Large-scale galaxy bias

A new review to appear in Physics Reports

Donghui Jeong (Penn State)

Collaborators:
Vincent Desjacques (Geneva, Technion)
Fabian Schmidt (MPA)

Large-Scale Galaxy Bias

Vincent Desjacques[a], Donghui Jeong[b], Fabian Schmidt[c]

[a]*Département de Physique Théorique and Center for Astroparticle Physics, Université de Genève, 24 quai Ernest Ansermet, CH-1221 Genève 4, Switzerland*
[b]*Department of Astronomy and Astrophysics, and Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA*
[c] *Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, 85748 Garching, Germany*

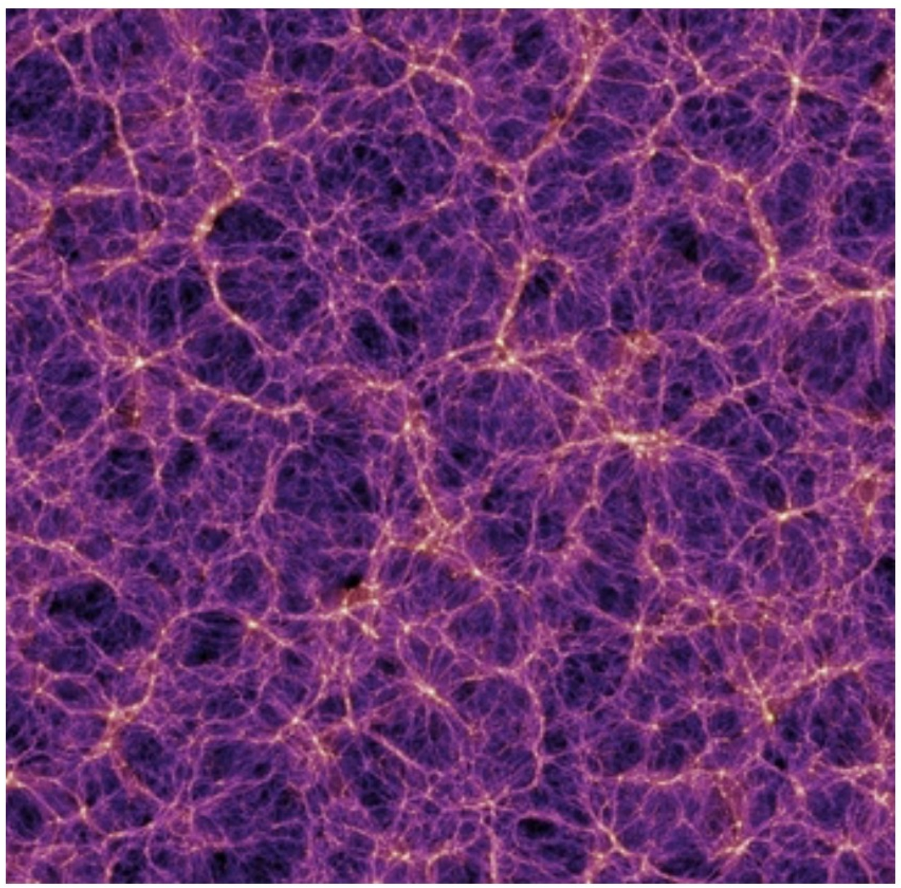
Abstract

We present a comprehensive review of galaxy bias, that is, the statistical relation between the distribution of galaxies and matter. Moreover, we focus on large scales where cosmic density fields are quasi-linear. On these scales, the clustering of galaxies can be described by a perturbative bias expansion, and the complicated physics of galaxy formation can be absorbed into a finite set of coefficients of the expansion, called *bias parameters*. The review begins with a pedagogical proof of this very important result, which forms the basis of the rigorous perturbative description of galaxy clustering, under the assumptions of General Relativity and Gaussian, adiabatic initial conditions. Key components of the bias expansion are quantities that are nonlocal in the matter density, in particular tidal fields and their time derivatives. This derivation is followed by a presentation of the peak-background split in its general form, which elucidates the physical meaning of the bias parameters, and a detailed description of the connection between bias parameters and galaxy (or halo) statistics. We then review the excursion set formalism and peak theory which provide predictions for the values of the bias parameters. In the remainder of the review, we consider the generalizations of galaxy bias required in the presence of various types of cosmological physics that go beyond the standard ΛCDM paradigm with adiabatic, Gaussian initial conditions: primordial non-Gaussianity, massive neutrinos, baryon-CDM isocurvature perturbations, dark energy, and modified gravity. Finally, we discuss how the description of galaxy bias in the galaxies' rest frame is related to observed clustering statistics measured from the observed angular positions and redshifts in actual galaxy catalogs.

Keywords:

What is galaxy bias?

What we can predict



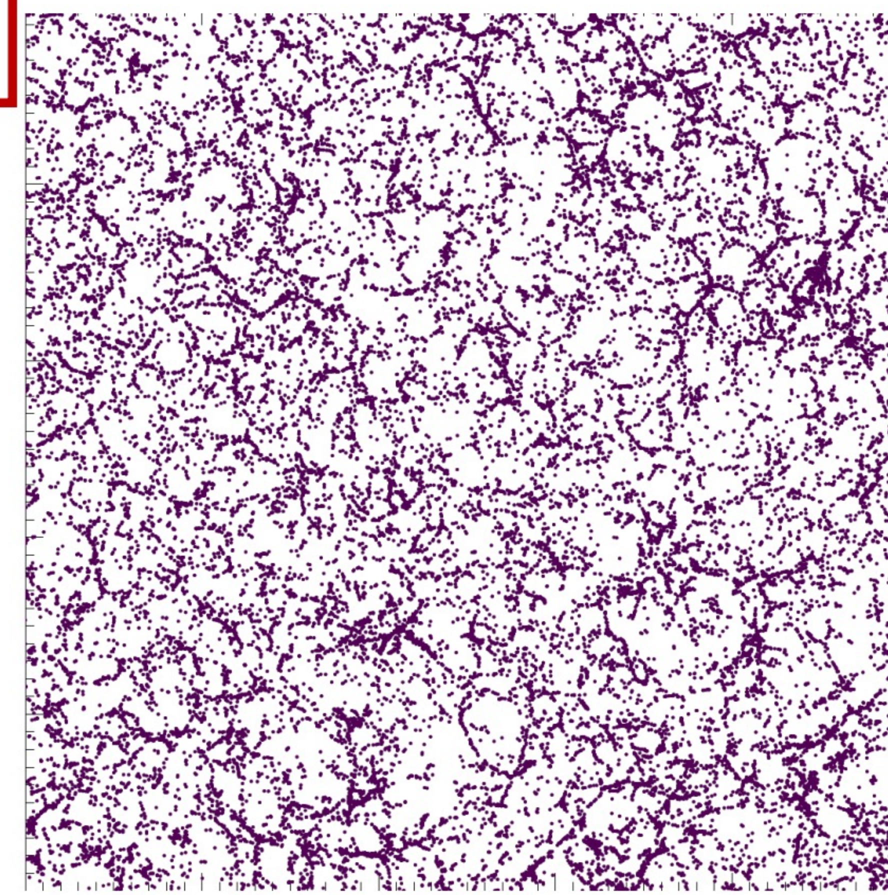
matter density: δ_m

Springel+ 2005

$$\delta_g(\boldsymbol{x}) = \mathcal{F}\left[\delta_m(\boldsymbol{x})\right]$$

- halo formation
- halo merger
- gas cooling
- star formation
- Feedback
- ...

What we observe



galaxy density: δ_g

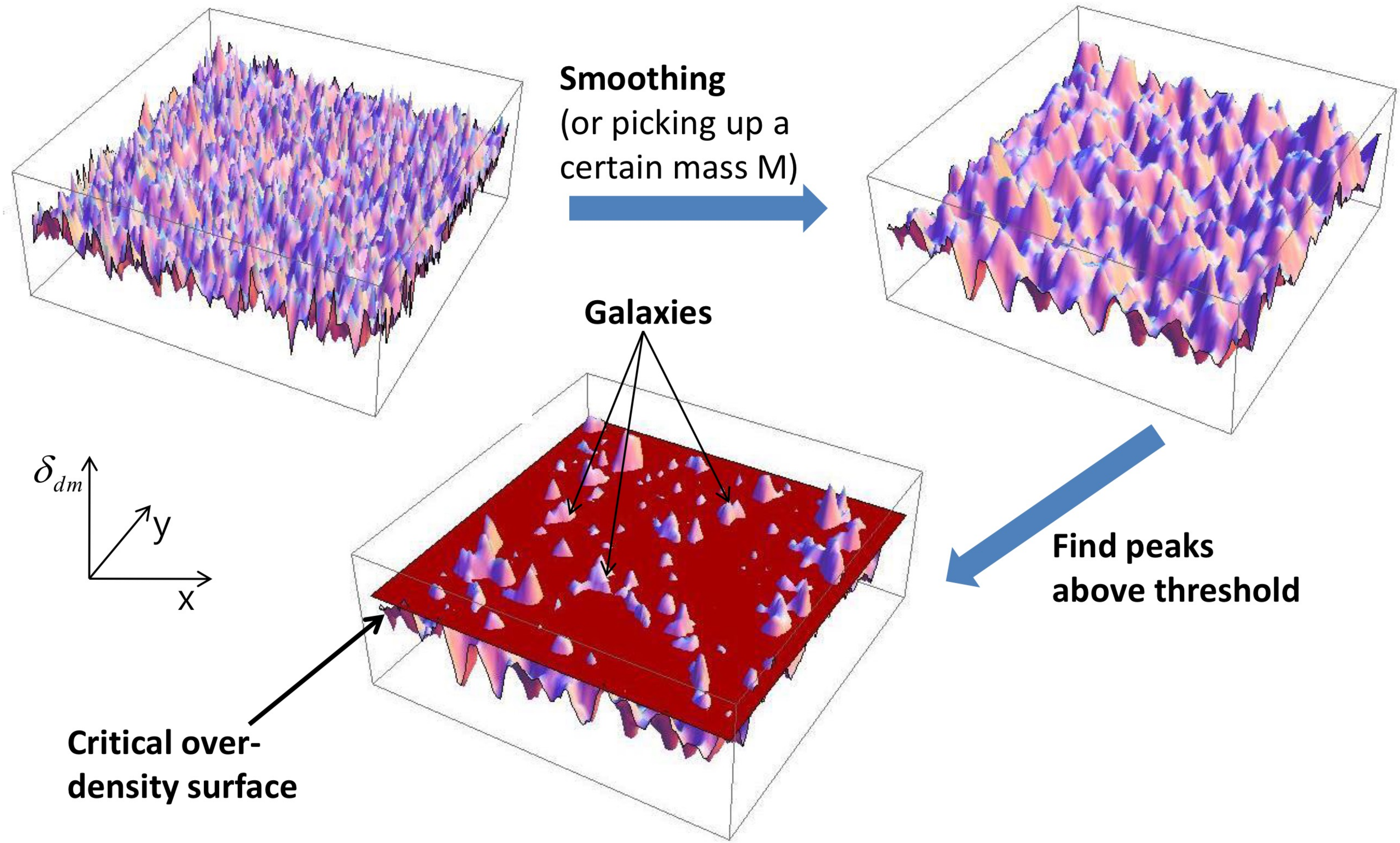
De Lucia+ 2007

Motivation and goal

- Therefore, exploiting galaxy survey data beyond BAO requires robust modeling of galaxy bias.
- Goal: to write the **observables** (galaxy two- and three-point correlation function) in terms of **matter correlation functions** and **a few bias parameters**; then, cosmological parameters can be measured after marginalizing over them.
- Possible, at least, in quasi-linear scales where cosmic density fields are well modeled by perturbation theory.

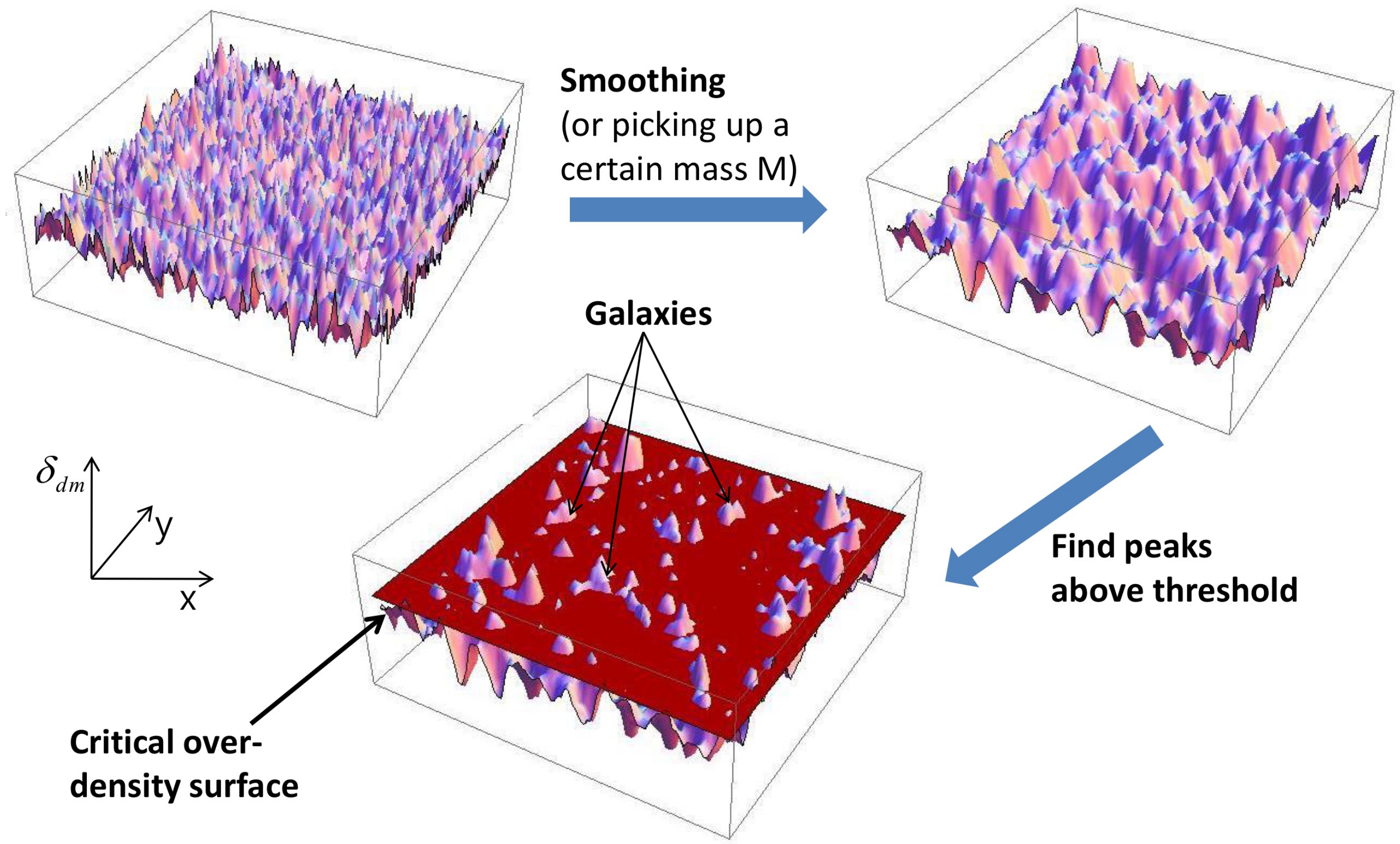
A simple model: thresholding

Kaiser (1984)



A simple model: thresholding

Essentially, excursion set, peak theory are extension of this.



Mathematically,

$$n^L_{\rm thr}(\boldsymbol{q}) = \Theta_H\left(\delta^{(1)}_R(\boldsymbol{q}) - \delta_c\right)$$

$\boldsymbol{q}$=Lagrangian coordinate

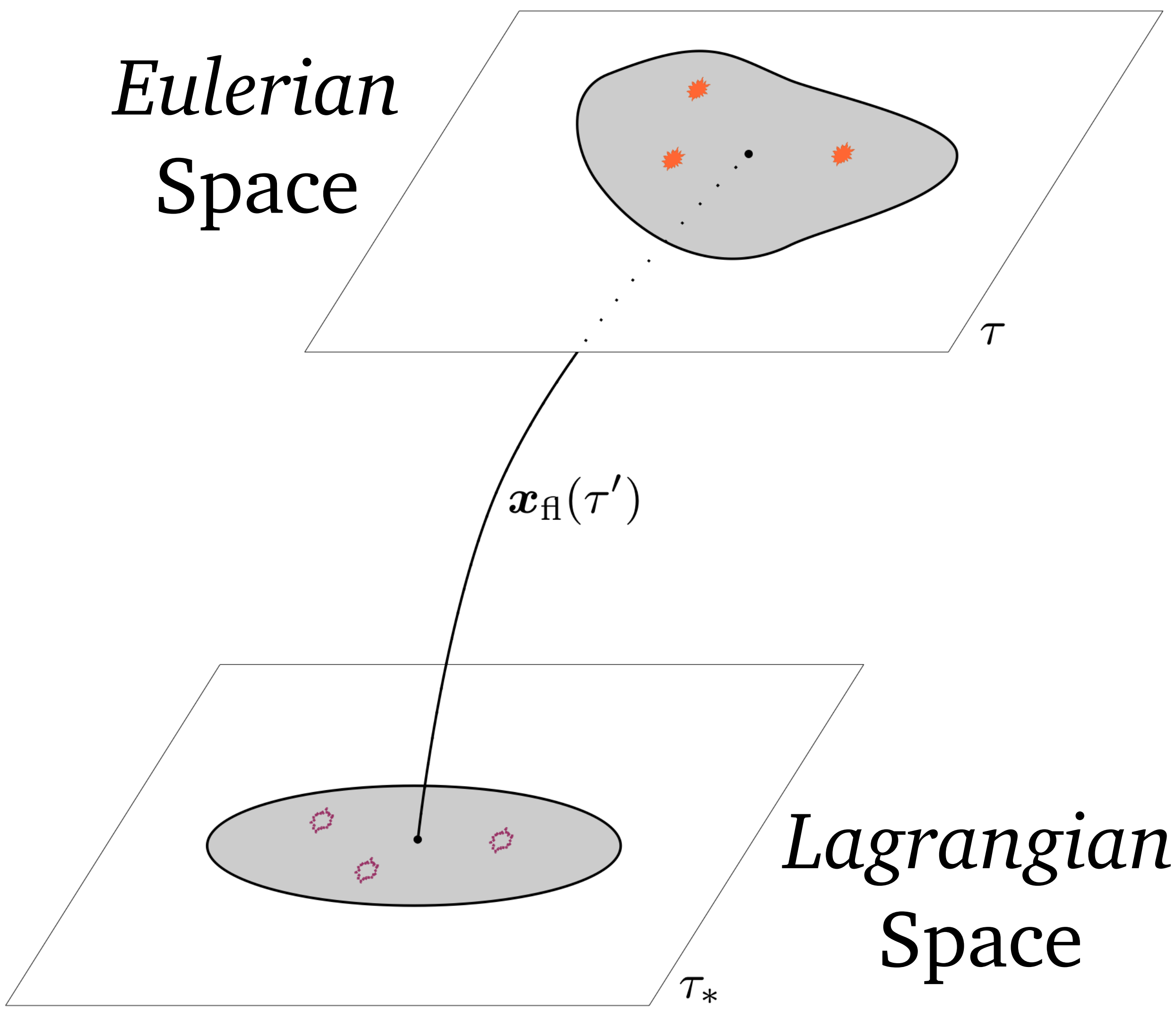
Thresholding leads to LIMD (Local-In-Matter-Density) bias:

$$\delta^L_g(\boldsymbol{q}) = \varepsilon(\boldsymbol{q}) + b^L_1\delta^{(1)}_R(\boldsymbol{q}) + \frac{b^L_2}{2}\left(\left[\delta^{(1)}_R(\boldsymbol{q})\right]^2 - \sigma^2(R)\right) + \cdots$$

ε = stochasticity; b_1^L, b_2^L = bias parameters

Lagrangian to Eulerian bias

How to connect the Lagrangian density to what we *observe*?



- Assumptions :
  - 1. Galaxies co-move with large-scale structure (no non-gravitational momentum transfer; good on large-scales)
  - 2. Galaxies form instantaneously at time $\tau = \tau_*$ (this can be think of as a 'Green's function' for galaxy bias)
- On large-scale where perturbation theory is valid (quasi-linear scales)

Lagrangian to Eulerian bias

Set of equations to solve

$$\frac{\partial \delta(\boldsymbol{x},\tau)}{\partial \tau} + \nabla \cdot [(1+\delta(\boldsymbol{x},\tau))\boldsymbol{v}(\boldsymbol{x},\tau)] = 0$$ Matter continuity eq.

$$\frac{\partial \delta_g(\boldsymbol{x},\tau)}{\partial \tau} + \nabla \cdot \left[\left(1+\delta_g(\boldsymbol{x},\tau)\right)\boldsymbol{v}(\boldsymbol{x},\tau)\right] = 0$$ Galaxy continuity eq.

$$\frac{\partial \boldsymbol{v}(\boldsymbol{x},\tau)}{\partial \tau} + [\boldsymbol{v}(\boldsymbol{x},\tau)\cdot\nabla]\,\boldsymbol{v}(\boldsymbol{x},\tau) + \mathcal{H}(\tau)\boldsymbol{x}(\boldsymbol{x},\tau) = -\nabla\Phi(\boldsymbol{x},\tau)$$ Euler eq.

$$\nabla^2\Phi(\boldsymbol{x},\tau) = \frac{3}{2}\mathcal{H}^2(\tau)\Omega_m(\tau)\delta(\boldsymbol{x},\tau)$$ Poisson eq.

$$\delta_g(\boldsymbol{x}_*,\tau_*) = b_1^L\delta_* + \varepsilon_* + \frac{1}{2}b_2^L\delta_*^2 + \frac{1}{6}b_3^L\delta_*^3 + \mathcal{O}(\delta^4)$$ Initial condition

Solution: local bias

$$\delta_g = b_1\delta + \varepsilon$$
$$+ \frac{1}{2}b_2\delta^2 + b_{K^2}(K_{ij})^2 + \varepsilon_\delta\delta$$
$$+ \frac{1}{6}b_3\delta^3 + b_{\delta K^2}\delta(K_{ij})^2 + b_{K^3}(K_{ij})^3 + b_{\mathrm{td}}O^{(3)}_{\mathrm{td}} + \varepsilon_{\delta^2}\delta^2 + \varepsilon_{K^2}(K_{ij})^2$$
$$+ \mathcal{O}(\delta^4)$$

McDonald & Roy (2009)

Local bias: all terms here are evaluated at the same location!
That is, galaxy density is determined by local quantities.

Solution: local bias

$$\begin{aligned}\delta_g = {} & b_1\delta + \varepsilon \\ & + \frac{1}{2}b_2\delta^2 + b_{K^2}(K_{ij})^2 + \varepsilon_\delta\delta \\ & + \frac{1}{6}b_3\delta^3 + b_{\delta K^2}\delta(K_{ij})^2 + b_{K^3}(K_{ij})^3 + b_{\rm td}O^{(3)}_{\rm td} + \varepsilon_{\delta^2}\delta^2 + \varepsilon_{K^2}(K_{ij})^2 \\ & + \mathcal{O}(\delta^4)\end{aligned}$$

linear order

quadratic order

cubic order

Solution: local bias

$$
\begin{aligned}
\delta_g = &\, b_1\delta + \varepsilon \\
&+ \frac{1}{2}b_2\delta^2 + b_{K^2}(K_{ij})^2 + \varepsilon_\delta\delta \\
&+ \frac{1}{6}b_3\delta^3 + b_{\delta K^2}\delta(K_{ij})^2 + b_{K^3}(K_{ij})^3 + b_{\rm td}O^{(3)}_{\rm td} + \varepsilon_{\delta^2}\delta^2 + \varepsilon_{K^2}(K_{ij})^2 \\
&+ \mathcal{O}(\delta^4)
\end{aligned}
$$

Terms previously known as ***local bias*** (Fry & Gaztanaga 1993)
Local in a sense that δ_g only depends on δ at the same position.
This, however, only captures a part of the full terms, and we call this
LIMD bias (Local-In-Matter-Density bias).

Solution: local bias

$$
\begin{aligned}
\delta_g = & \; b_1 \delta + \varepsilon \\
& + \frac{1}{2} b_2 \delta^2 + b_{K^2} (K_{ij})^2 + \varepsilon_\delta \delta \\
& + \frac{1}{6} b_3 \delta^3 + b_{\delta K^2} \delta (K_{ij})^2 + b_{K^3} (K_{ij})^3 + b_{\mathrm{td}} O^{(3)}_{\mathrm{td}} + \varepsilon_{\delta^2} \delta^2 + \varepsilon_{K^2} (K_{ij})^2 \\
& + \mathcal{O}(\delta^4)
\end{aligned}
$$

Terms proportional to the **local tidal field K_{ij}**, and **local density δ**.

Solution: local bias

$$\delta_g = b_1\delta + \varepsilon + \frac{1}{2}b_2\delta^2 + b_{K^2}(K_{ij})^2 + \varepsilon_\delta \delta + \frac{1}{6}b_3\delta^3 + b_{\delta K^2}\delta(K_{ij})^2 + b_{K^3}(K_{ij})^3 + b_{\rm td}O^{(3)}_{\rm td} + \varepsilon_{\delta^2}\delta^2 + \varepsilon_{K^2}(K_{ij})^2 + \mathcal{O}(\delta^4)$$

Terms proportional to the **local tidal field K_{ij}**, and **local density δ**.

The new term that is not proportional to the local density field nor local tidal field. ***Therefore, the general bias expansion should NOT be the Taylor-type expansion of local density field and local tidal field!***

Solution: local bias

$$\delta_g = b_1\delta + \varepsilon + \frac{1}{2}b_2\delta^2 + b_{K^2}(K_{ij})^2 + \varepsilon_\delta \delta + \frac{1}{6}b_3\delta^3 + b_{\delta K^2}\delta(K_{ij})^2 + b_{K^3}(K_{ij})^3 + b_{\rm td}O^{(3)}_{\rm td} + \varepsilon_{\delta^2}\delta^2 + \varepsilon_{K^2}(K_{ij})^2 + \mathcal{O}(\delta^4)$$

Terms proportional to the **local tidal field K_{ij}**, and **local density δ**.

This term is, in fact, **the convective derivative of the local tidal field**, the rate of change of the local tidal field measured by a local observer moving with the galaxies.

Solution: local bias

$$\delta_g = b_1\delta + \varepsilon + \frac{1}{2}b_2\delta^2 + b_{K^2}(K_{ij})^2 + \varepsilon_\delta\delta + \frac{1}{6}b_3\delta^3 + b_{\delta K^2}\delta(K_{ij})^2 + b_{K^3}(K_{ij})^3 + b_{\rm td}O^{(3)}_{\rm td} + \varepsilon_{\delta^2}\delta^2 + \varepsilon_{K^2}(K_{ij})^2 + \mathcal{O}(\delta^4)$$

Terms proportional to the **local tidal field K_{ij}**, and **local density δ**.

This term is, in fact, **the convective derivative of the local tidal field**.

We then have **higher-order stochastic bias terms**, which are proportional to ε for galaxies form at a single time, but must treat as independent parameters for general galaxy samples.

General local bias expansion

$$\delta_g(\boldsymbol{x}, \tau) = \sum_O b_O O(\boldsymbol{x}, \tau) + \varepsilon(\boldsymbol{x}, \tau) + \sum_O \varepsilon_O(\boldsymbol{x}, \tau) O(\boldsymbol{x}, \tau)$$

- Here, O's are operators that the galaxy density contrast depends upon.
  ε, ε_O are stochastic variables.

- Which operators should be included?

  - Key principle: **All observables of the local observers can affect the evolution of galaxies**: hence, ***all local-observable operators must be included!***

  - In plane vanilla cosmology, these are ***density(δ), tidal field (K_{ij}) and their convective derivatives*** (along the trajectory of local comoving observers).

  - In perturbation theory, each order contains unique time dependence; including higher-order terms effectively includes time evolution as well.

General local bias expansion

$$\delta_g(\boldsymbol{x},\tau) = \sum_O b_O O(\boldsymbol{x},\tau) + \varepsilon(\boldsymbol{x},\tau) + \sum_O \varepsilon_O(\boldsymbol{x},\tau) O(\boldsymbol{x},\tau)$$

- Here, O's are operators that the galaxy density contrast depends upon. ε, ε_O are stochastic variables.

- What other operators do we discuss in the review?

  - Primordial non-Gaussianities (7) : φ_L at initial time

  - Massive neutrino fluctuations (8.1) : no new term, but scale-dependent b_1

  - Baryon-CDM relative velocity (8.2) : δ_{bc}, θ_{bc} at recombination

  - Clustering dark energy (8.3) : $\delta_{DE} \sim (1+w)$, scale-dependent b_1 near k_J.

Galaxy power spectrum

$$P_g(k,\tau) = (b_1)^2 P_L(k,\tau) + P_\varepsilon$$

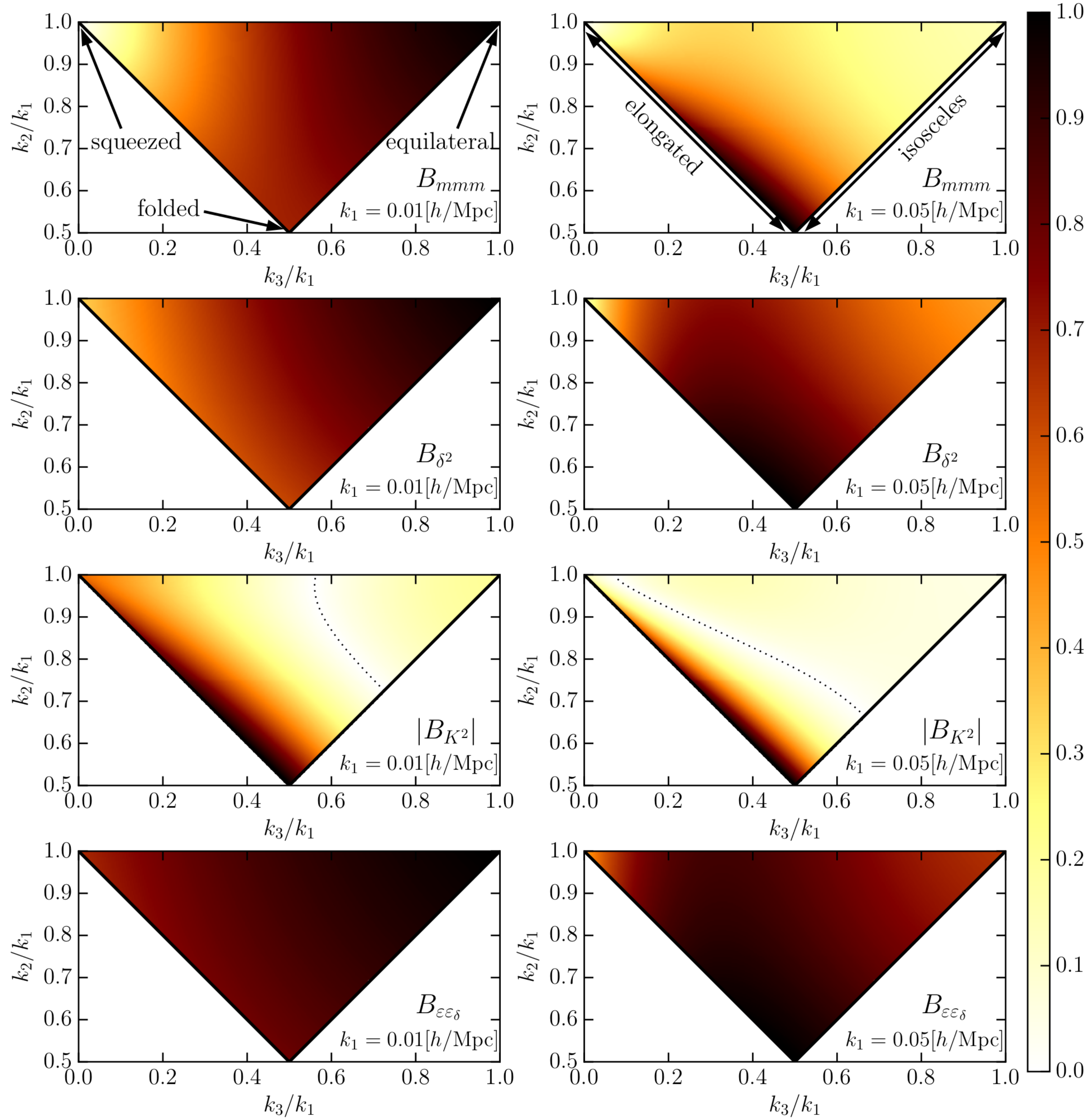
- Power spectrum = Two-point function in Fourier space
- In the leading order expression, linear bias b_1 is strongly degenerated with the amplitude of power spectrum.
- Amplitude of power spectrum can be an important dynamical probe of dark energy! That is, it measures how much dark energy retards the linear growth of structure.

Galaxy bispectrum

$$B_g(k_1,k_2,k_3) = 2b_1^3\left[\frac{5}{7}+\frac{1}{2}(\hat{k}_1\cdot\hat{k}_2)\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)+\frac{2}{7}(\hat{k}_1\cdot\hat{k}_2)^2\right]P_L(k_1)P_L(k_2)+(2\text{ cyclic})$$
$$+\,b_1^2\left[b_2+2b_{K^2}\left((\hat{k}_1\cdot\hat{k}_2)^2-\frac{1}{3}\right)\right]P_L(k_1)P_L(k_2)+(2\text{ cyclic})$$
$$+\,2b_1P_{\varepsilon\varepsilon_\delta}\left[P_L(k_1)+P_L(k_2)+P_L(k_3)\right]+B_\varepsilon$$

- Bispectrum = three-point function in Fourier space
- Statistical homogeneity dictates the triangular condition: $\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3=0$
- Five free parameters (b_1, b_2, b_{K2}, B_ε, $P_{\varepsilon\varepsilon\delta}$), but many many triangles with unique configuration dependence!

Shape of galaxy bispectrum



Projection: measuring bias

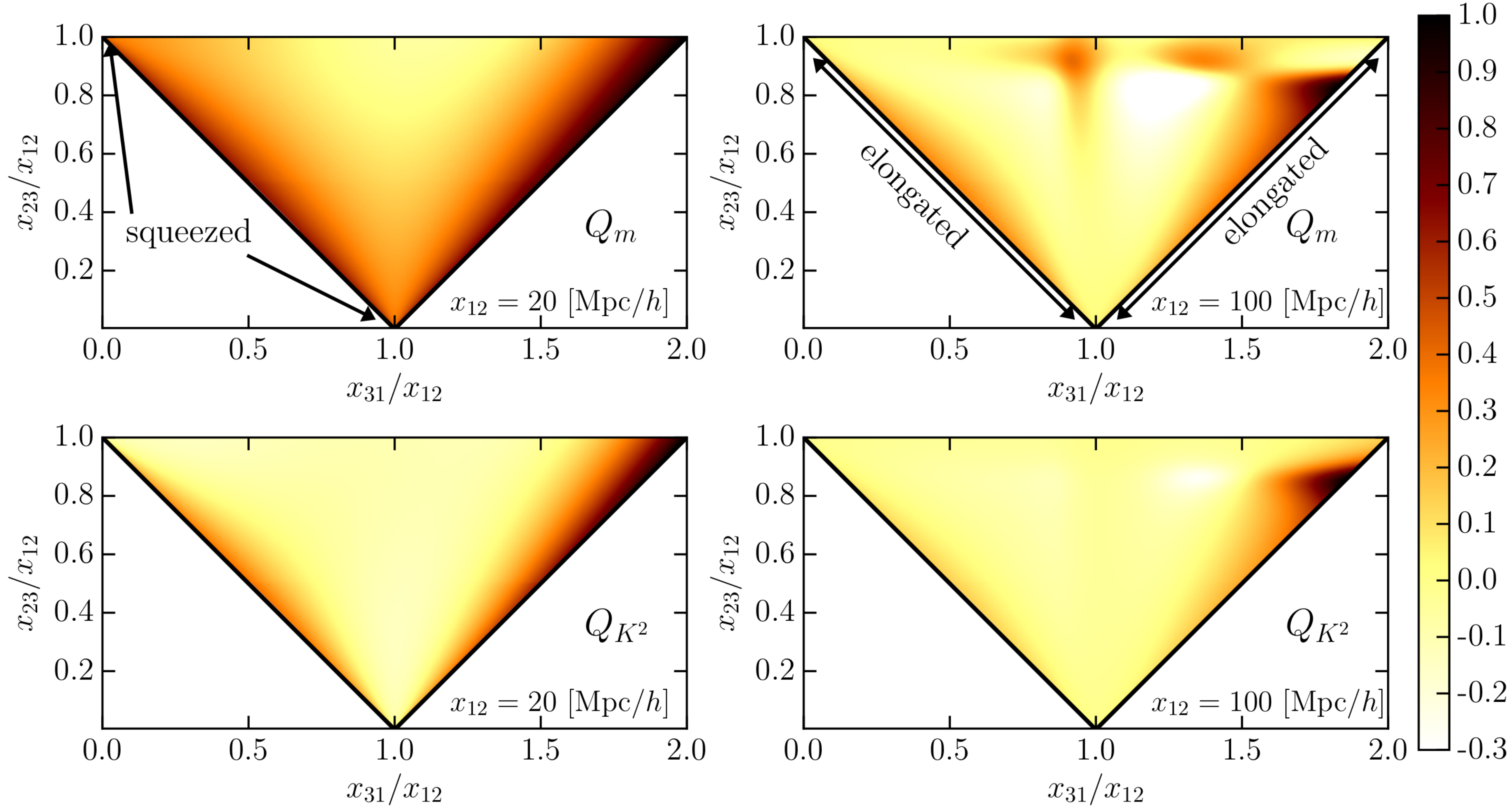
Table 5: Projected uncertainties of the deterministic bias marameters, b_1, b_2, b_{K^2}, and the amplitude of linear power specgrum, $\ln \mathcal{A}$, from current and upcoming galaxy surveys. For all case, we assume $b_1 = 1.5$, and calculate b_2 from the fitting formula in Tab. 6 ($b_2 \simeq -0.69$) and b_{K_2} by assuming Lagrangian LIMD biasing which yields $b_{K^2} = -2/7(b_1 - 1) \simeq -0.14$. Future galaxy surbeys can measure the galaxy bias and the amplitude of the linear power spectrum to a few percent-level accuracy.

survey	redshift	Volume	$10^4 n_g$	$k_{\rm max} = 0.1\, h/{\rm Mpc}$				$k_{\rm max} = 0.2\, h/{\rm Mpc}$			
	z_{cent}	${\rm Gpc}^3/h^3$	$h^{-3}{\rm Mpc}^3$	$100\sigma_{b_1}$	$100\sigma_{b_2}$	$\sigma_{b_{K^2}}$	$\sigma_{\ln \mathcal{A}}$	$100\sigma_{b_1}$	$100\sigma_{b_2}$	$\sigma_{b_{K^2}}$	$\sigma_{\ln \mathcal{A}}$
eBOSS (LRG)	0.8	6.1	4.4	32	45	0.30	0.43	7.0	4.5	0.059	0.093
eBOSS (QSO)	1.4	39	1.5	38	51	0.36	0.51	11	6.5	0.092	0.15
HETDEX	2.7	2.7	3.6	190	260	1.8	2.6	59	35	0.49	0.79
PFS	1.5	8.7	4.6	47	66	0.44	0.62	11	6.7	0.089	0.14
DESI	1.1	40	3.3	18	25	0.17	0.25	4.4	2.7	0.037	0.059
WFIRST	1.9	13	12	35	49	0.32	0.46	6.8	4.4	0.056	0.091
Euclid	1.4	63	5.2	15	20	0.14	0.20	3.3	2.1	0.027	0.044

- Measure bias parameters to a few percent from galaxy surveys.
- Combining Pk and Bk measures the amplitude of fluctuations.

BAO from 3PCF!

$$\begin{aligned}\xi^{(3)}_{hhh}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3) =& b_1^3\xi^{(3)}_m(x_{12},x_{23},x_{31}) + b_1^2 b_2 \xi^{(3)}_{b_2}(x_{12},x_{23},x_{31}) + b_1^2 b_{K^2}\xi^{(3)}_{b_{K^2}}(x_{12},x_{23},x_{31})\\ =& b_1^3\xi^{(3)}_m(x_{12},x_{23},x_{31}) + b_1^2\left\{ b_2\xi_{\rm L}(x_{23})\xi_{\rm L}(x_{31}) + 2b_{K^2}\xi^{(0)}_2(x_{13})\xi^{(0)}_2(x_{23})\left[\mu^2_{23,31}-\frac{1}{3}\right] + (2\ \text{cyclic})\right\}\end{aligned}$$



What else in the review?

- Renormalization of bias:
  Observed bias parameters are not the same as the bare bias parameters in the expansion, but the renormalized one.

- Correlation function bias = PBS bias:
  Bias that we measure from correlation functions are ***exactly*** the same as what we define through Peak-Background Split method: that is, bias is the fractional change of number density of galaxies as a response to the change of background cosmology.

- Other method of measuring bias with explicit smoothing:

  - Bias from the moments of density field = bias from the scatter plot (matter density vs. halo density)

What else in the review?

- Theoretical methods of calculating bias:
  - Excursion set approach (Bond+)
  - Peak theory (BBKS)
  - Excursion set peak/Peak-Patch
- Connecting galaxy clustering to large-scale structure observables:
  - Halo Occupation Distribution
  - Effect of Astrophysical selection (tidal alignment; radiative transfer)
  - general relativistic projection effect
  - Redshift-space distortion
  - wide-angle (or light-cone) effect

Large-Scale Galaxy Bias

Vincent Desjacques[a], Donghui Jeong[b], Fabian Schmidt[c]

[a]*Département de Physique Théorique and Center for Astroparticle Physics, Université de Genève, 24 quai Ernest Ansermet, CH-1221 Genève 4, Switzerland*
[b]*Department of Astronomy and Astrophysics, and Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA*
[c] *Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, 85748 Garching, Germany*

Abstract

We present a comprehensive review of galaxy bias, that is, the statistical relation between the distribution of galaxies and matter. Moreover, we focus on large scales where cosmic density fields are quasi-linear. On these scales, the clustering of galaxies can be described by a perturbative bias expansion, and the complicated physics of galaxy formation can be absorbed into a finite set of coefficients of the expansion, called *bias parameters*. The review begins with a pedagogical proof of this very important result, which forms the basis of the rigorous perturbative description of galaxy clustering, under the assumptions of General Relativity and Gaussian, adiabatic initial conditions. Key components of the bias expansion are quantities that are nonlocal in the matter density, in particular tidal fields and their time derivatives. This derivation is followed by a presentation of the peak-background split in its general form, which elucidates the physical meaning of the bias parameters, and a detailed description of the connection between bias parameters and galaxy (or halo) statistics. We then review the excursion set formalism and peak theory which provide predictions for the values of the bias parameters. In the remainder of the review, we consider the generalizations of galaxy bias required in the presence of various types of cosmological physics that go beyond the standard ΛCDM paradigm with adiabatic, Gaussian initial conditions: primordial non-Gaussianity, massive neutrinos, baryon-CDM isocurvature perturbations, dark energy, and modified gravity. Finally, we discuss how the description of galaxy bias in the galaxies' rest frame is related to observed clustering statistics measured from the observed angular positions and redshifts in actual galaxy catalogs.

Keywords:
