Joint constraints on galaxy bias and σ_8 through the N-pdf of the galaxy number density

based on Arnalte-Mur et al. (2016) JCAP, 3, 5

Pablo Arnalte-Mur, **Vicent J. Martínez**, Patricio Vielva,
José L. Sanz, Enn Saar, Silvestre Paredes, and Lluís Hurtado-Gil

May 25, 2016

STATISTICAL CHALLENGES in 21st CENTURY COSMOLOGY

Observatori Astronòmic
Vniversitat de València

Vniversitat de València

Outline

1. Introduction
   - Mesuring bias
   - Our goal
2. The galaxy density N-pdf method
3. Data
   - SDSS main galaxy sample
   - Lognormal simulations
   - Las Damas mock catalogues
4. Test of the method
   - Test on lognormal simulations
   - Test on Las Damas simulations
5. Results for the SDSS data
6. Conclusions

Introduction: Measuring bias

In most of the literature, the bias has been estimated using:

- the 3D-correlation function of galaxies: (Fry & Gaztanaga, 1993; Gaztanaga & Frieman, 1994; Croft et al., 1999);
- counts-in-cells statistics: Sigad et al. (2000); Colless et al. (2001); Marinoni et al. (2005); Kovač et al. (2011); Di Porto et al. (2014);
- the projected correlation function: Norberg et al. (2002); Zehavi et al. (2005, 2011); Arnalte-Mur et al. (2014);
- the bi-spectrum: Guo & Jing (2009); Pollack et al. (2014);
- higher order moments of the galaxy distribution: Szapudi (1998); Verde et al. (2002); Swanson et al. (2008); McBride et al. (2011);
- the use of multivariate probability distributions, typically Gaussian or lognormal distributions, which are well suited priors for Bayesian analyses: Jasche & Wandelt (2013); Ata et al. (2015); Granett et al. (2015); Jasche et al. (2015).

Introduction: Our goal

Our method relies on the use of the whole set of N-pdf of the galaxy number density fluctuations:

This multivariate probability density function depends on the bias parameter and the correlations of the underlying cold dark matter fluctuations.

The estimation

- The N-pdf of the galaxy number density field as a function of the bias and the σ_8 parameters is, in fact, the likelihood of the data (i.e., the galaxy catalogue) given these parameters.
- Therefore optimal parameter estimation can be performed via the maximum-likelihood.

The galaxy density N-pdf method

Observable

Galaxy overdensity field in N cubic cells $\rightarrow \{\Delta_{g,i}\}_{i=1}^{N} \rightarrow \Delta_{g,i} = \frac{\rho_{g,i}}{\bar{\rho_g}}$

Model: probability distribution function of $\vec{\Delta}_g$

- Constant, linear bias: $\delta_g = b\delta_m$
- Log-normal distribution model for the matter density field (Coles & Jones, 1991; Kayo et al., 2001)
- Difference w.r.t. "traditional" counts-in-cells: we take into account the cells' positions and the corresponding correlations between them

Model: probability distribution function of $\vec{\Delta}_g$

$$f(\vec{\Delta}_g) = \frac{1}{(2\pi)^{N/2}\,|\mathbf{H_0}|^{1/2}} \frac{\exp\left(-\frac{1}{2}\vec{y}^t \cdot \mathbf{H_0}^{-1} \cdot \vec{y}\right)}{\prod_{j=1}^{N}\left(\Delta_{g,j} + b - 1\right)}\ ,\ \text{where:}$$

$$y_i = \log\left[\frac{\sqrt{1+\sigma_m{}^2}}{b}\left(\Delta_{g,i} + b - 1\right)\right]$$

$$\mathbf{H_0} = \log(1 + \mathbf{C_m})\ ,\ C_{m,ij} = \xi_m(|\vec{r}_i - \vec{r}_j|)$$

Model: probability distribution function of $\vec{\Delta}_g$

$$f(\vec{\Delta}_g) = \frac{1}{(2\pi)^{N/2} \left|\mathbf{H_0}\right|^{1/2}} \frac{\exp\left(-\frac{1}{2}\vec{y}^t \cdot \mathbf{H_0}^{-1} \cdot \vec{y}\right)}{\prod_{j=1}^{N}\left(\Delta_{g,j} + b - 1\right)} \text{, where:}$$

$$y_i = \log\left[\frac{\sqrt{1+\sigma_m{}^2}}{b}\left(\Delta_{g,i} + b - 1\right)\right]$$

$$\mathbf{H_0} = \log(1 + \mathbf{C_m}) \,,\ C_{m,ij} = \xi_m(|\vec{r}_i - \vec{r}_j|)$$

Parameters of the model

- Galaxy bias $\rightarrow b$
- Cosmological parameters $(\theta_{\text{cosmo}}) \rightarrow \xi_m(r) \rightarrow \mathbf{C_m}$

Parameter estimation

We use a Bayesian approach, knowing that the model distribution function is the *likelihood of the data ($\vec{\Delta}_g$) given a model:*

$$\mathcal{L}(\vec{\Delta}_g|b,\theta_{\text{cosmo}}) = f(\vec{\Delta}_g)$$

- Fixing a fiducial cosmological model:

$$p(b|\vec{\Delta}_g) \propto \mathcal{L}(\vec{\Delta}_g|b)p(b)$$

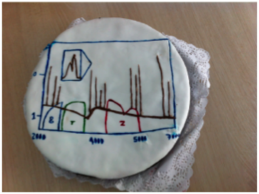
- Allowing σ_8 to vary $\rightarrow \mathbf{C_m} = \left(\frac{\sigma_8}{\sigma_8^{\text{fid}}}\right)^2 \mathbf{C_m}^{\text{fid}}$:

$$p(b,\sigma_8|\vec{\Delta}_g) \propto \mathcal{L}(\vec{\Delta}_g|b,\sigma_8)p(b,\sigma_8)$$

$\rightarrow$ We assume flat priors in all cases, $p(b)$, $p(b,\sigma_8)$
$\rightarrow$ We use a Fisher matrix approach to estimate parameters' uncertainties

A digression: why Bayesian?

Remember this morning talk by Roberto Trotta: the cake analogy:
Hopefully cakes could incorporate very informative priors



Bayesian cake

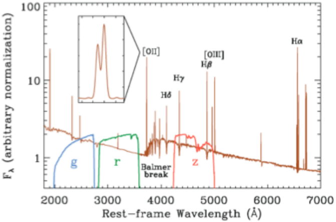


Figure 1.2: Rest-frame spectrum of an ELG showing the blue stellar continuum, the prominent Balmer break, and the numerous strong nebular emission lines. The inset shows a zoomed-in view of the [OII] doublet, which DESI (see Section 1.7) is designed to resolve over the full redshift range of interest, $0.6 < z < 1.6$. The figure also shows the portion of the rest-frame spectrum the DECam grz optical filters would sample for such an object at redshift $z = 1$. Figure from the *DESI Science final design report* at http://desi.lbl.gov/tdr/.

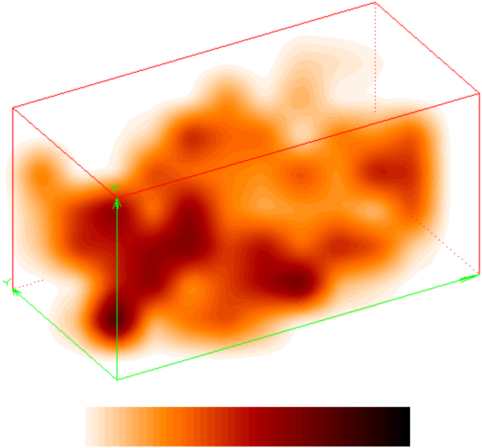
SDSS main galaxy sample

- We used the data provided by the New York University Value Added Catalogue [Blanton et al. (2005)]. We have selected a volume limited sample with $M_r \leq -20$ in the redshift range $0.033 < z < 0.106$, where M_r is the $K + E$ corrected r-band absolute magnitude.
- We study this field using a grid of cubic cells with a side of $30\, h^{-1}\,\mathrm{Mpc}$.
- We compute the completeness for each of the cells (c_i) as the combination of the radial and the angular selection functions and keep only those cells with $c_i \geq 0.8$.
- We obtain the number of galaxies n_i in each of these accepted cells, and estimate the galaxy number density for each cell as

$$\rho_{g,i} = \frac{n_i}{c_i V_c}$$

where $V_c = (30\, h^{-1}\,\mathrm{Mpc})^3$ is the volume of a cell.

SDSS main galaxy sample

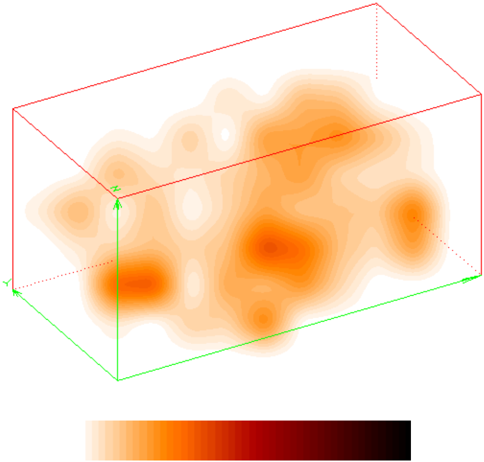


3D projection of the galaxy number density field corresponding to the SDSS catalogue. Colour palette used in the projection corresponds to densities $1 \leq \Delta_g \leq 2$ from left to right.

Lognormal simulations

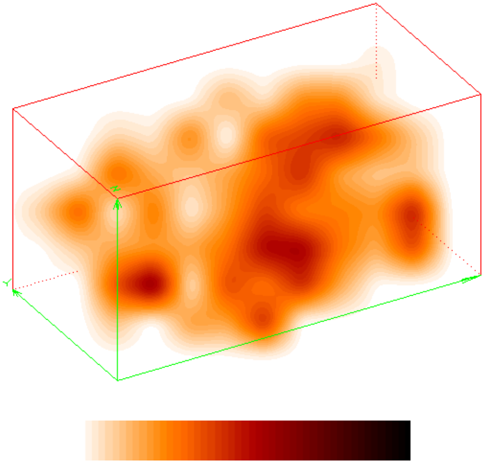
- We generate the lognormal simulations taking into account that a lognormal random field is a local transformation of a Gaussian field.
- We compute the matter correlation function $\xi_m(r)$ via a Fourier transform of $P_m(k)$.
- We then generate the Gaussian random field $\delta_{0,i}$ in a cubic grid using the standard method of generating Gaussian Fourier modes with variances given by $P_0(k)$ and then performing a FFT.
- Finally, we obtain the corresponding matter density field $\Delta_{m,i}$.
- Given a value for the galaxy bias b, we can generate the corresponding galaxy density field $\Delta_{g,i}$. We have explored four input values of the bias: $b = 0.5, 1.0, 1.5, 2.0$.

Lognormal simulations



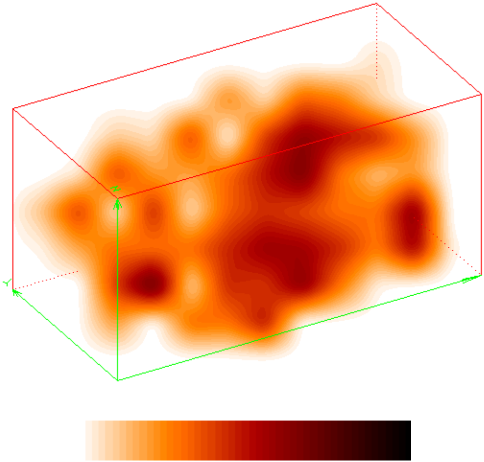
Bias $b = 0.5$

Lognormal simulations



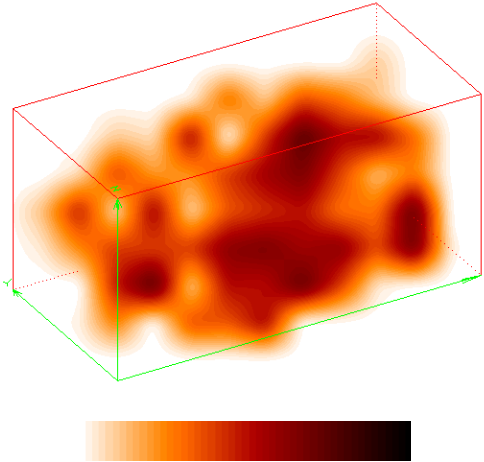
Bias $b = 1.0$

Lognormal simulations



Bias $b = 1.5$

Lognormal simulations

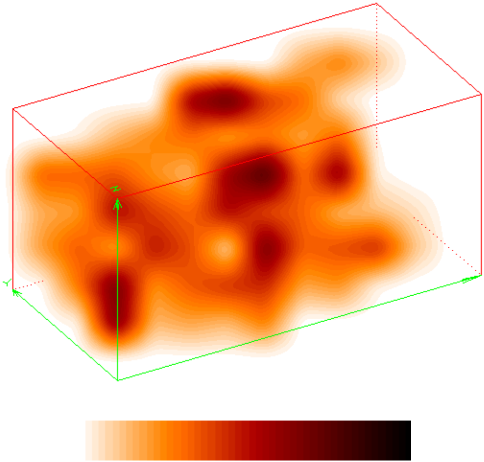


Bias $b = 2.0$

Las Damas mock catalogues

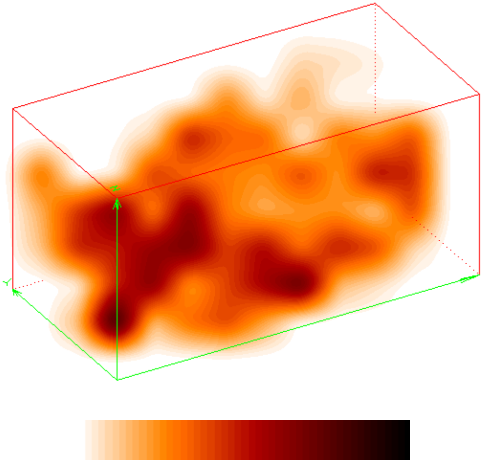
- We use the set of galaxy mocks obtained from the Las Damas simulations http://lss.phy.vanderbilt.edu/lasdamas [McBride et al. (2009)], and in particular the *gamma* release.
- The simulations are populated by galaxies using the halo occupation distribution (HOD) formalism, with the HOD parameters tuned to reproduce the observed number density and projected correlation function $w_p(r_p)$ (at scales $r_p \in [0.3, 30]\, h^{-1}\,\mathrm{Mpc}$, as studied in Zehavi et al. (2011)) of the corresponding SDSS catalogues.
- We end up with a total of 120 mock galaxy catalogues and we compute the galaxy density field $\Delta_{g,i}$.

Las Damas mock catalogues



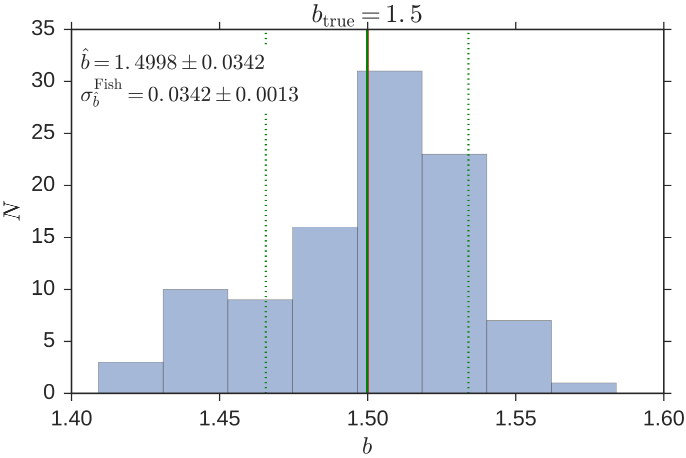
Las Damas

SDSS main galaxy sample



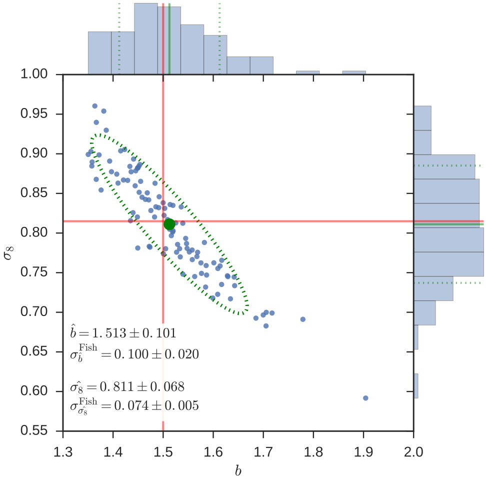
SDSS

Test on lognormal simulations



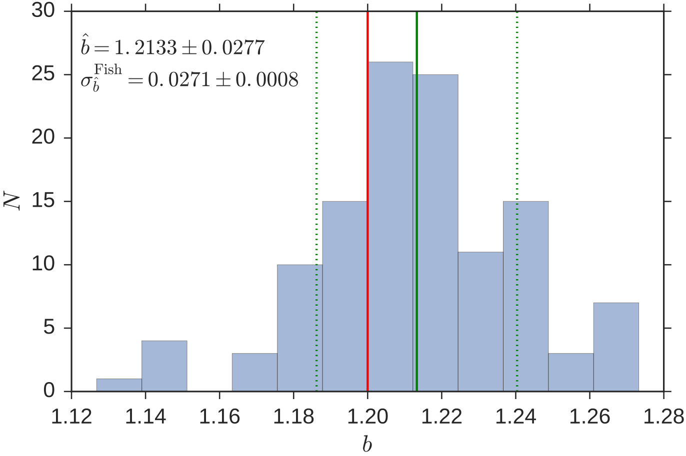
Distribution of the maximum-likelihood estimate of the bias $\hat{b}$. The amplitude of the matter power spectrum is fixed.

Test on lognormal simulations



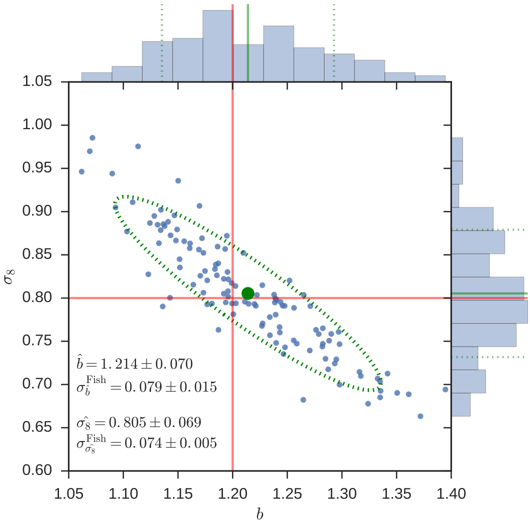
Distribution of the maximum-likelihood estimates of the parameters b, σ_8, for the case in which we allow both parameters to change.

Test on Las Damas simulations



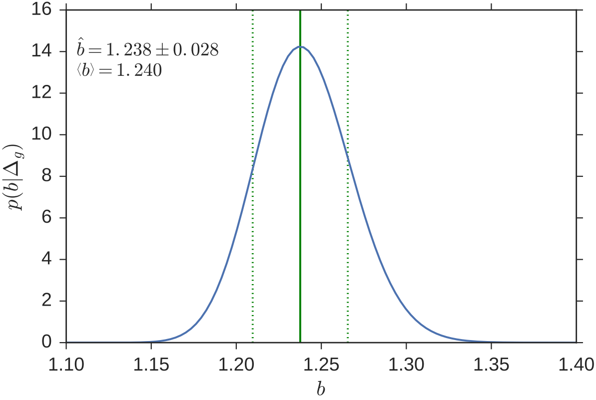
Distribution of the maximum-likelihood estimates of the bias, for the 120 realizations of the Las Damas mocks, for the case in which σ_8 is fixed.

Test on Las Damas simulations



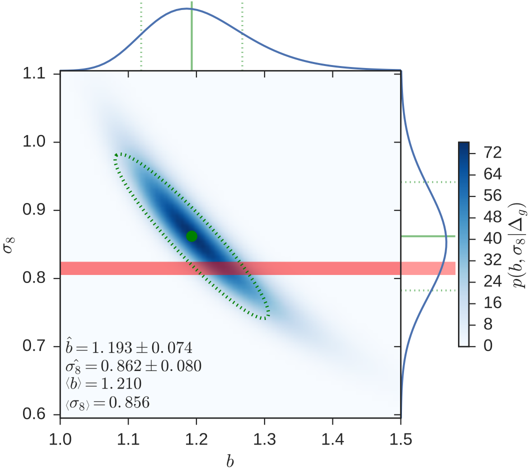
Distribution of the maximum-likelihood joint estimates of b and σ_8.

Results for the SDSS data



Posterior probability distribution of the bias parameter obtained by analysing the SDSS catalogue for the case in which σ_8 is fixed at its fiducial value, $\sigma_8^{\rm fid} = 0.8149$.

Results for the SDSS data



Joint posterior probability distribution of b and σ_8 for the SDSS. Top and right sub-panels are marginalized probability distributions.

Results for the SDSS data. Model selection

- *Null* hypothesis H_0: bias is fixed at $b \equiv 1$, only σ_8 is allowed to vary
- *Alternative* hypothesis H_1: Both bias b and σ_8 are allowed to vary.

→ AIC: *Akaike information criterion*

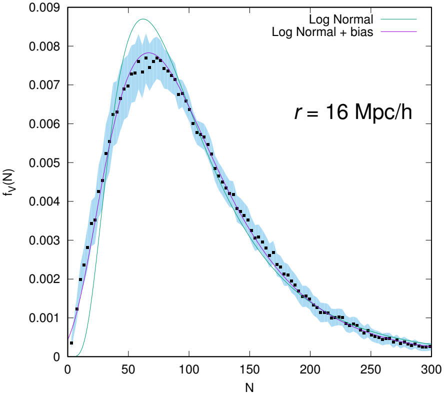
→ BIC: *Bayesian information criterion*

→ GLRT: *Generalized likelihood ratio test*

Criterion	
AIC(H_0) − **AIC**(H_1)	12.6
BIC(H_0) − **BIC**(H_1)	8.22
GLRT: $\log\left(\frac{L_1}{L_0}\right)$ at $\nu = 1$	7.29

The different criteria provide either 'substantial' or 'strong' evidence in favour of the biased model (H_1).

Results for the SDSS data. Counts-in-cells



For the counts-in-cells distribution function $f_V(N)$ (1-pdf) [See Hurtado-Gil et al. (in prep.)]

$$f(\Delta) = \frac{1}{\sqrt{2\pi H_0}} \frac{\exp\left(-\frac{1}{2}\frac{y^2}{H_0}\right)}{\Delta + b - 1} \quad \text{where} \quad H_0 = \log(1 + C)$$

Conclusions

- We have presented a full description of the N-probability density function of the galaxy number density fluctuations in terms of the galaxy bias and σ_8.
- Parameter estimation can be performed via the maximum-likelihood, even more, Bayesian inference can be also performed providing a *prior*.
- The methodology has been tested with both ideal log-normal realizations and mocks derived from the Las Damas project, showing that the maximum-likelihood estimates are unbiased.
- We have applied our formalism to the SDSS main sample. For a volume-limited subset with magnitude $M_r < -20$. We obtain $\hat{b} = 1.193 \pm 0.074$ and $\hat{\sigma_8} = 0.862 \pm 0.080$.
- We show that the *alternative* hypothesis (H_1 of a galaxy bias parameter b given by the maximum-likelihood estimator) is favoured with respect to a no-biasing scenario given by the *null* hypothesis H_0 of $b \equiv 1$.

Arnalte-Mur, P., Martínez, V. J., Norberg, P., et al. 2014, Mon. Not. Roy. Astron. Soc. , 441, 1783

Ata, M., Kitaura, F.-S., & Müller, V. 2015, Mon. Not. Roy. Astron. Soc. , 446, 4250

Blanton, M. R., Schlegel, D. J., Strauss, M. A., et al. 2005, AJ, 129, 2562

Coles, P. & Jones, B. 1991, Mon. Not. Roy. Astron. Soc. , 248, 1

Colless, M., Dalton, G., Maddox, S., et al. 2001, Mon. Not. Roy. Astron. Soc. , 328, 1039

Croft, R. A. C., Dalton, G. B., & Efstathiou, G. 1999, Mon. Not. Roy. Astron. Soc. , 305, 547

Di Porto, C., Branchini, E., Bel, J., et al. 2014, ArXiv e-prints

Fry, J. N. & Gaztanaga, E. 1993, Astrophys. J. , 413, 447

Gaztanaga, E. & Frieman, J. A. 1994, Astrophys. J. Let. , 437, L13

Granett, B. R., Branchini, E., Guzzo, L., et al. 2015, Astron. Astrophys. , 583, A61

Guo, H. & Jing, Y. P. 2009, Astrophys. J. , 702, 425

Jasche, J., Leclercq, F., & Wandelt, B. D. 2015, JCAP, 1, 36

Jasche, J. & Wandelt, B. D. 2013, Astrophys. J. , 779, 15

Kayo, I., Taruya, A., & Suto, Y. 2001, Astrophys. J. , 561, 22

Kovač, K., Porciani, C., Lilly, S. J., et al. 2011, Astrophys. J. , 731, 102

Marinoni, C., Le Fèvre, O., Meneux, B., et al. 2005, Astron. Astrophys. , 442, 801

McBride, C., Berlind, A., Scoccimarro, R., et al. 2009, in Bulletin of the American Astronomical Society, Vol. 41, American Astronomical Society Meeting Abstracts #213, 253

McBride, C. K., Connolly, A. J., Gardner, J. P., et al. 2011, Astrophys. J. , 739, 85

Norberg, P., Baugh, C. M., Hawkins, E., et al. 2002, Mon. Not. Roy. Astron. Soc. , 332, 827

Pollack, J. E., Smith, R. E., & Porciani, C. 2014, Mon. Not. Roy. Astron. Soc. , 440, 555

Sigad, Y., Branchini, E., & Dekel, A. 2000, Astrophys. J. , 540, 62

Swanson, M. E. C., Tegmark, M., Blanton, M., & Zehavi, I. 2008, Mon. Not. Roy. Astron. Soc. , 385, 1635

Szapudi, I. 1998, Mon. Not. Roy. Astron. Soc. , 300, L35

Verde, L., Heavens, A. F., Percival, W. J., et al. 2002, Mon. Not. Roy. Astron. Soc. , 335, 432

Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, Astrophys. J. , 736, 59

Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2005, Astrophys. J. , 630, 1