models that include non-standard effects such as massive neutrinos or modified gravity models. The purpose of this workshop is to gather active researchers in the development of efficient analytical methods for the computation of the statistical properties of the large-scale structure of the Universe. It will provide the opportunity for participants to present and discuss the merits and scopes of the different perturbation theory approaches that have been put forward in recent years.

Main topics will include:
- hardcore methods of perturbation theory
- application to redshift-space distortions
- biasing mechanisms and properties of halos
- construction of modified gravity and dark energy models
- impact of massive neutrinos on the development of large-scale structure
- computations of covariances

Eminent scientists in the field will animate the school. These include:

The scientific program will gradually be established, based on the proposals of accepted contributions.

Organization Committee
Francis Bernardeau (IPhT Saclay, FR), Takahiro Nishimichi (IPMU & IAP, Tokyo JP and Paris FR), Patrick Valegeas (IphT Saclay FR)

Application and registration:

No registration fees
Deadline for applications to April 7th, 2013

Francis Bernardeau
Institut de Physique Théorique de Saclay
91191 Gif-sur-Yvette
Francis.bernardeau@cea.fr
Tel 0169088116

Cargèse International School 2013
PT Chat at Cargèse
April 30 – May 3, 2013
Moving from « naked » correlation functions to « dressed » or regularised correlation functions

e.g. clipped density field, or peaks, etc.
Building dressed correlation functions

\[ \langle \rho(x_1) \rho(x_2) \rangle = 1 + \xi(|x_1 - x_2|) \]

\[ \langle \hat{\rho}_\theta(x_1) \hat{\rho}_\theta(x_2) \rangle = \langle \hat{\rho}_\theta(x_1) \rangle \langle \hat{\rho}_\theta(x_2) \rangle \left[ 1 + \hat{\xi}(|x_1 - x_2|) \right] \]

This form assumes that the «dressed» density is defined locally (like the local density in a spherical region, from higher order derivative, etc.) and that the scale at which it is defined is much smaller than the separation.

Perturbatively, it collects contributions coming from higher correlation function in the squeezed limit.
If one knows how to compute the dressed density, then the bias factor is given by the linear response of the dressed density with a large-scale variation.

Are there quantities that are better suited for such calculations, taking into account non-linear evolution?

The new pot = the large deviation principle.
Large-deviation theory, one step beyond the central limit theorem.

It addresses the question: what is the most likely way for an unlikely event to happen?

Can serve as a computational method and/or guideline for quantities of interest.
Basics of theory of large deviation functions

Review paper by Hugo Touchette, ‘09

One example: tossing coins and taking the average number of heads

\[
 x = \frac{1}{n} \sum_{n} t_n 
\]

\[
 P_n(> x) \asymp \exp(-nI(x)) 
\]

Put a threshold at a fixed position

Central limit theorem: \( I(x) = 2(x - 0.5)^2 \)

Exact result: \( I(x) = x \log[x] + (1 - x) \log[1 - x] + \log[2] \)

The cumulant generating function: \( \varphi(\lambda) = \log \left( e^\lambda / 2 + 1/2 \right) \)

Cramér’s Theorem: both are Legendre transform of one-another
Key theorems: from rate function to scaled cumulant generating functions

The **Contraction Principle**: the rate function of an unlikely event is the rate function of the most likely configuration for it to happen.

For a mapping $x \rightarrow y$ we have, $I(y) = \inf_{x, x \rightarrow y} I(x)$

that is the rate function for $y$ is the smallest rate function (the most probable) of the values (configurations) that lead to $y$.

The **Gärtner-Ellis Theorem** (Cramér’s Theorem for IID): the rate function is the Legendre-Fenchel transform of the (scaled) cumulant generating function

$$I(\rho) = \sup_{\lambda} [\lambda \rho - \varphi(\lambda)]$$

Under some regularity conditions, this relation can be inverted in

$$\varphi(\lambda) = \sup_{\rho} [\lambda \rho - I(\rho)]$$

The scaled cumulant generating function:

$$\varphi(\lambda) = \lim_{\langle \rho^2 \rangle_c \rightarrow 0} \langle \rho^2 \rangle_c \sum_{p=1}^{\infty} \frac{\langle \rho^p \rangle_c}{p!} \left( \frac{\lambda}{\langle \rho^2 \rangle_c} \right)^p = \lambda + \frac{\lambda^2}{2} + S_3 \frac{\lambda^3}{3!} + \ldots$$
Large-Deviation Principle in the context of large-scale structure cosmology

Discrete or continuous sets of Gaussian variables obey the Large Deviation Principle: their rate function is a simple quadratic form.

One needs a mapping... (a priori non-linear and non-local)
An explicit large-deviation regime

If one restricts the ensemble of realisations to spherically symmetric configurations, one can define a set of random variables - the densities in concentric shells - for which we know the rate function and the mapping into their nonlinear values.

The collection \( \{\delta_{\mathrm{lin}}(\theta_i)\}_{1 \leq i \leq N} \) of correlated gaussian random variables obeys the LDP with rate function:

\[
I(\delta_{<}^{\mathrm{lin}}(\theta_1), \ldots, \delta_{<}^{\mathrm{lin}}(\theta_N)) = \frac{\sigma^2(\theta_N)}{2} \sum_{i,j} \Xi_{ij} \delta_{<}^{\mathrm{lin}}(\theta_i) \delta_{<}^{\mathrm{lin}}(\theta_j)
\]

where \( \Xi = \Sigma^{-1} \), and \( \sigma^2(\theta_N) = \Sigma_{NN} \).
The spherical collapse: the solution for specific initial conditions (with adiab. modes)

\[
\frac{d^2 R}{dt^2} = - \frac{GM(< R)}{R^2}
\]

The radius evolution

The exact non-linear mapping for spherically symmetric initial profile (for growing mode setting)

Note that this mapping is independent on the small scale physics (with baryons, shell crossings, etc.)

Is it good enough for spherically symmetric observables? Not necessarily (e.g. Zel’dovich approximation, FB, Reimberg, in prep.)
There exists a mapping which maps the initial radii into the nonlinear ones

\[ \delta_<(\vartheta) = \zeta(\delta_\text{lin}^<) \]

\[ \vartheta = \theta \zeta^{-1/D}(\delta_\text{lin}^<) \]

The scaled cumulant generating function of any functional of the nonlinear density profile is then given by,

\[ \varphi(\lambda) = \sup_{\delta_\text{lin}^<} \left[ \lambda \hat{\rho}\{\delta_<(\vartheta)\} - I(\delta_\text{lin}(\theta_1), \ldots, \delta_\text{lin}(\theta_N)) \right] \]

\[ \hat{\rho}\{\delta_<(\vartheta)\} \] does not have to be local, linear or defined from a discrete number of shells.
Consequences in the context of LSS cosmology are at least 2 folds:

- **you do not need to impose** \( \delta(x) \) **to be small everywhere, only the variance has to be small;**
- **you have a possible working procedure provided you can identify the most likely initial configuration and its probability (rate function).**

Such an identification can be done for configurations with enough symmetries: in practice with spherical (or cylindrical) symmetry.
Standard result: the cumulants of the top-hat smoothed density

scaled cumulant GF is Legendre T. of rate function:

\[
\varphi(\lambda) = \lim_{\langle \rho^2 \rangle_c \to 0} \langle \rho^p \rangle_c \sum_{p=1}^{\infty} \frac{\langle \rho^p \rangle_c}{p!} \left( \frac{\lambda}{\langle \rho^2 \rangle_c} \right)^p = \lambda + \frac{\lambda^2}{2} + S_3 \frac{\lambda^3}{3!} + \ldots
\]

Average of (combination of) tree order expression of the p-point correlation functions in spherical cells.

**Expression of** \( S_p = \lim_{\langle \delta^2 \rangle_c \to 0} \frac{\langle \delta^p \rangle_c}{\langle \delta^2 \rangle_c^{p-1}} = \text{tree order expr.} \)

\[
\langle \delta^3 \rangle = 6 \int \frac{dk_1}{(2\pi)^3} P(k_1) P(k_2) \times F_2(k_1, k_2) W(k_1 R) W(k_2 R) W(|k_1 + k_2| R) \propto \langle \delta^2 \rangle^2
\]

\[
S_3 = \frac{34}{7} + \gamma_1,
S_4 = \frac{60712}{1323} + \frac{62 \gamma_1}{3} + \frac{7 \gamma_1^2}{3} + \frac{2 \gamma_2}{3},
S_5 = \frac{200575880}{305613} + \frac{1847200 \gamma_1}{3969} + \frac{6940 \gamma_1^2}{63} + \frac{235 \gamma_1^3}{27} + \frac{1490 \gamma_2}{63} + \frac{50 \gamma_1 \gamma_2}{9} + \frac{10 \gamma_3}{27},
\]

\[
\gamma_p = \frac{d^p \log \sigma^2(R_0)}{d \log^p R_0}.
\]

1-cell density cumulants (FB '94)

it has a non trivial dependence on the wave vectors through the functions \( F_3 \) and \( F_2 \)
Application 1: 1-cell PDF and stats

FB Pichon, Codis ’13

The inverse Laplace transform,

\[ P(\hat{\rho}_1) = \int_{-i\infty}^{+i\infty} \frac{d\lambda_1}{2\pi i} \exp(-\lambda_1\hat{\rho}_1 + \varphi(\lambda_1)) \]

\[ R = 10 \: h^{-1} \: \text{Mpc} \]
Computation of the 1-cell density PDF, LDP applied to $\mu = \log \hat{\rho}$

The general expression of the PDF

$$\mathcal{P}_{R,\mu}(\mu) d\mu = \int_{-i\infty}^{+i\infty} \frac{d\lambda}{2\pi i} \exp[-\lambda \mu + \phi_{R,\mu}(\lambda)],$$

$$\mathcal{P}_R(\hat{\rho}) d\hat{\rho} = \mathcal{P}_{R,\mu}(\log(\hat{\rho})) \frac{d\hat{\rho}}{\hat{\rho}},$$

The saddle point expression

$$\mathcal{P}_R(\hat{\rho}) = \sqrt{\frac{\Psi''[\hat{\rho}] + \Psi'[\hat{\rho}] / \hat{\rho}}{2\pi}} \exp(-\Psi_R[\hat{\rho}])$$

\[\text{Uhlemann, Codis, FB, Pichon, Reimberg '15}\]
The 2-cell probability distribution function

FB, Pichon, Codis '13
Uhlemann, Codis, FB, Pichon, Reimberg '15

Choice of variables

\[ \mu_1 = \log \left( r^3 \hat{\rho}_2 + \hat{\rho}_1 \right), \]
\[ \mu_2 = \log \left( r^3 \hat{\rho}_2 - \hat{\rho}_1 \right), \]

Saddle point expression

\[ \mathcal{P}_{R_1,R_2}(\hat{\rho}_1,\hat{\rho}_2) = \frac{\exp[-\Psi_{R_1,R_2}(\hat{\rho}_1,\hat{\rho}_2)]}{2\pi} \]
\[ \times \sqrt{\det \left[ \frac{\partial^2 \Psi_{R_1,R_2}}{\partial \mu_i \partial \mu_j} \right]} \det \left[ \frac{\partial \mu_i}{\partial \hat{\rho}_j} \right] \]
Figure 3. Density profiles in underdense (solid light blue), overdense (dashed purple) and all regions (dashed blue) for cells of radii $R_1 = 10 \, \text{Mpc}/h$ and $R_2 = 11 \, \text{Mpc}/h$ at redshift $z = 0.97$. Predictions are successfully compared to measurements in simulations (points with error bars).
It opens the way to build constrained PDFs.

**Figure 11.** Top panel: PDF of the slope, of the slope when the inner density is below one and of the slope when the inner density is above one. Error bars represent the error on the mean as measured in our simulation, red lines represent the numerical integration while blue lines are the log-mass saddle approximation given by equation (30). The agreement is very good for the whole range of density and slope probed by the simulation. Bottom panel: residuals of measured slope PDFs compared to the log-mass saddle approximation corresponding to the blue lines in the top panel.
A realistic Mass-aperture statistics

\[ \varphi(\lambda) = -\inf_{\delta_{\text{lin}}} \left[ \lambda M_{\text{ap}}(\delta_{\text{lin}}) + \frac{\sigma_F^2}{2} \int d\theta \int d\theta' \delta_{\text{lin}}(\theta) \delta_{\text{lin}}(\theta') \xi(\theta, \theta') \right] \]

\[ M_{\text{ap}} = \int d^2\vartheta \, W(\vartheta) \frac{\gamma_t}{1 - \kappa} \]

- Gaussian profile
- Taking into account the fact that what we measure is the reduced shear (i.e. a non-linear functional of the profile)

FIG. 5: The effective cumulant generating functions \( \varphi_{\text{eff}}^\kappa \) and \( \varphi_{\text{eff}}^g \) satisfying Eq. (51). The projection factor \( w_{\text{eff}} = 0.1 \) is used on the \( \varphi_{\text{eff}}^g \) data.
Towards a complete theory of cell density statistics...

**Joint PDFs read in the large-separation limit (no finite separation effects)**

\[
P(\{\hat{\rho}_k\}, \{\hat{\rho}'_k\}; r_e) = P(\{\hat{\rho}_k\})P(\{\hat{\rho}'_k\}) \left[1 + \xi(r_e)b(\{\hat{\rho}_k\})b(\{\hat{\rho}'_k\})\right]
\]

**Correlation of measured density probabilities in different locations**

\[
\left\langle \hat{P}(\hat{\rho}_i)\hat{P}(\hat{\rho}_j) \right\rangle = \bar{P}(\hat{\rho}_i)\bar{P}(\hat{\rho}_j)(1 + \xi b_i b_j)
\]
Results for the bias for the density and for the slope.

Consistency relations

\[
\int_{0}^{\infty} \, d\rho \, b(\rho) \, P(\rho) = 0
\]

\[
\int_{0}^{\infty} \, d\rho \, \rho \, b(\rho) \, P(\rho) = 1
\]

For the slope
A regime of large-deviation functions can be identified in LSS cosmology.

- Observables can be related to joint PDFs of the density in concentric cells but also to the cumulant generating function.

Perspectives - what are the domains of application ?:

- These calculations can be applied to 3D and projected mass maps, and to joint density of multiple tracers;
- biasing of over-dense/under-dense regions can also be computed = statistical properties of clipped regions;
- it can be applied to some non-linear transforms of the density field;
- other configuration/geometries ?