Large-scale Structure Systematics

Michael D. Schneider

March 26, 2014

LLNL-PRES-XXXXXX

Lawrence Livermore National Laboratory

DARK ENERGY SURVEY

LSST Large Synoptic Survey Telescope

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

The new discovery space for stage III/IV dark energy surveys

Larger surveyed volume $\Rightarrow$ new physics?

THE NEW DISCOVERY SPACE FOR STAGE III/IV DARK ENERGY SURVEYS

Larger surveyed volume $\Rightarrow$ new physics?

- Clustering dark energy
- Early dark energy
- Primordial non-Gaussianity

THE NEW DISCOVERY SPACE FOR STAGE III/IV DARK ENERGY SURVEYS

Larger surveyed volume ⇒ new physics?

- Clustering dark energy
- Early dark energy
- Primordial non-Gaussianity

Distinguish dark energy from modified gravity?

The new discovery space for Stage III/IV dark energy surveys

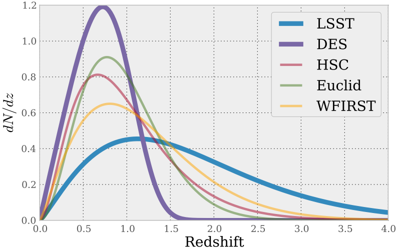
Larger surveyed volume $\Rightarrow$ new physics?

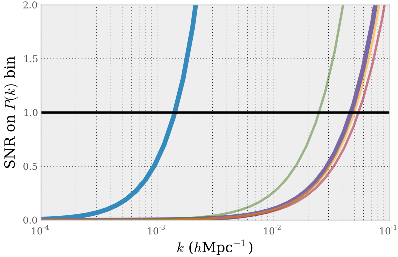
- Clustering dark energy
- Early dark energy
- Primordial non-Gaussianity

Distinguish dark energy from modified gravity?

- Need more information than just BAOs.
- Confidence in BAO systematics implies confidence in broad-band power, photo-z's, & gravitational growth via cross-correlations.

The new discovery space for stage III/IV dark energy surveys

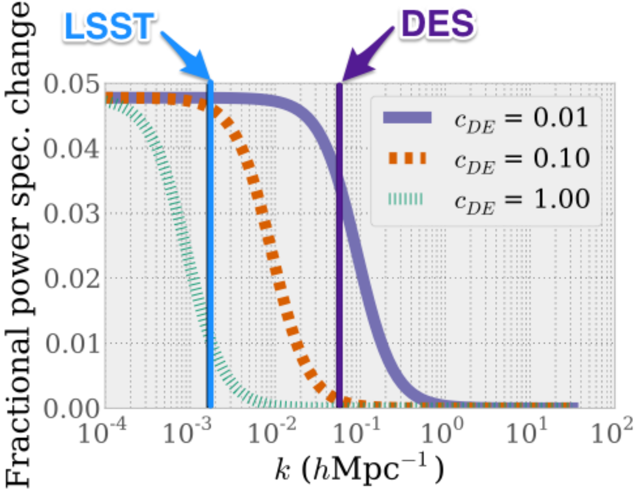




	DES	LSST	HSC	Euclid	WFIRST
Area (sq. deg.)	5,000	20,000	1,500	15,000	2,000
min k (hMpc^{-1})	0.049	0.0017	0.060	0.028	0.049

Clustering DE increases the large-scale matter power spectrum

Transition scale set by DE Jeans length ↔ effective sound speed c_{DE}

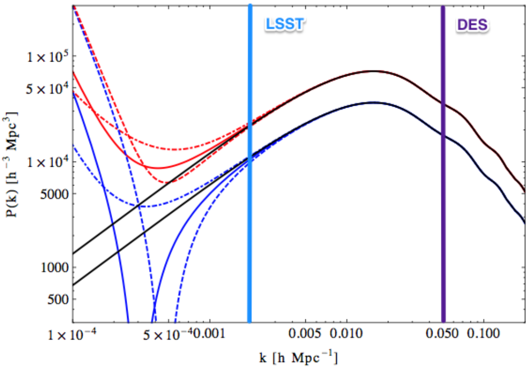


Primordial non-Gaussianity also increases large-scale power of biased matter tracers

From 'standard' inflation, $f_{\rm NL}^{\rm loc.} \lesssim 1$.

- line-of-sight
- transverse
- $f_{\rm NL}^{\rm loc.} = \pm 0.5$ (dot-dashed / dashed)

Baldauf et al. (2011)



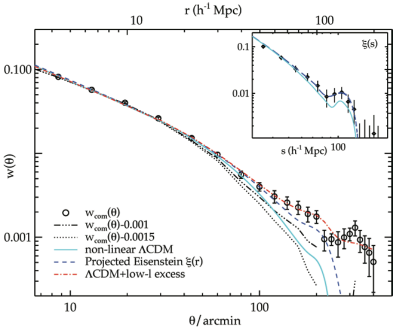
Difficult measurement $\rightarrow$ LSST may bound interesting range.

Given photo-z errors, the angular (cross-)correlation functions are a primary probe of LSS.

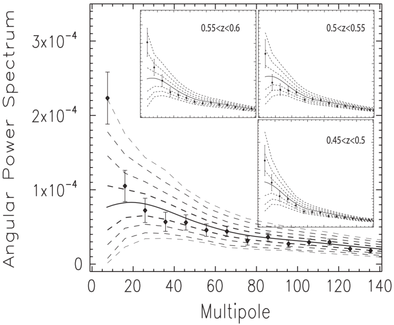
$$w(\theta) \equiv \left\langle \delta_g(\vec{\theta} + \vec{\theta}')\delta_g(\vec{\theta}') \right\rangle \approx \frac{DD(\theta, \theta + \Delta\theta)}{RR(\theta, \theta + \Delta\theta)} \tag{1}$$

Excess clustering detections

Sawangwit et al. (2011)



Thomas et al. (2010)

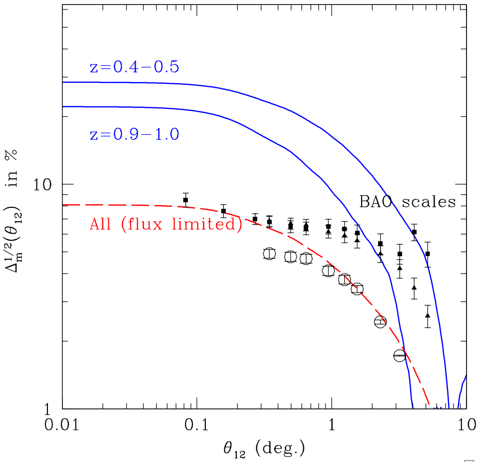


LRG, SDSS DR5, photo-*z*

700k LRG, SDSS DR7, photo-*z*, 4 bins

DES calibration of large-scale systematic clustering

Credit: DES-LSS WG, incl. Anne Bauer, Eli Rykoff, Eduardo Rozo, Nacho Sevilla, Aurelieb Benoit-Levy



Classes of systematics

Courtesy Ofer Lahav

- Instrumental
  - Photo-z
  - Star / galaxy separation
  - Sky background, seeing
  - Dithering / footprint
  - Image quality (photometry)
- Astrophysical
  - Galaxy biasing
  - Lensing magnification
  - Redshift-space distortions (RSD)
  - Galactic extinction
- Theoretical
  - Cosmological model & parameters
  - Covariances / nonlinear mode coupling
- Analysis methods (Percival: "misused statistics")
  - Look-elsewhere effect

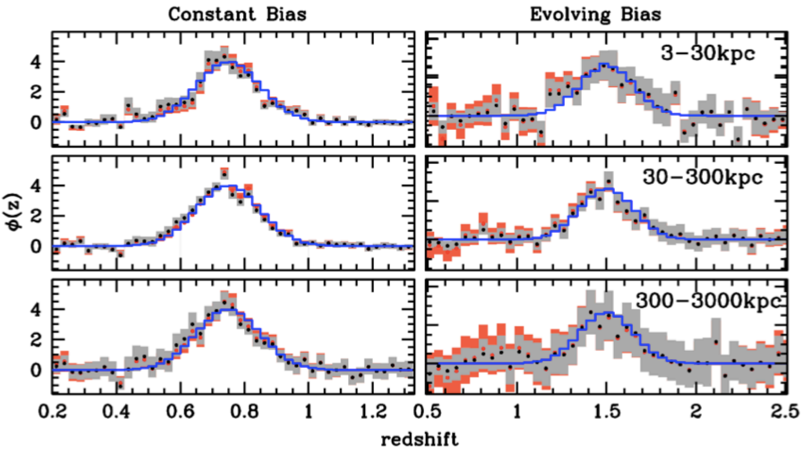
Photometric redshifts

Photo-z errors cause:

- Finite resolution along line-of sight
  - No radial BAOs
  - Redshift-space distortions systematic rather than signal
  - Limited modes for z-dependent dark energy
- Systematic errors in cosmology (when photo-z error distribution unknown)
  - Confusion of signals (clustering / lensing)
  - Limits on self-calibration through cross-correlations / joint probes
  - Skewed BAO peak (Zhan & Knox 2006, Simpson+2009) $\rightarrow \sigma_z \lesssim 10^{-3}$

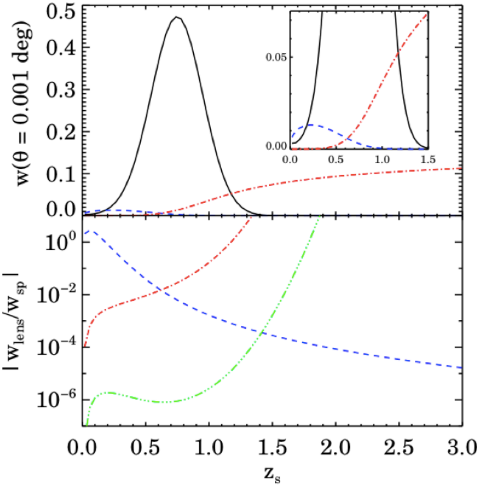
PHOTO-z CALIBRATION FROM ANGULAR CROSS-CORRELATIONS

Schmidt et al. (2013), Menard et al. (2013)



LENSING MAGNIFICATION IS A CONTAMINANT FOR PHOTO-z CALIBRATIONS

Matthews & Newman (DESC meeting Dec. 2013)



Star / galaxy separation

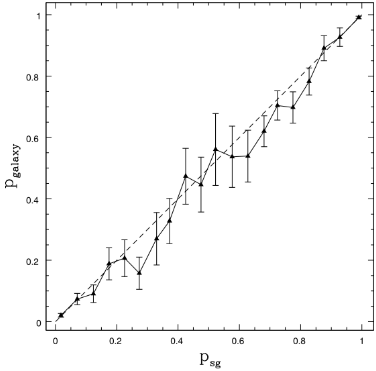
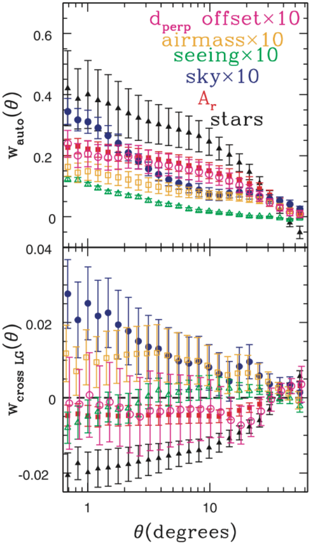


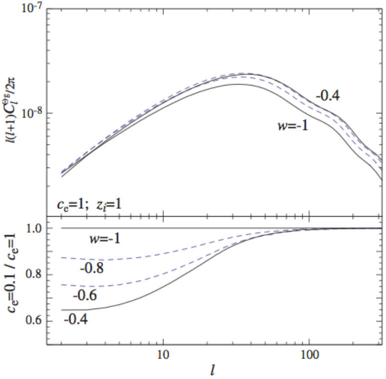
Figure 2. The fraction of objects that are galaxies versus the value of the star/galaxy parameter (p_{sg}). The dashed line displays the relationship $p_{galaxy} = p_{sg}$. Errors are Poissonian.

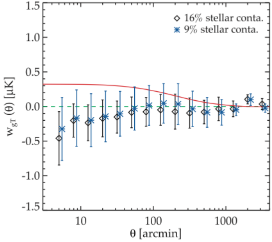


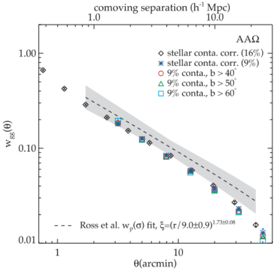
CMB / galaxy cross-correlations

Sawangwit et al. (2009)

Hu & Scranton (2004)







"…it is not only the accuracy of the z_{phot} that is important, but also the probability that an object is a galaxy."

– Ross et al. (2011)
Ameiliorating systematic uncertainties in the angular clustering of galaxies: a study using the SDSS-III

"…it is not only the accuracy of the z_{phot} that is important, but also the probability that an object is a galaxy."

"…identifying robust methods of assigning probabilities that objects are galaxies should be a major focus of forthcoming photometric redshift surveys."

– Ross et al. (2011)

Ameiliorating systematic uncertainties in the angular clustering of galaxies: a study using the SDSS-III

Methods for systematics correction

Ross et al. (2011)

1. Mask (e.g. extinction, regions around bright stars)
2. Apply weights to galaxies – What about correlated systematics?
3. Calibrate with cross-correlations

Methods for systematics correction

Ross et al. (2011)

1. Mask (e.g. extinction, regions around bright stars)
2. Apply weights to galaxies – What about correlated systematics?
3. Calibrate with cross-correlations

Correcting for known systematics

Agarwal et al. (2013),

$$\delta_{g,\text{obs}}(\vec{\theta}) = \delta_{g,\text{true}}(\theta) + \sum_{i=1}^{N_{\text{sys}}} \epsilon_i(\theta)\delta_i(\vec{\theta}) + u(\vec{\theta}) \quad (2)$$

- $\delta_i(\vec{\theta})$: known systematic i angular overdensity
- $\epsilon_i(\theta)$: amplitude of systematic i
- $u(\vec{\theta})$: unknown systematics

Correcting for known systematics

Agarwal et al. (2013),

$$\delta_{g,\mathrm{obs}}(\vec{\theta}) = \delta_{g,\mathrm{true}}(\theta) + \sum_{i=1}^{N_{\mathrm{sys}}} \epsilon_i(\theta)\delta_i(\vec{\theta}) \tag{2}$$

- $\delta_i(\vec{\theta})$: known systematic i angular overdensity
- $\epsilon_i(\theta)$: amplitude of systematic i

Use all tomographic cross-correlations to solve for $\epsilon_i(\theta)$.

Correcting for known systematics

Agarwal et al. (2013),

$$\delta_{g,\mathrm{obs}}(\vec{\theta}) = \delta_{g,\mathrm{true}}(\theta) + \sum_{i=1}^{N_{\mathrm{sys}}} \epsilon_i(\theta)\delta_i(\vec{\theta}) \tag{2}$$

- $\delta_i(\vec{\theta})$: known systematic i angular overdensity
- $\epsilon_i(\theta)$: amplitude of systematic i

Use all tomographic cross-correlations to solve for $\epsilon_i(\theta)$.
Problem: no information left to correct for *unknown* sytematics.

To confidently remove unknown systematics we have to
throw away data.
What's the best (i.e. least damaging) way to do this?

To confidently remove unknown systematics we have to *throw away data*.
What's the best (i.e. least damaging) way to do this?

Insight from a different problem: astrophysical systematics in cross-correlations.

To confidently remove unknown systematics we have to *throw away data*.
What's the best (i.e. least damaging) way to do this?

Insight from a different problem: astrophysical systematics in cross-correlations.

- The observed δ_g is a sum of the intrinsic clustering of galaxies, distortions due to lensing magnification, and shot noise,

$$n(\mathbf{x}) = \bar{n} + \delta n_g(\mathbf{x}) + \delta n_\mu(\mathbf{x}) + \epsilon \tag{3}$$

To confidently remove unknown systematics we have to *throw away data*.
What's the best (i.e. least damaging) way to do this?

Insight from a different problem: astrophysical systematics in cross-correlations.

- The observed δ_g is a sum of the intrinsic clustering of galaxies, distortions due to lensing magnification, and shot noise,

$$n(\mathbf{x}) = \bar{n} + \delta n_g(\mathbf{x}) + \delta n_\mu(\mathbf{x}) + \epsilon \tag{3}$$

- The galaxy angular (cross-)correlation functions in tomographic bins have in principle the following contributions:

$$w(\theta) = w_{gg}(\theta) + w_{g\mu}(\theta) + w_{\mu g}(\theta) + w_{\mu\mu}(\theta) + w_{\mathrm{SN}} \tag{4}$$

To confidently remove unknown systematics we have to *throw away data*.
What's the best (i.e. least damaging) way to do this?

Insight from a different problem: astrophysical systematics in cross-correlations.

- The observed δ_g is a sum of the intrinsic clustering of galaxies, distortions due to lensing magnification, and shot noise,

$$n(\mathbf{x}) = \bar{n} + \delta n_g(\mathbf{x}) + \delta n_\mu(\mathbf{x}) + \epsilon \quad (3)$$

- The galaxy angular (cross-)correlation functions in tomographic bins have in principle the following contributions:

$$w(\theta) = w_{gg}(\theta) + w_{g\mu}(\theta) + w_{\mu g}(\theta) + w_{\mu\mu}(\theta) + w_{\mathrm{SN}} \quad (4)$$

- Hard to separate terms with photo-z errors.

Optimal cross-correlations

Treat the Limber equation,

$$w(\theta) = \int dz\, W_A(z) \left[W_B(z) \xi(r(\theta, z)) \right] \tag{5}$$

as an integral equation for the foreground window $W_A(z)$.
Solve for the optimal $W_A(z)$ that:

1. Maximizes the signal-to-noise ratio of $w(\theta)$
2. Isolates separate physical contributions to $w(\theta)$[1] (e.g. clustering vs. lensing)

Improve self-calibration of systematics → particularly photo-z errors.

Schneider (2014) arXiv: 1401.0537

[1] Similar to Joachimi & Schneider (2010) "Intrinsic alignment boosting"

Optimal redshift weighting

The Limber equation is a
Fredholm integral equation of the 1st kind,

$$w_{XY}(\theta) = \int_0^{\chi_\infty} d\chi\, W_X(\chi) K(\chi, \theta), \tag{6}$$

where $K(\chi, \theta) \equiv W_Y(\chi)\xi(\chi\theta)$ and $W_X(\chi)$ is to be optimized.

Use eigenvectors of K to describe the optimal solution space.
But, $K(\chi, \theta)$ is not Hermitian ($K(\chi, \theta) \neq K(\theta, \chi)$).

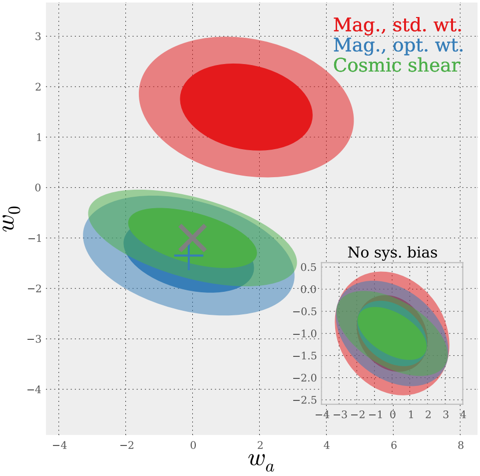
A better kernel is,

$$C(\chi, \chi') \equiv \int d\theta K(\chi, \theta)\, K(\chi', \theta), \tag{7}$$

$$\int d\chi\, C(\chi, \chi')\, \psi(\chi) = \lambda \psi(\chi'). \tag{8}$$

Lensing magnification vs galaxy bias systematic error

DE Figure of Merit from magnification ≈ 80% that from shear.



Lesson:
The eigenvectors of the integration kernel provide a natural basis for optimized correlation function estimators.

- Generally, can expand the (unknown) systematic error in an orthogonal basis,

$$u^{\alpha}(\theta) = \sum_i u_i^{\alpha} \Phi_i(\theta) \tag{9}$$

- The clustering (cross-)correlation is similarly expanded as,

$$w_{\text{true}}(\theta) = \sum_i w_i \eta_i(\theta) \tag{10}$$

- The best choice for systematics modeling is $\Phi_i = \eta_i$.
    - Don't care about modes orthogonal to $\eta_i(\theta)$
    - Prior information: select subset of $\{\eta_i\}$ or bound the volume on $\{u_i\}$.

To derive η in $w_{\text{true}}(\theta) = \sum_i w_i \eta_i(\theta)$, use the auxiliary symmetric kernel,

$$C_{\text{aux}}(\theta, \theta') \equiv \int d\chi\, K(\chi, \theta)\, K(\chi, \theta'),$$

$$\int d\chi\, C_{\text{aux}}(\theta, \theta')\, \eta(\theta) = \lambda \eta(\theta'), \tag{11}$$

the original source kernel can be reconstructed,

$$K(\chi, \theta) = \sum_{i=1}^{\infty} \sqrt{\lambda_i} \psi_i(\chi) \eta_i(\theta). \tag{12}$$

(Remember $w(\theta) = \int d\chi\, W(\chi) K(\chi, \theta)$)

A filter for angle-dependent systematics in the correlation function

Define,

$$\hat{w}(\theta) = \int d\phi \, w_{\text{obs}}(\phi) f(\theta, \phi). \tag{13}$$

Can we find an f such that $\hat{w}(\theta) \approx w_{\text{true}}(\theta)$?

A FILTER FOR ANGLE-DEPENDENT SYSTEMATICS IN THE CORRELATION FUNCTION

Define,

$$\hat{w}(\theta) = \int d\phi \, w_{\text{obs}}(\phi) f(\theta, \phi). \quad (13)$$

Can we find an f such that $\hat{w}(\theta) \approx w_{\text{true}}(\theta)$?

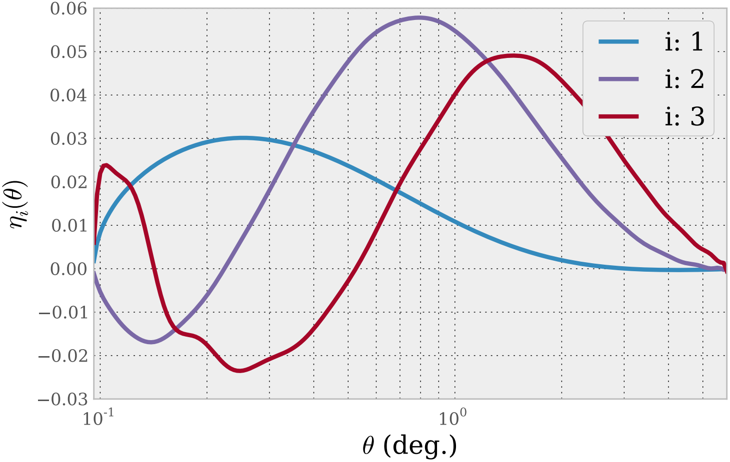
Define $\tilde{\eta}_i$ such that,

$$\int d\theta \, \tilde{\eta}_i(\theta) w_u(\theta) = 0. \quad (14)$$

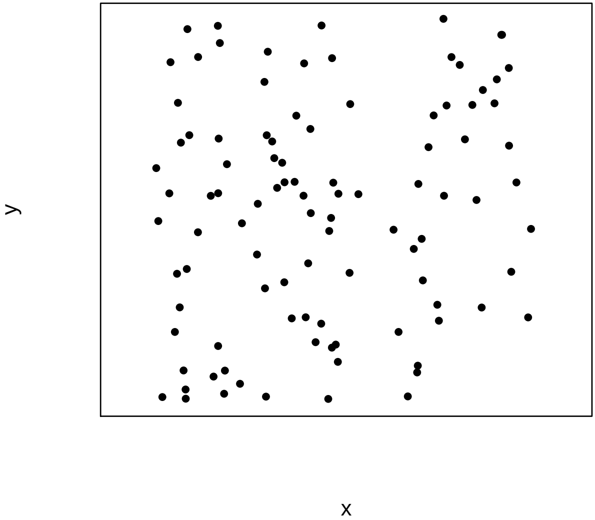
Then let,

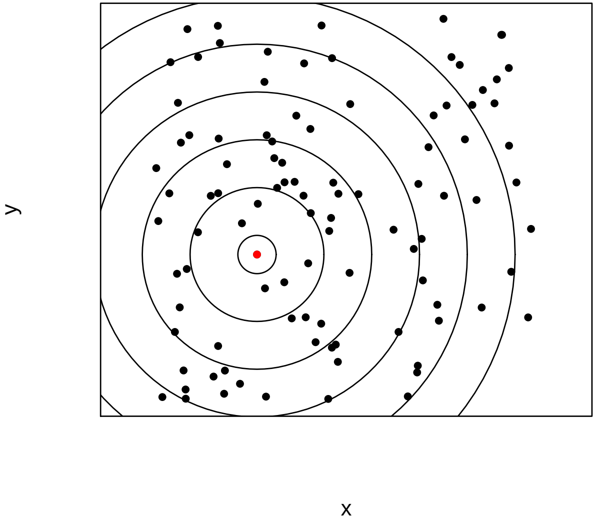
$$f(\theta, \phi) = \sum_{i=1}^{n} \sqrt{\lambda_{f,i}} \, \tilde{\eta}_i(\theta) \tilde{\eta}_i(\phi). \quad (15)$$

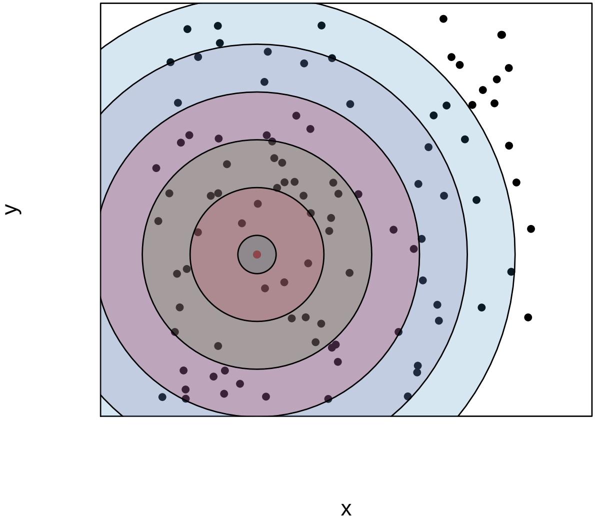
$\Rightarrow f$ **selects components of** $w(\theta)$ **orthogonal to the systematics model.**



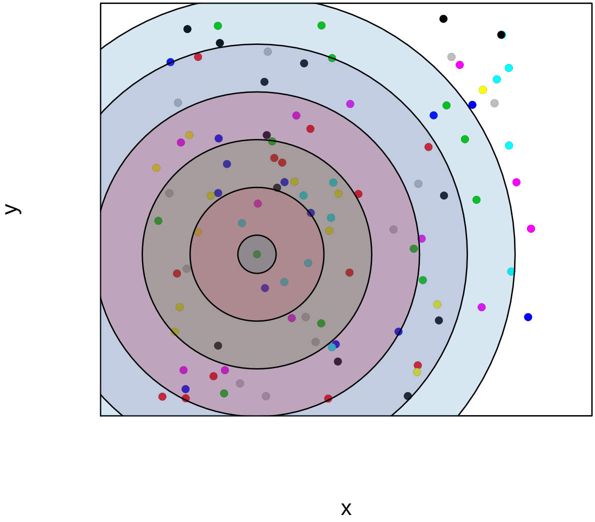
0.06
0.05
0.04
0.03
0.02
0.01
0.00
−0.01
−0.02
−0.03
$\eta_i(\theta)$
10^{-1}
10^{0}
θ (deg.)
i: 1
i: 2
i: 3


y
x


y
x



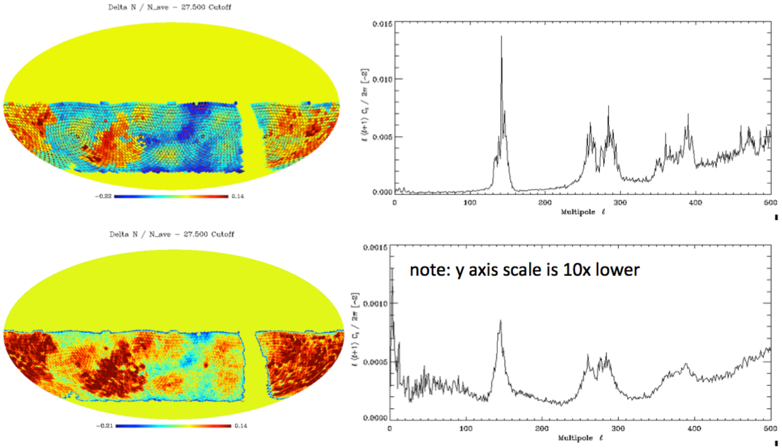
y
x



y
x

Filtering the correlation function has limits.
It's better to avoid imprinting systematic correlations wherever possible.

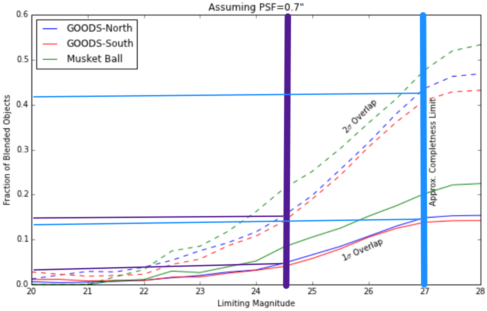
DITHERING / FOOTPRINT



HEALPIX shows the angular power spectrum of δN/<N>. Features at 100<l<300 cause systematic errors in B.A.O. studies of dark energy, **but large dithers reduce these by a factor of 10!!!**

Photometry & cadence

- Same patch of sky should ideally sample different airmass, pointing configurations, seeing.
- *'Catastrophic' blending* of objects could be a problem for photometry in LSST (but probably not DES).



Covariances

Cosmological parameter inference requires specifying the covariance matrix of the angular correlations for evaluating the likelihood function.

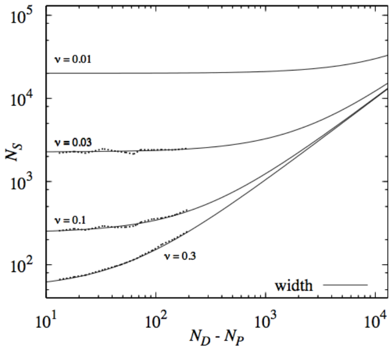
$$\hat{C}(\theta_i, \theta_j) = \frac{1}{N_r} \sum_{k=1}^{N_r} \left(w_k(\theta_i) - \bar{w}(\theta_i)\right) \left(w_k(\theta_j) - \bar{w}(\theta_j)\right) \tag{16}$$

Challenges:

- Require large numbers of realizations, N_r
- Number of θ bins becomes large with tomography / joint probes analyses
- Covariance can be a function of cosmology

Covariance errors propagate to parameter errors

Taylor & Joachimi (2014)

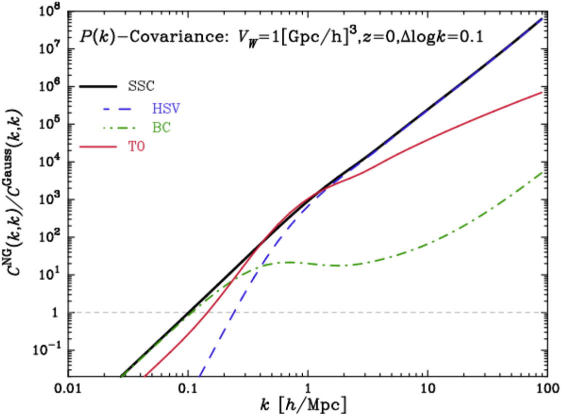


$\nu \equiv$ fractional bias in parameter variance.

Scales with (num. bins) / (num. samples) (Dodelson & Schneider, 2013)

Super-sample covariance

A new (and important) contribution to the sample covariance from long-wavelength density perturbations.
Takada & Hu (2013)



Simulations - novel approaches

- Emulators
  - Heitmann, Habib et al. – sub-percent accuracy in the power spectrum, mass function in highly nonlinear regime.
- Mapping across cosmologies
  - Angulo & White (2010) – Explore cosmological models with a single high-res simulation?
- Mode-resampling
  - Schneider+(2011) – Sample Gaussian statistics of large-scale modes in an N-body simulation in post-processing.
  - Explore models with different long-wavelength perturbations.

SUMMARY

- Many systematics introduce spurious large-angle clustering
- Some of the most exciting discoveries could come from large angles!

Synergies between DES and LSST for LSS:

- DES will teach LSST:
  - Where new algorithms and systematics mitigation approaches are needed.
  - Important systematics we may not have thought of yet.
- LSST can give DES:
  - Alternate data management approaches (e.g. Jim Bosch's talk yesterday on probabilistic PSF estimation)
  - PhoSim – coupling of catalog generation and image systematics

Tools – Some useful codes I've noted

STOMP – "A library for doing astrostatistics on the celestial sphere" https://code.google.com/p/astro-stomp/

CHOMP – "Object oriented, cosmology and halo model theoretical prediction library" https://code.google.com/p/chomp/

GalSim – "The modular galaxy image simulation toolkit" https://github.com/GalSim-developers/GalSim

emcee – "extensible, pure-Python implementation of Goodman & Weare's Affine Invariant Markov chain Monte Carlo (MCMC) Ensemble sampler" http://dan.iel.fm/emcee/current/

Tractor – "astronomical source detection, separation, and photometry" http://thetractor.org/

ALPT – Kitaura / Neyrinck

Agarwal, N., Ho, S., Myers, A. D., Seo, H. J., Ross, A. J., Bahcall, N., Brinkmann, J., Eisenstein, D. J., Muna, D., Palanque-Delabrouille, N., Pâris, I., Petitjean, P., Schneider, D. P., Streblyanska, A., Weaver, B. A., & Yèche, C. 2013

Dodelson, S. & Schneider, M. D. 2013, arXiv, D88, 063537

Menard, B., Scranton, R., Schmidt, S., Morrison, C., Jeong, D., Budavari, T., & Rahman, M. 2013, arXiv.org, 4722

Ross, A. J., Ho, S., Cuesta, A. J., Tojeiro, R., Percival, W. J., Wake, D., Masters, K. L., Nichol, R. C., Myers, A. D., de Simoni, F., Seo, H. J., Hernandez-Monteagudo, C., Crittenden, R., Blanton, M., Brinkmann, J., da Costa, L. A. N., Guo, H., Kazin, E., Maia, M. A. G., Maraston, C., Padmanabhan, N., Prada, F., Ramos, B., Sanchez, A., Schlafly, E. F., Schlegel, D. J., Schneider, D. P., Skibba, R., Thomas, D., Weaver, B. A., White, M., & Zehavi, I. 2011, Monthly Notices of the Royal Astronomical Society, 417, 1350

Sawangwit, U., Shanks, T., Abdalla, F. B., Cannon, R. D., Croom, S. M., Edge, A. C., Ross, N. P., & Wake, D. A. 2011, MNRAS, 416, 3033

Sawangwit, U., Shanks, T., Cannon, R. D., Croom, S. M., Ross, N. P., & Wake, D. A. 2009, arXiv.org, 2228

Schmidt, S. J., Ménard, B., Scranton, R., Morrison, C., & McBride, C. K. 2013, Monthly Notices of the Royal Astronomical Society

Takada, M. & Hu, W. 2013, arXiv.org

Taylor, A. & Joachimi, B. 2014

Thomas, S. A., Abdalla, F. B., & Lahav, O. 2010, arXiv.org, 241301