Galaxy Clustering Beyond the Power Spectrum or The slow coming-of-age of the galaxy bispectrum



Emiliano Sefusatti

Astronomical Observatory of Trieste



Paris-Saclay Astroparticle Symposium 2022 Wednesday November 9th



Cosmological constraints from the galaxy power spectrum





$$P_g(k) = P_{template}(k/\alpha)$$

$$\alpha = \frac{D_V(z)r_s^{\text{fid}}(z_d)}{D_V^{\text{fid}}(z)r_s(z_d)} \qquad D_V(z) = \left[(1+z)^2 D_A^2(z)\frac{cz}{H(z)}\right]^{1/3}$$

Geometrical probe of expansion history

Cosmological constraints from the galaxy power spectrum



$$P_g(k,\mu) \simeq (b + f\mu^2)^2 P_L(k) = \sum_{\ell=0,2,4} P_\ell(k) \mathcal{L}_\ell(\mu)$$

Dynamical probe of expansion history

Cosmological constraints from the galaxy power spectrum



The galaxy power spectrum in Perturbation Theory

$$\begin{split} \delta_{h}^{G}(\mathbf{x}) &= b_{1}\delta(\mathbf{x}) + b_{\nabla^{2}\delta}\nabla^{2}\delta(\mathbf{x}) + \epsilon(\mathbf{x}) + \frac{b_{2}}{2}\delta^{2}(\mathbf{x}) + b_{\mathcal{G}_{2}}\mathcal{G}_{2}(\mathbf{x}) + \epsilon_{\delta}(\mathbf{x})\delta(\mathbf{x}) \\ &+ \frac{b_{3}}{6}\delta^{3}(\mathbf{x}) + b_{\mathcal{G}_{3}}\mathcal{G}_{3}(\mathbf{x}) + b_{(\mathcal{G}_{2}\delta)}\mathcal{G}_{2}(\mathbf{x})\delta(\mathbf{x}) + b_{\Gamma_{3}}\Gamma_{3}(\mathbf{x}) + \epsilon_{\delta^{2}}(\mathbf{x})\delta^{2}(\mathbf{x}) + \epsilon_{\mathcal{G}_{2}}(\mathbf{x})\mathcal{G}_{2}(\mathbf{x}) \\ P_{h}^{G}(k) &= b_{1}^{2}\left[P_{0}(k) + P_{m}^{1-\mathrm{loop}}(k)\right] + b_{1}b_{2}\mathcal{I}_{\delta^{2}}(k) + 2b_{1}b_{\mathcal{G}_{2}}\mathcal{I}_{\mathcal{G}_{2}}(k) \\ &+ \frac{1}{4}b_{2}^{2}\mathcal{I}_{\delta^{2}\delta^{2}}(k) + b_{\mathcal{G}_{2}}^{2}\mathcal{I}_{\mathcal{G}_{2}\mathcal{G}_{2}}(k) + b_{2}b_{\mathcal{G}_{2}}\mathcal{I}_{\delta^{2}\mathcal{G}_{2}}(k) + 2b_{1}(b_{\mathcal{G}_{2}} + \frac{2}{5}b_{\Gamma_{3}})\mathcal{F}_{\mathcal{G}_{2}}(k). \end{split}$$



A glimpse ... *in real space*

The galaxy power spectrum in Perturbation Theory

Test of 1-loop power spectrum in EFTofLSS HOD galaxies (CMASS, LOWZ) 566 Gpc³ $h^{-3} \sim 100$ times BOSS

$$P_{\ell}(k) = P_{\ell}^{\text{tree}}(k) + P_{\ell}^{\text{loop}}(k) + P_{\ell}^{\text{ctr}}(k)$$



Nishimichi et al. (2020)

3 bias parameters $\delta_g(\mathbf{k}) = b_1$ 3 counterterms + cosmological parameters

$$\delta_g(\mathbf{k}) = b_1 \delta(\mathbf{k}) + \frac{b_2}{2} \delta^2(\mathbf{k}) + b_{\mathcal{G}_2} \mathcal{G}_2(\mathbf{k})$$

Non-Gaussianity





(An incomplete list)

The galaxy bispectrum: $\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3) \rangle = \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B(k_1, k_2, k_3)$



Non-Gaussianity



The power spectrum cannot distinguish a cosmological simulation from a (properly designed) random walk

ES & Scoccimarro (2005)

Perturbation Theory model

Anisotropies

Window convolution

Covariance

State-of-the-art: recent BOSS analyses

Beyond ΛCDM

Euclid

Bispectrum S/N is comparable to the power spectrum

but the signal is distributed over a large number of triangular configurations

To extract enough information we must **get to small scales!**

$$\left(\frac{S}{N}\right)_{P}^{2} = \sum_{k}^{k_{\text{max}}} \frac{P^{2}(k)}{\Delta P^{2}(k)}$$
$$\left(\frac{S}{N}\right)_{B}^{2} = \sum_{\text{triangles}}^{k_{\text{max}}} \frac{B^{2}(k_{1}, k_{2}, k_{3})}{\Delta B^{2}(k_{1}, k_{2}, k_{3})}$$





The **matter** bispectrum



Fry (1984)

The **matter** bispectrum



The **matter** bispectrum at one-loop



Alkhanishvili *et al.* (2019)

Galaxy bias

Non-Gaussianity from nonlinear bias

$$\delta_g(\boldsymbol{k}) = b_1 \delta(\boldsymbol{k}) + \frac{b_2}{2} \delta^2(\boldsymbol{k}) + b_{\mathcal{G}_2} \mathcal{G}_2(\boldsymbol{k})$$

quadratic bias, local & nonlocal

 $\rightarrow \langle \delta_g \delta_g \delta_g \rangle = b_1^3 \langle \delta \delta \delta \rangle + b_1^2 b_2 \langle \delta \delta \delta^2 \rangle + \dots \\ B_g(k_1, k_2, k_3) = b_1^3 B_m(k_1, k_2, k_3) + b_1^2 b_2 P_L(k_1) P_L(k_2) + \\ + 2 b_1^2 b_{\mathcal{G}_2} S(\mathbf{k_1}, \mathbf{k_2}) P_L(k_1) P_L(k_2) + \text{perm.} + \text{loop corrections}$

Galaxy bias

Non-Gaussianity from nonlinear bias

$$\delta_g(\mathbf{k}) = b_1 \delta(\mathbf{k}) + \frac{b_2}{2} \delta^2(\mathbf{k}) + b_{\mathcal{G}_2} \mathcal{G}_2(\mathbf{k})$$

quadratic bias, local & nonlocal

$$\rightarrow \quad \langle \delta_g \delta_g \delta_g \rangle = b_1^3 \langle \delta \delta \delta \rangle + b_1^2 b_2 \langle \delta \delta \delta^2 \rangle + \dots \\ B_g(k_1, k_2, k_3) = b_1^3 B_m(k_1, k_2, k_3) + b_1^2 b_2 P_L(k_1) P_L(k_2) + \\ + 2 b_1^2 b_{\mathcal{G}_2} S(\mathbf{k_1}, \mathbf{k_2}) P_L(k_1) P_L(k_2) + \text{perm.} + \text{loop corrections}$$



$$Q_g(k_1, k_2, k_3) = \frac{1}{b_1}Q(k_1, k_2, k_3) + \frac{b_2}{b_1^2}$$

This allows to break the degeneracy between b_1 and A_s in the power spectrum, as $P_g(k) \simeq b_1^2 P_L(k) \sim b_1^2 A_s$ but also to determine b_2 Measurements of the galaxy bispectrum in N-body simulations identify a problem in our understanding of galaxy bias:

In a **local bias** model, the linear b_1 bias determined from the power spectrum was inconsistent with the one determined from the bispectrum

$$\begin{split} P_g(k) &= b_1^2 \, P_L(k) \\ B_g(k_1,k_2,k_3) &= b_1^3 \, B_m(k_1,k_2,k_3) + \\ &+ b_1^2 \, b_2 \, P_L(k_1) \, P_L(k_2) + \text{perm.} \end{split}$$

Nonlocal bias Chan *et al.* (2012), Baldauf *et al.* (2012)



$$\delta_g(\mathbf{k}) = b_1 \delta(\mathbf{k}) + \frac{b_2}{2} \delta^2(\mathbf{k}) + b_{\mathcal{G}_2} \mathcal{G}_2(\mathbf{k})$$

P at 1-loop, *B* tree-level Halos test on $1000 h^{-3}$ Gpc³ of cumulative volume



The bispectrum is expected to reduce degeneracies in the power spectrum loop corrections

$$\begin{split} b_1, b_2, b_{\mathcal{G}_2} \\ & \checkmark \\ P_{\ell}(k) = P_{\ell}^{\text{tree}}(k) + P_{\ell}^{\text{loop}}(k) + P_{\ell}^{\text{ctr}}(k) \end{split}$$

Limited reach on such a large volume: $k_{max}^B \simeq 0.09 \, h \text{Mpc}^{-1}$

Significant improvement over *P*, but in real space!

See also Ivanov et al. (2022)

Test of 1-loop bispectrum bias model in real space

```
HOD galaxies (CMASS, LOWZ) & halos 6 \,\text{Gpc}^3 h^{-3}
```

8 parameters (tree-level *B*) 15 parameters (one-loop *B*)



Eggemeier et al. (2021)

One-loop corrections greatly extend the reach of the model and its potential to constrain its parameters (despite their larger number)

but, again, this is still real space ...

Galaxy density in redshift space: more nonlinearity

$$\delta_s(\mathbf{k}) = Z_1(\mathbf{k})\delta_L(\mathbf{k}) + \int d^3q \, Z_2(\mathbf{q}, \mathbf{k} - \mathbf{q})\delta_L(\mathbf{q})\delta_L(\mathbf{k} - \mathbf{q}) + \dots$$

$$\begin{split} Z_{1}(\mathbf{k}) &= b_{1} + f\mu^{2} \,, \\ Z_{2}(\mathbf{k}_{1}, \mathbf{k}_{2}) &= \frac{b_{2}}{2} + b_{1}F_{2}(\mathbf{k}_{1}, \mathbf{k}_{2}) + b_{\mathcal{G}_{2}}S(\mathbf{k}_{1}, \mathbf{k}_{2}) + f\mu_{12}^{2}G_{2}(\mathbf{k}_{1}, \mathbf{k}_{2}) + \\ &+ \frac{f\mu_{12}k_{12}}{2} \begin{bmatrix} \mu_{1}}{k_{1}}Z_{1}(\mathbf{k}_{2}) + \frac{\mu_{2}}{k_{2}}Z_{1}(\mathbf{k}_{1}) \end{bmatrix} & \begin{array}{c} \textit{Redshift-space} \\ \textit{PT kernels} \\ \end{split}$$



 $B_s(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2 Z_1(\mathbf{k}_1) Z_1(\mathbf{k}_2) Z_2(\mathbf{k}_1, \mathbf{k}_2) P_L(k_1) P_L(k_2) + 2 \text{ perm.}$

Galaxy bispectrum in redshift space

Galaxy density in redshift space: more nonlinearity

$$\delta_s(\mathbf{k}) = Z_1(\mathbf{k})\delta_L(\mathbf{k}) + \int d^3q \, Z_2(\mathbf{q}, \mathbf{k} - \mathbf{q})\delta_L(\mathbf{q})\delta_L(\mathbf{k} - \mathbf{q}) + \dots$$

$$\begin{split} Z_1(\mathbf{k}) &= b_1 + f\mu^2 \,, \\ Z_2(\mathbf{k}_1, \mathbf{k}_2) &= \frac{b_2}{2} + b_1 F_2(\mathbf{k}_1, \mathbf{k}_2) + b_{\mathcal{G}_2} S(\mathbf{k}_1, \mathbf{k}_2) + f\mu_{12}^2 G_2(\mathbf{k}_1, \mathbf{k}_2) + \\ &+ \frac{f\mu_{12}k_{12}}{2} \begin{bmatrix} \mu_1 \\ k_1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ k_2 \end{bmatrix} + \frac{\mu_2}{k_2} Z_1(\mathbf{k}_1) \end{bmatrix} & \begin{array}{c} \textit{Redshift-space} \\ \textit{PT kernels} \\ \end{split}$$

$$\Rightarrow B_{s}(\mathbf{k}_{1}, \mathbf{k}_{2}, \mathbf{k}_{3}) = B_{s}^{(\text{det})}(\mathbf{k}_{1}, \mathbf{k}_{2}, \mathbf{k}_{3}) + B_{s}^{(\text{stoch})}(\mathbf{k}_{1}, \mathbf{k}_{2}, \mathbf{k}_{3})$$

$$B_{s}^{(\text{det})}(\mathbf{k}_{1}, \mathbf{k}_{2}, \hat{n}) = 2 Z_{1}(\mathbf{k}_{1}) Z_{1}(\mathbf{k}_{2}) Z_{2}(\mathbf{k}_{1}, \mathbf{k}_{2}) P_{L}(k_{1}) P_{L}(k_{2}) + 2 \text{ perm.}$$

$$B_{s}^{(\text{stoch})}(\mathbf{k}_{1}, \mathbf{k}_{2}, \hat{n}) = \frac{1}{\bar{n}} \left[(1 + \alpha_{1}) b_{1} + (1 + \alpha_{3}) f \mu^{2} \right] Z_{1}(\mathbf{k}_{1}) P_{L}(k_{1}) + 2 \text{ perm.} + \frac{1 + \alpha_{2}}{\bar{n}^{2}}$$

$$Lot's of fun here!$$

$$B_s(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = B_s(k_1, k_2, k_3, \theta_1, \phi_{12})$$

The orientation of the triangle w.r.t. the line-of-sight now matters

Different choices are possible (see e.g. Hashimoto *et al.*, 2017, Gualdi & Verde, 2020)

We follow Scoccimarro *et al.* (1999), with the FFT-based estimator of Scoccimarro (2015).



$$B_{s}(k_{1}, k_{2}, k_{3}, \theta_{1}, \phi_{12}) = \sum_{\ell, m} B_{\ell, m}(k_{1}, k_{2}, k_{3}) Y_{\ell, m}(\theta_{1}, \phi_{12})$$
$$\mu_{1} \equiv \mu \equiv \cos \theta_{1}$$

Redshift-space: bispectrum *monopole*



FIG. 7. Same as Fig. 5 but with the covariance rescaled by 100 to match the BOSS survey volume.

Ivanov *et al.* (2022)



Test of bispectrum multipoles: halos on $1000 h^{-3}$ Gpc³ of eumulative volume



M

See also Gualdi & Verde (2020), Gualdi et al. (2021), D'Amico et al. (2022)

Redshift-space: bispectrum monopole & quadrupole



D'Amico *et al.* (2022)

The convolution of the bispectrum prediction with the window function is a problem ..

$$\tilde{B}(\mathbf{k}_1, \mathbf{k}_2) = \int \frac{d^3 p_1}{(2\pi)^3} \int \frac{d^3 p_2}{(2\pi)^3} B_W(\mathbf{k}_1 - \mathbf{p}_1, \mathbf{k}_2 - \mathbf{p}_2) B(\mathbf{p}_1, \mathbf{p}_2)$$

Three approaches so far:

- "Tree-level" approximation (Gil-Marín *et al.*, 2015) $\tilde{B} \simeq 2Z_1(\mathbf{k}_1)Z_1(\mathbf{k}_2)Z_2(\mathbf{k}_1, \mathbf{k}_2)\tilde{P}(k_1)\tilde{P}(k_2) + \text{perm}.$
- Windowless estimator (Philcox, 2021)
- Exact convolution (Pardede *et al.*, 2022)



Pardede, Rizzo et al. (2022)

The bispectrum signal is **distributed over a large number of configurations**



A robust, numerical estimates of such a **large covariance matrix** requires a **large number of mocks**

CMASS sample (z_{eff}=0.57)

Covariance

As an alternative, a **theoretical covariance** should be considered ...

Gaussian covariance of bispectrum multipoles in a box (no window) Rizzo, Moretti, Pardede *et al.* (2022) See also Sugiyama *et al.* (2022)



Or **compression methods** Gualdi *et al.* (2018; 2019)

- The model tree-level *vs* **one-loop** *vs* **phenomenological**
- Anisotropy monopole *vs* monopole + quadrupole
- Window function convolution approximated *vs* **exact** *vs* **windowless**
- Covariance numerical *vs* theoretical (?)

BOSS

Analysis of BOSS data: Gil-Marin et al. (2017)

- *data:* monopole (825 triangles, $\Delta k = 0.01h \,\mathrm{Mpc}^{-1}$)
- model: fit to N-body + tree-level bias & RSD (+AP) $0.03 h \text{Mpc}^{-1} \le k \le 0.18 h \text{Mpc}^{-1}$ $0.03 h \text{Mpc}^{-1} \le k \le 0.22 h \text{Mpc}^{-1}$
- *window*: approximation

$$\widetilde{B} \simeq Z_1(\mathbf{k}_1) Z_1(\mathbf{k}_2) Z_2(\mathbf{k}_1, \mathbf{k}_2) \widetilde{P}(k_1) \widetilde{P}(k_2)$$

- *covariance*: numerical (2048 Patchy mocks)
- analysis: template fitting

 $\{b_1, b_2, A_{\text{noise}}, \sigma_{\text{FoG}}^P, \sigma_{\text{FoG}}^B, f, \sigma_8, \alpha_{\parallel}, \alpha_{\perp}\}.$



Significant improvement, up to 50% (for CMASS) on RSD parameters

Analysis of BOSS data: Philcox & Ivanov (2022)

- *data:* monopole (62 triangles, $\Delta k = 0.01hMpc^{-1}$, $0.01 \le k \le 0.08 hMpc^{-1}$)
- *model*: tree-level
- *window*: windowless estimator
- *covariance*: numerical (2048 Patchy mocks)
- analysis: full-shape
 3/4 cosmo + 13 bias/noise
 parameters



13% improvement on σ_8

Analysis of BOSS data: D'Amico et al. (2022)

- data: monopole & quadrupole (150 triangles for B_0 , 9 for B_2 , $\Delta k = 0.02 h \text{Mpc}^{-1}$, $0.02 \le k \le 0.21 h \text{Mpc}^{-1}$ for CMASS)
- *model*: 1-loop for B_0 , tree-level for B_2
- *window*: approximation
- *covariance*: numerical (2048 Patchy mocks)
- analysis: full-shape
 3 cosmo + 12 bias/noise
 parameters



Significant improvement (30% for σ_8) from one-loop B_0 , rather than B_2

Beyond ΛCDM

Beyond ACDM: Tests with Primordial non-Gaussainity

Test of the power spectrum & bispectrum model in real space

Eos simulations, $80 h^{-3} \,\mathrm{Gpc}^3$ Halo catalogs

Significant improvement (factor of 5) over power spectrum only

Also from the reduction of the $f_{\rm NL} - b_{\phi}$ degeneracy



Moradinezhad *et al*. (2021)

BOSS analysis beyond ACDM: Primordial non-Gaussainity

- 1. The bispectrum greatly improves constraints on local PNG and ...
- 2. ... it *allows* those on single-field inflation models



D'Amico *et al.* (2022) Cabass *et al.* (2022A, 2022B)



Other directions include, for instance, a bispectrum dipole from GR effects See e.g. Clarkson *et al.* (2019)

Euclid

Fully anisotropic bispectrum: $B_s(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$

 $k_{\rm max} = 0.15 \, h \, {\rm Mpc^{-1}}$



Significant improvement on dark energy equation of state parameter

• The model

tree-level PT vs one-loop PT vs phenomenological

we probably need to go beyond tree-level, but loop + AP integrations are challenging

Anisotropy

monopole *vs* **monopole** + **quadrupole** *we already have multipoles estimators, so* ...

Window function

approximated *vs* **exact** *vs* **windowless** *it would be* very *nice to test both exact convolution and windowless*

Covariance

numerical *vs* theoretical *probably a theoretical approach or a compression method are inevitable*

Alternative estimators

Skew-spectra (Schmittfull *et al.*, 2015; Moradinezhad *et al.* 2020; ...) Tri-polar Spherical Harmonic Decomposition (Sugiyama *et al.*, 2017) Modal estimator (Fergusson *et al.*, 2012; Byun *et al.*, 2021) ... and more ...