Analysis of the response of void BAO to systematic effects in the SDSS observations using mock datasets

Daniel Forero-Sánchez[†]

[†]Laboratory of Astrophysics (LASTRO)

École Polytechnyque Fédérale de Lausanne (EPFL)
Lausanne, Suisse

January 6, 2020

Outline I

Introduction I

- The use of voids + matter ones $\Rightarrow \sim 10\%$ improvement of error & 20% survey size increase.
- Voids could be less sensitive to systematic effects than matter.
- We want to use mocks to check if there is a difference in the BAO peak shift due to systematical effects between matter tracers and voids.
- Systematics can greatly affect the measurement of the 2PCF in real data

Voids I

Definition is controversial, we use a **geometrical** one.
Delaunay Triangulation $\longrightarrow$ Voids are circumspheres in the simplices with tracers as vertices.

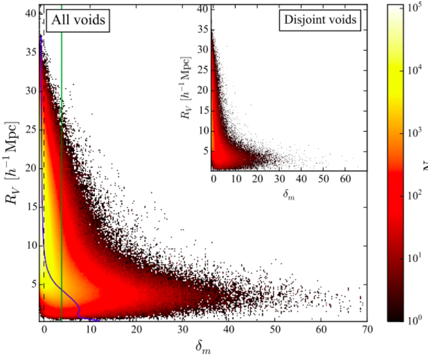
Small voids

- Radius cut: $R_c < 8\,h^{-1}\text{Mpc}$
- Located in regions with high δ_{dm}
- Correlated with galaxy sample
- Are actually Dark Matter

Big voids

- Radius cut: $R_c > 15.5\,h^{-1}\text{Mpc}$
- Located in regions with low δ_{dm}
- Anticorrelated with galaxy sample
- Are actually empty

Voids II



Number of voids, N_V, as a function of Dark Matter density, δ_{dm}, and void radius, R_V. Taken from (Zhao et al., 2016)

How do we know?

- Large voids are constrained to regions with low matter density.
- Small void population spans a wide range of densities.

Data I

We use 1000 **EZ mocks** Zhao and Chuang (2020).

- Displacement field from Zel'dovich
- PDF extraction from n-body simulation
- Halo position assignment

Table: Fiducial cosmology parameters used to produce the EZmocks used in this work.

Parameter	Value
Ω_m	0.307115
Ω_b	0.048206
h	0.6777
σ_8	0.8225
n_s	0.9611

Systematics: Fiber collisions I

- Objects too close in the sky ($r_{cp} < 62''$) can't be seen at the same time due o width of the fiber.
- One is measured by the plate. The other one(s) could still be measured by other plate.
- Define $\mathrm{TSR} = \frac{\text{measured targets}}{\text{total number of targets}}$, per sector.
- Compensate this effect with $w_{cp} \approx \mathrm{TSR}^{-1}$
  Actually $w_{cp} = \frac{\text{total number of targets}}{\text{measured targets}}$ but defined per **collision group**.

Systematics: Redshift Failures I

- Errors in the spectroscopic pipeline $\Rightarrow$ $\mathrm{SSR} < 1$.
- Two sources of error: Observational conditions & Position of fiber
- This is corrected by the weight

$$w_{noz} \equiv (\mathrm{SSR}_{\mathrm{obs}}\mathrm{SSR}_{\mathrm{pos}})^{-1}.$$

Systematics: Angular Photometric I

- Represented as HEALPIX maps containing different photometric parameters, p_i.
- Parameters are combined as

$$y^k = \epsilon + \sum_i c_i p_i^k$$

where the model weights, c_i are optimized (per chunk) such that $n_{\mathrm{dat},k} \approx n_{\mathrm{ran},k} y^k$, where k indicates the pixels.

- The photometric weights designed to partially correct for these effects are defined as

$$w_{\mathrm{systot}} = \left(y^k\right)^{-1}.$$

Systematics: Normalization I

- Completeness weights are defined as:

$$w_{\mathrm{comp}} = w_{\mathrm{systot}} w_{cp} w_{noz}.$$

- Some normalization is done (on w_{systot} before w_{comp} and on w_{noz} after).
- Invalid objects are set $w_{cp}, w_{noz} = 0$.
- Only elements with $\mathrm{SSR} \geq 0$, $z \in (0.6, 1.1)$ and completeness $> 50\%$ are selected.
- The dependence on $n(z)$ is corrected by

$$w_{\mathrm{FKP}} \equiv \frac{1}{1 + n(z)P_0}; \quad P_0 = 4000\, h^{-3}\mathrm{Mpc}^3.$$

Systematics: Application I

- We compute $w_{\text{comp}}^{\text{ALLSYST}} = w_{\text{systot}} w_{cp} w_{noz} w_{\text{FKP}}$ and the effective number of tracers in each chunk

$$n_{eff} = \sum_{i=1}^{N_{chunk}} w_{\text{comp},i}^{\text{ALLSYST}},$$

where N_{chunk} is the number of tracers in the chunk considered.

- We then compute

$$w_{\text{comp}}^{\text{PARTIAL}} = w_{\text{FKP}} \prod_{s \in \mathcal{S}'} w_s,$$

Systematics: Application II

where $\mathcal{S}' \subseteq \mathcal{S}$ is the subset of systematic effects to be considered and $\mathcal{S} = \{\mathrm{systot},\ noz,\ cp\}$. In this case we compute the new effective number of tracers

$$n'_{eff} = \sum_{i=1}^{N_{chunk}} w^{\mathrm{PARTIAL}}_{\mathrm{comp},i}.$$

- Objects in the catalog with $w^{\mathrm{PARTIAL}}_{\mathrm{comp}} = 0$ or $\texttt{veto} = 0$ (such as objects outside the survey area or masked by bright stars) are removed.
- We then normalize the completeness weights in each chunk by the corresponding n_{eff}/n'_{eff} to keep the original effective number of tracers.

Catalog generation I

- Export catalogs with

$$\mathrm{RA},\ \mathrm{DEC},\ z,\ w_{cp}w_{\mathrm{FKP}},\ w_{cp},\ w_{\mathrm{FKP}},\ n(z)$$

- Use DIVE to extract void catalogs with (we use $\Omega_m = 0.31$)

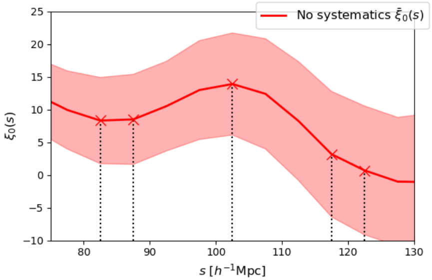
$$\mathrm{RA},\ \mathrm{DEC},\ z,\ R$$

- Mask void catalogs
- Create void randoms
  - Combine 100 void mocks
  - Divide into z bins
  - divide each bin in R bins
- Split "vertically" into $\mathrm{RA},\ \mathrm{DEC} \mid z,\ R$.
- Shuffle one of the two halves.
- Recombine halves and all bins.
- Randomly choose 2 700 000 with $R > R_c (= 15.5\, h^{-1}\mathrm{Mpc})$

Radius cut I

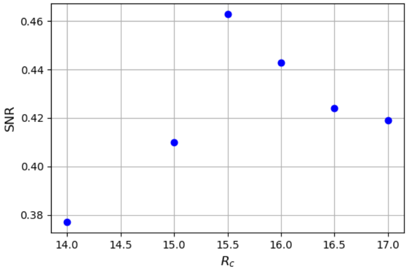
Analyze signal-to-noise ratio $(\mathrm{SNR} \equiv \frac{\langle S \rangle}{\sigma_S})$ for 100 void mocks with different R_c. Use definition of signal, S, as in Liang, Zhao, Chuang, Kitaura, and Tao (2016).

$$S = \xi_0(s^{\mathrm{BAO}}) - \frac{\xi_0(s_1^{\mathrm{dl}}) + \xi_0(s_2^{\mathrm{dl}}) + \xi_0(s_1^{\mathrm{dr}}) + \xi_0(s_2^{\mathrm{dr}})}{4} \quad (1)$$



$s_1^{\mathrm{dl}} = 82.5$, $s_2^{\mathrm{dl}} = 87.5$, $s_1^{\mathrm{BAO}} = 102.5$, $s_1^{\mathrm{dr}} = 117.5$ and $s_2^{\mathrm{dr}} = 122.5\,h^{-1}\mathrm{Mpc}$ respectively from left to right.

Radius cut II



SNR vs low radius cut R_c. Upper cut is always set to $50\,h^{-1}$Mpc. Maximum SNR is obtained for $R_c = 15.5\,h^{-1}$Mpc.

The BAO model I

We use the model in Zhao et al. (2018):

$$\xi_t(s) = \int \frac{k^2 \mathrm{d}k}{2\pi^2} \frac{\sin ks}{ks} P_t(k) \exp\left(-k^2 a^2\right), \quad a = 1\, h^{-1}\mathrm{Mpc} \tag{2}$$

$$P_t(k) = \left\{ [P_{\rm lin}(k) - P_{\rm nw}(k)] \exp\left(\frac{-\Sigma_{\rm nl}^2 k^2}{2} \right) + P_{\rm nw}(k) \right\} \frac{P_{\rm t,nw}(k)}{P_{\rm lin,nw}(k)}, \tag{3}$$

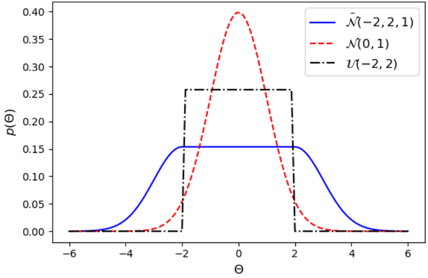
$$\xi_{\rm model}(s) \equiv B^2 \xi_t(\alpha s) + A(s), \quad A(s) = \frac{a_1}{s^2} + \frac{a_2}{s} + a_3 \tag{4}$$

	Galaxies	Voids
$\frac{P_{\rm t,nw}(k)}{P_{\rm lin,nw}(k)}$	1	$1 + ck^2$

Parameter fitting I

Bayesian inference for parameter Θ:

$$p(\Theta|X) = \frac{p(X|\Theta)p(\Theta)}{p(X)} = \frac{\mathcal{L}(X|\Theta)p(\Theta)}{\mathcal{Z}}, \tag{5}$$



Examples of the different kinds of priors used in the fitting.

Parameter fitting II

Voids

$$p(\Sigma_{\rm nl}) = \mathcal{U}(0, 20) \tag{6}$$
$$p(B) = \mathcal{N}(2, 0.15) \tag{7}$$
$$p(\alpha) = \mathcal{U}(0.8, 0.12) \tag{8}$$
$$p(c) = \bar{\mathcal{N}}(-500, 1000, 100) \tag{9}$$

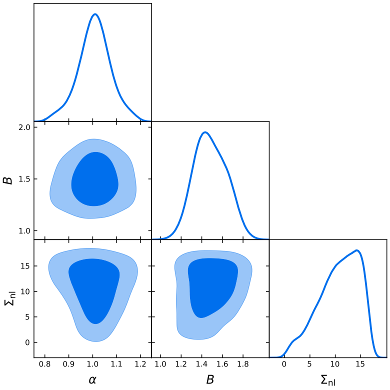
Galaxies

$$p(\Sigma_{\rm nl}) = \mathcal{U}(5, 17) \tag{10}$$
$$p(B) = \bar{\mathcal{N}}(1.4, 1.6, 0.12) \tag{11}$$
$$p(\alpha) = \mathcal{U}(0.8, 0.12) \tag{12}$$

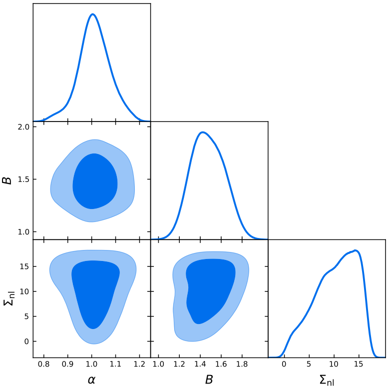
EPFL

Results: Mean 2PCF Galaxies I

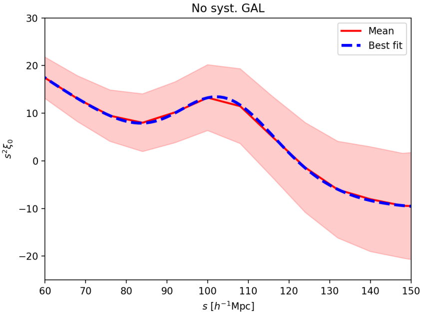
No systematics



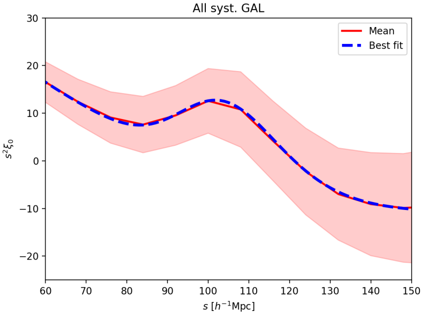
All systematics



Results: Mean 2PCF Galaxies II



$\chi^2_{\rm best}/$dof$= 0.04$

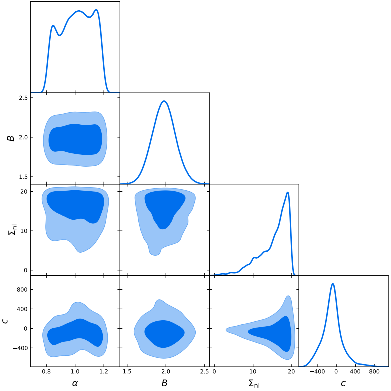


$\chi^2_{\rm best}/$dof $= 0.06$

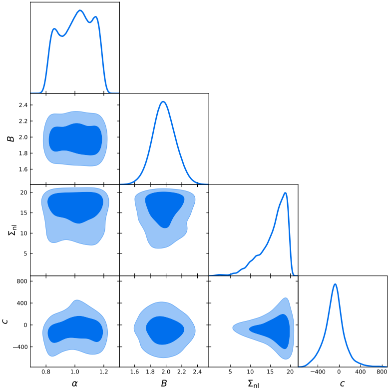
EPFL

Results: Mean 2PCF Voids I

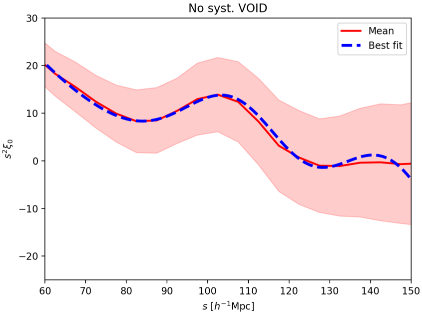
No systematics



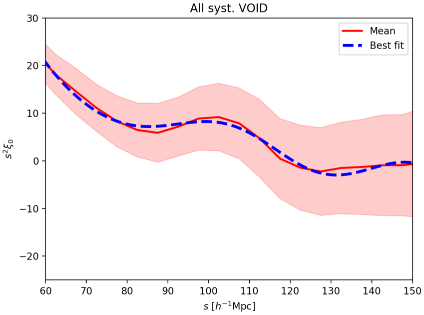
All systematics



Results: Mean 2PCF Voids II



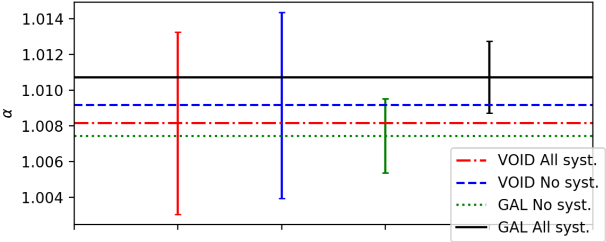
$\chi^2_{\rm best}$/dof= 0.36



$\chi^2_{\rm best}$/dof= 0.42

Results: Mean2PCF Comparison I

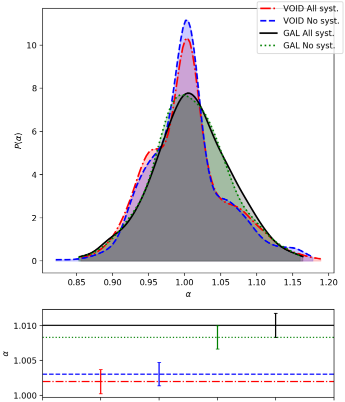
	Galaxies	Voids
$\alpha_{\rm all} - \alpha_{\rm none}$	$(3.34 \pm 2.89) \times 10^{-3}$	$(-9.9 \pm 72.9) \times 10^{-4}$



Fit results for the dilation parameter α when using the mean of the mocks.

Results: Individual 2PCF I

	Galaxies	Voids
$\alpha_{\text{all}} - \alpha_{\text{none}}$	$(1.89 \pm 2.39) \times 10^{-3}$	$(-1.07 \pm 2.39) \times 10^{-3}$



Fit results for the α parameter from fitting each of the mocks individually. Top panel shows the distributions obtained. Bottom panel shows the mean values and the (scaled) standard deviation as error.

Conclusions I

- The fit to the mean 2PCF shows a smaller shift due to systematics in the void case.
- The posteriors $p(\alpha|X)$ in the void case have large widths that make it difficult to be confident in the conslusion above.
- The individual fits show smaller uncertainty in the void measurement and still show the difference in the peak shift.
- Priors can be tweaked to get a better $p(\alpha|X)$ for voids but reduced χ^2 is already close to 1.
- Dataset is noisy as shown by the low SNR
- The analysis should be repeated with other tracers with better SNR (e.g. LRG)

References I

Liang, Y., Zhao, C., Chuang, C.-H., Kitaura, F.-S., & Tao, C. (2016). Measuring baryon acoustic oscillations from the clustering of voids. *Monthly Notices of the Royal Astronomical Society*, *459*(4), 4020–4028. doi:10.1093/mnras/stw884

Zhao, C. & Chuang, C.-h. (2020). The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey : mock catalogues for the eBOSS final Data Release. *in preparation*, *10*(December 2019), 1–10.

Zhao, C., Chuang, C.-h., Kitaura, F.-s., Liang, Y., Pellejero-Ibanez, M., Tao, C., … Yepes, G. (2018). Improving baryon acoustic oscillation measurement with the combination of cosmic voids and galaxies. *19*(February), 1–19. arXiv: 1802.03990. Retrieved from http://arxiv.org/abs/1802.03990

References II

Zhao, C., Tao, C., Liang, Y., Kitaura, F.-S., & Chuang, C.-H. (2016). DIVE in the cosmic web: voids with Delaunay triangulation from discrete matter tracer distributions. *Monthly Notices of the Royal Astronomical Society*, *459*(3), 2670–2680. doi:10.1093/mnras/stw660