Comprehensive Forecasts for WL Surveys

G. Bernstein (U Penn)

23 August 2007

Motivation & Problems

- When you're asking for several $100M, you'd better understand how well your measurement will work! *(including checking your calculations)*
- Many idealizations in most forecasts, which may lead to substantial overestimates of capabilities:
  - There are { no intrinsic alignments | only II | intrinsic alignments have simple functional form}
  - The Universe is described by GR with an isotropic DE with w0/wa eqn of state.
  - Galaxy densities are linearly biased w.r.t. mass with pure Poisson noise (e.g. fully correlated)
  - Photo-z errors {don't exist | are just biases | have Gaussian distributions}.
  - All of the galaxies (or all in a bin) have the same shear calibration errors, same intrinsic correlations, same bias & correlation coefficients w.r.t. mass and to each other.

Challenges

- *The "dark energy," modifications to GR, and many systematic effects are free functions of redshift, scale, galaxy type, etc. Assuming that they have particular functional forms can make them artificially easy to distinguish.*
- Problems that are easily solved in isolation may be difficult to solve in combination. Example:
  - "Well known" that power-spectrum tomography only requires photo-z's accurate enough to generate 3 or so bins (e.g. Hu).
  - "Well known" that intrinsic alignments (at least II) are easily removed with use of photo-z's
  - But Bridle & King note that tomography with simultaneous intrinsic-alignment rejection places much stronger demands on photo-z.

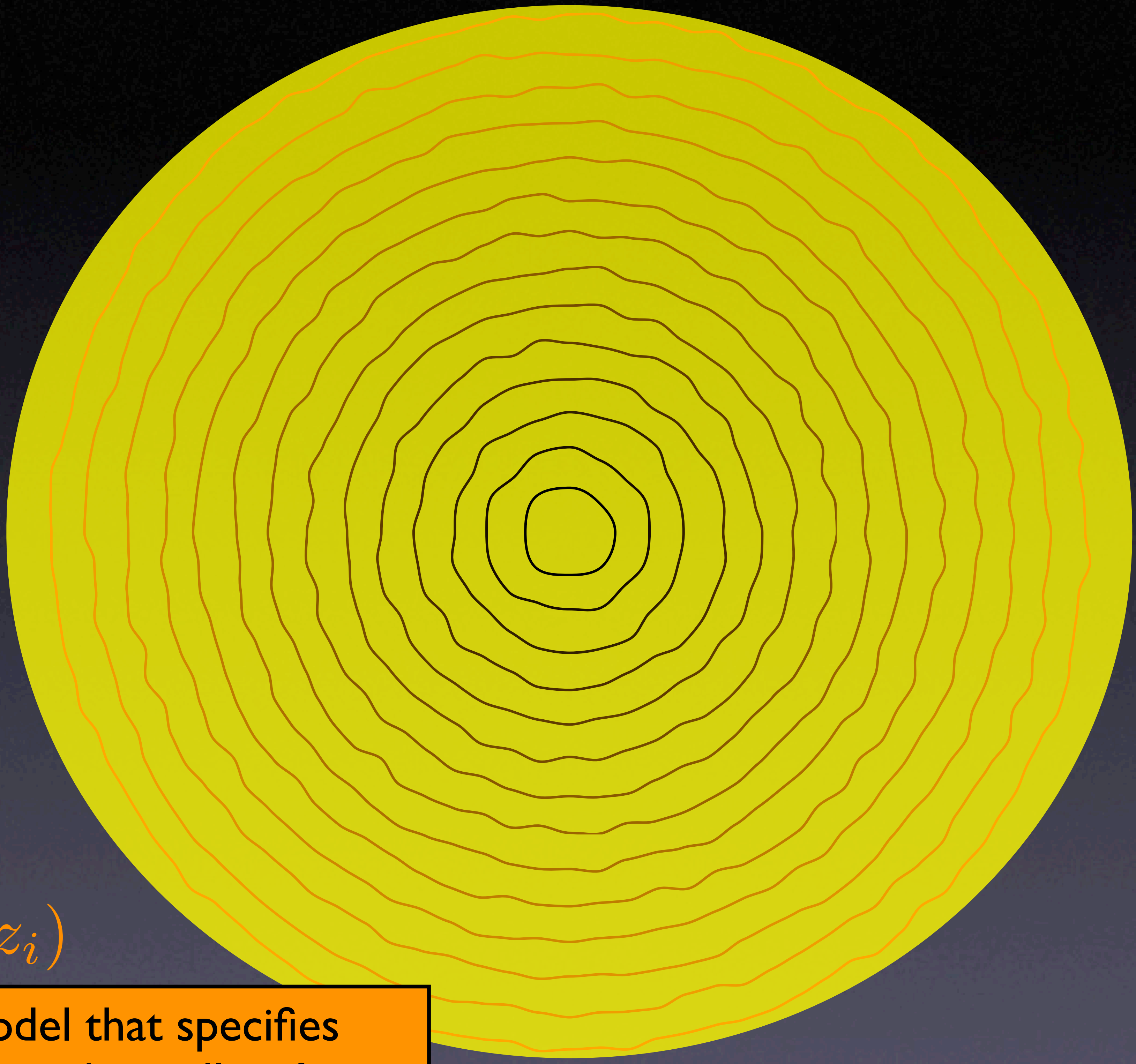
Desire a WL forecast that includes all important systematics. Also treat dark energy, gravity alterations, and systematics as generally or non-parametrically as possible.

Benefits

- Most forecasts make use only of power-spectrum tomography, to some predetermined maximal l. Much more info is available from WL survey:
  - High-l lensing carries information, even though must marginalize over theoretical power-spectrum uncertainties.
  - Galaxy-galaxy, galaxy-shear, and shear-shear 2-pt info can work very well together (Hu & Jain, Zhan)
  - 3-point shear info (even ignoring SGG, SSG, GGG signals, plus III, GII, GGI contaminants....) (Takada & Jain)
  - Peak statistics (clusters) (Hennawi & Spergel, Marion & Bernstein, Wang et al, ...)
- Using multiple statistics can isolate systematics and reduce their detriment. Example: galaxy-shear correlations reduce impact of intrinsic correlations?

Generalized WL Analysis

- The Universe has FRW metric with some curvature and mass densities ω_k, ω_m
- Galaxies are assumed to line on a series of shells at nominal redshifts z_i or a_i
- Each shell has
  - Ang-diam distance D_i
  - Comoving thickness $\Delta\chi_i$
  - A mass distribution m_{ilm}
  - which is produced from some power spectrum: $P_m(l/D_i, z_i)$



We may have some cosmological model that specifies distances, power, etc, but for now leave them all as free parameters and see what combinations are constrained by observations.

Generalized WL Analysis (2)

- If light travels on geodesics of perturbed RW metric:

$$ds^2 = (1+2\psi)dt^2 - a^2(t)(1+2\phi)d\mathbf{x}^2$$

- *and* the matter field sources all the potential fluctuations
- *and* GR is correct about Poisson equation with $\phi = -\psi$
- then gravitational lensing induces a convergence (and matching E-mode shear) on shell *i* of

$$\kappa_i^{\mathrm{lens}} = \sum_{j<i} A_{ij} \frac{3\omega_m D_j \, \Delta\chi_j}{2a_j} m_j$$

- with

$$A_{ij} \equiv \begin{cases} \frac{D_{ij}}{D_i} \approx (1 - D_j/D_i)(1 - \omega_k D_i D_j/2) & i > j, \\ 0 & i \leq j. \end{cases}$$

- We may add additional terms to Aij to reflect non-GR Poisson eqns or clustered dark energy.

Robustness of lensing cosmology

Change in expansion history at some z alters

Change in growth or bias at some z affects

Intrinsic alignments alter these

Error in shear calibration at some z affects

Curvature affects

$$\begin{pmatrix} \text{shear}_1 \\ \text{shear}_2 \\ \text{shear}_3 \\ \vdots \\ \text{shear}_n \end{pmatrix} = \begin{pmatrix} 0 & A_{12} & A_{13} & \cdots & A_{1n} \\ 0 & 0 & A_{23} & & A_{2n} \\ 0 & 0 & 0 & & A_{3n} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \times \begin{pmatrix} \text{mass}_1 \\ \text{mass}_2 \\ \text{mass}_3 \\ \vdots \\ \text{mass}_n \end{pmatrix}$$

$$A_{ij} = (1 - D_i/D_j)\,(1 - \omega_k D_i D_j/2) \qquad (i < j)$$

Generalized WL Analysis: Observables

- We use photo-z or other information to divide the observed galaxies (or 21-cm data, other observables) into *sets*. For each set we observe
  - A ***density*** field $g_{\alpha lm}$
  - and/or a ***lensing*** field $\kappa_{\alpha lm}$
- Each galaxy in set α has probability $p_{\alpha i}$ of being on redshift shell i.
- At each spherical harmonic, the observable quantities are

$$g_\alpha = \sum_i p_{\alpha i} g_{\alpha i}$$

$$\kappa_\alpha = \sum_i p_{\alpha i} \kappa_{\alpha i} = \sum_i p_{\alpha i} \left[(1 + f_{\alpha i}) \sum_{j<i} A_{ij} \frac{3\omega_m D_j \Delta\chi_j}{2a_j} m_j + \kappa_{\alpha i}^{\mathrm{int}} \right].$$

- where each *subset* of the sources has its own density, shear cal error, and intrinsic-alignment "convergence" signal.

Likelihood for density/shear

- On each redshift shell i there are variables $\{m_i, g_{\alpha i}, \kappa_{\alpha i}^{\rm int}\}$
- Assume Gaussian and Limber, *i.e.:*
  - There is no correlation between distinct redshift shells or spherical harmonics
  - Within a shell, the above variable have multivariate Gaussian distribution.
- Then the likelihood is fully described by the covariances which are power spectra:

$$\langle m_i^2 \rangle \equiv P_{im}$$
$$\langle g_{i\alpha}^2 \rangle \equiv P_{i\alpha\alpha}^{gg} = (b_{\alpha i}^g)^2 P_i^m + \frac{1}{n_{\alpha i}}$$
$$\langle (\kappa_{i\alpha}^{\rm int})^2 \rangle \equiv P_{i\alpha\alpha}^{\kappa\kappa} = (b_{\alpha i}^\kappa)^2 P_i^m + \left(\sigma_\gamma^2/\rho\right)_{\alpha i}$$
$$\langle m_i g_{i\alpha} \rangle \equiv P_{i\alpha}^{mg} = b_{\alpha i}^g r_{\alpha i}^g P_i^m$$
$$\langle m_i \kappa_{i\alpha} \rangle \equiv P_{i\alpha}^{m\kappa} = b_{\alpha i}^\kappa r_{\alpha i}^\kappa P_i^m$$

- Note appearance of II and GI terms in analogy to bias and correlation coefficients of galaxy clustering.
- Since the observables are linear functions of these variables, they also have a multivariate, zero-mean Gaussian likelihood. Usual Fisher matrix and data-fitting techniques will apply.

Parameter Hell

- A beautiful non-parametric likelihood expression for the likelihood of 2-point survey statistics. But aside from the desirable cosmological parameters (curvature, distances) we have a vast number of nuisance parameters:
  - Probability *p* for *every subset* of galaxies (*i.e.* photo-z error distributions)
  - bias and correlation coefficient of galaxy density of *every subset* of galaxies which is also function of k.
  - bias and correlation coefficient of intrinsic "lensing" signal of *every subset* of galaxies which is also function of k.
  - even worse: there are huge number of *cross-correlation* terms

$$\langle g_{i\alpha} g_{i\beta} \rangle \equiv P^{gg}_{i\alpha\beta} = b^g_{\alpha i} b^g_{\beta i} r^{gg}_{i\alpha\beta} P^m_i$$
$$\langle g_{i\alpha} \kappa_{i\beta} \rangle \equiv P^{g\kappa}_{i\alpha\beta} = b^g_{\alpha i} b^\kappa_{\beta i} r^{g\kappa}_{i\alpha\beta} P^m_i$$
$$\langle \kappa_{i\alpha} \kappa_{i\beta} \rangle \equiv P^{\kappa\kappa}_{i\alpha\beta} = b^\kappa_{\alpha i} b^\kappa_{\beta i} r^{\kappa\kappa}_{i\alpha\beta} P^m_i$$

  - Structure-formation models will never be able to tell us what all these should be; neither will we be able to conduct an analysis in which they are all free, so some model with limited number of parameters will be necessary. But it shows that present modeling is grossly simplified even in Gaussian approximation.

Prior Salvation

- N-body modeling will provide prior on the mass power (exact prior at low k, weaker at high k). Will depend on:
  - Primordial spectral parameters $\{A_s, n_s\}$
  - Transfer function parameters $\{\omega_m, \omega_b, \omega_\nu\}$
  - Growth function at each redshift shell $\{G_i\}$
  - maybe some more, for modified gravity.
- Physical model for galaxy biases/covariances may be possible: halo model in Hu & Jain, but substantial additional "adjustment" parameters likely.
- Intrinsic alignment models of useful precision are not likely, so we'll need generic functional models (polynomials, etc.).
- Note that the observational Fisher matrix already includes all "galaxy-galaxy" lensing and "cross-correlation" information.
- Also if there are spectroscopic survey samples, they become "sets" and their cross-correlation with source galaxies incorporates Newman's technique for constraining source redshifts.

Prior from Redshift Survey

- An *unbiased* and *complete* redshift survey of N_α galaxies in set α yields observables $N_{\alpha i}$ counts of galaxies in each subset. The likelihood is

$$L(\{N_{\alpha i}\}|\{p_{\alpha i}\}) = \prod_{i,\alpha} p_{\alpha i}^{N_{\alpha i}}$$

- and the corresponding Fisher matrix over the probabilities is

$$F_{\alpha i,\beta j} = \left\langle \frac{-\partial^2 \log L}{\partial p_{\alpha i} \partial p_{\beta j}} \right\rangle = \delta_{\alpha\beta}\delta_{ij}\frac{N_\alpha}{p_{\alpha i}}$$

- Any stronger prior constraints on the $p_{\alpha i}$ would have to come from assuming that spectra of un-surveyed galaxy sets are exactly the same as those in survey sets. How will we be able to quantify our trust in this?

Three-point information

- The expected bispectrum can also be expressed in terms of the same dark-energy-agnostic parameters as the general 2-point function; in approximation of Scoccimarro & Couchman, 3-pt is a function of the linear & non-linear power spectra.
- Implement as in Takada & Jain, gaussian error model for the bispectrum components. Ignore intrinsic 3-pt correlations, etc., for sanity
- Takada & Jain (in prep) show that Gaussian-uncertainty approximation isn't too bad. Just add a new 3-pt Fisher matrix to the 2-pt matrix.
- More complex than PS, likelihood does not separate into ilm terms; must limit to <~5 redshift bins for feasible computation.

Peak-Counting Information

- Presume that peaks in convergence distribution are located using a multiscale filter matched to angular and redshift dependence of clusters (as per Marion & Bernstein). Observables are peak counts vs convergence strength and redshift above safe detection threshold.
- Sheth-Tormen formula gives convergence peak counts as simple function of linear power spectrum plus growth and distances to redshift shells - again dark-energy agnostic parameter set.
- Build Fisher matrix assuming Poisson errors (and clustering).
- Takada & Bridle show that cluster counts are reasonably independent of 2-point statistics (but bispectrum?)

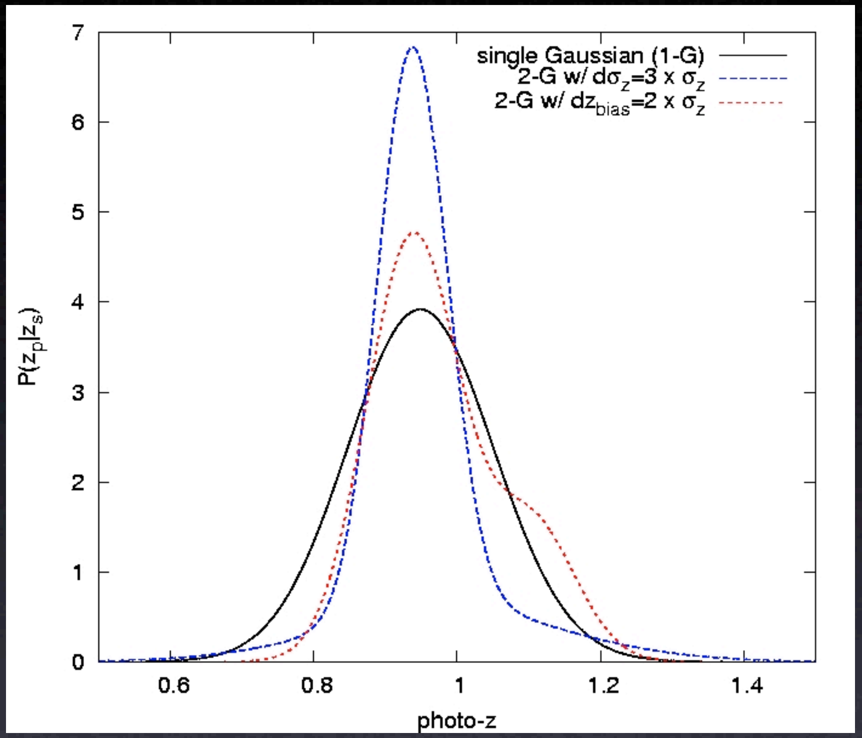
Results

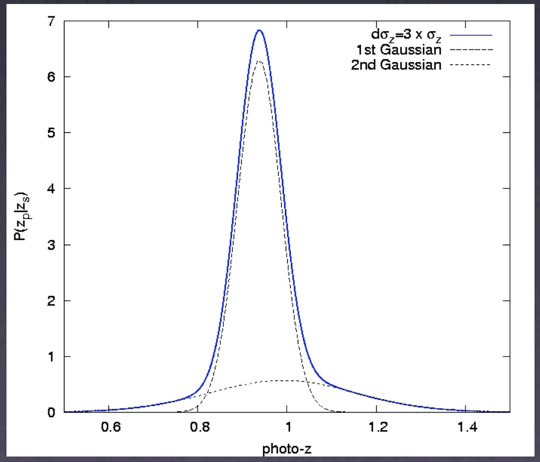
- None.

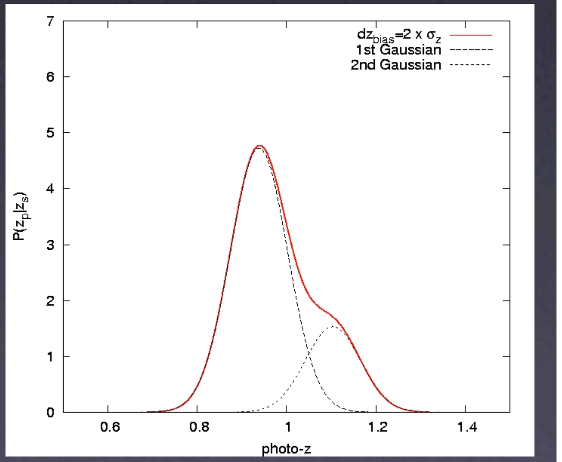
Results on spectroscopic prior: Zhaoming Ma

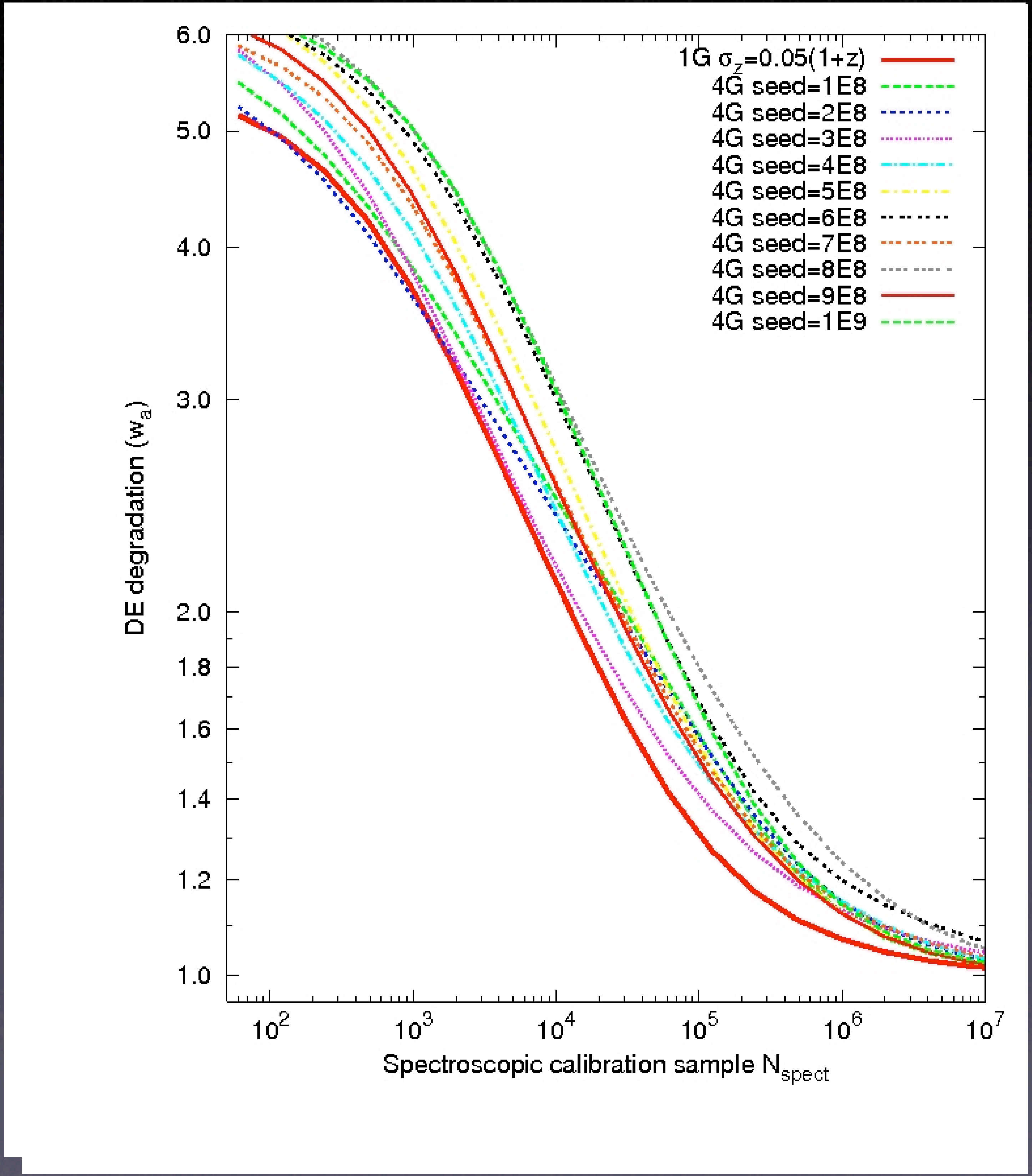
- Use simplified case of power-spectrum tomography, no intrinsic alignments, no shear cal errors, simple DE model.
- But now allow photo-z error distribution to be more complex than Gaussian, formalism allows arbitrary parametric function. Choose sum of up to 4 Gaussians, so 8 parameters per dz=0.1 instead of just 2 params.
- Drop assumption that true n(z) is known; we only know distribution of *photo-z's*.
- Now plot dark-energy uncertainty versus size of unbiased spectroscopic redshift survey.

Results on spectroscopic prior: Zhaoming Ma









Results on spectroscopic prior: Zhaoming Ma

- Use of single-Gaussian model underestimates size of required training set by few-x, or overestimates DETF FOM by 40-100%, depending upon true fiducial model.
- Few 10,000 - 100,000 spectra are sweet spot for LSST or SNAP-scale survey (but must be complete!)
- 6- and 8-parameter photo-z distributions are not really worse than 4-param distributions; convergences occurs since WL does not care about detailed structure in z.
- Full simulation under development will treat catastrophic errors, non-parametric (discrete) z distributions, inclusion of bispectrum, cross-correlations, intrinsic alignments, etc.

(Near!) Future results

- Quantify the constrained distance-growth combinations from a WL + galaxy survey of given size and source density, coupled with spectro survey of given size.
- Can construct a figure of merit from this information, or project it onto any chosen model.
  - Example: Albrecht & Bernstein, project DE proposals onto a DE model with arbitrary w function.
- Optimize suites of experiments, determine required photo-z calibration size and shear-systematic levels.
- Later: same code can analyse SKA lensing, recombination-era or CMB lensing constraints, or any combination.

Example: Albrecht's EOS PCA metrics

- Describe w(a) as a stepwise function.
- Express constraints as a series of eigenfunctions
- About 2 eigenfunctions constrained to <0.5 accuracy when current experiments are complete.
- 5-7 useful constraints with future data
- A. Albrecht & GMB (2007), extension of *Dark Energy Task Force* work.
- In this view, space gain has larger numerical value.