We present a framework for analyzing weak gravitational lensing survey data, including lensing and source- density observables, plus spectroscopic redshift calibration data. All two-point observables are predicted in terms of parameters of a perturbed Robertson–Walker metric, making the framework independent of the models for gravity, dark energy, or galaxy properties. For Gaussian fluctuations, the two-point model determines the survey likelihood function and allows Fisher matrix forecasting. The framework includes nuisance terms for the major systematic errors: shear measurement errors, magnification bias and redshift calibration errors, intrinsic galaxy alignments, and inaccurate theoretical predictions. We propose flexible parameterizations of the many nuisance parameters related to galaxy bias and intrinsic alignment. For the first time, we can integrate many different observables and systematic errors into a single analysis. As a first application of this framework, we demonstrate that: uncertainties in power-spectrum theory cause very minor degradation to cosmological information content; nearly all useful information (excepting baryon oscillations) is extracted with ≈ 3 bins per decade of angular scale; and the rate at which galaxy bias varies with redshift, substantially influences the strength of cosmological inference. The framework will permit careful study of the interplay between numerous observables, systematic errors, and spectroscopic calibration data for large weak lensing surveys.

Key words: gravitational lensing – cosmological parameters – relativity

Online-only material: color figures

1. INTRODUCTION

Weak gravitational lensing of background sources can produce exceptionally strong constraints on cosmological parameters and tests of general relativity. Initial analyses considered the two-point correlation function (or, equivalently, power spectrum) of the shear pattern induced on a single population of background galaxies (Miralda-Escudé 1991; Kaiser 1992; Blandford et al. 1991). A wealth of new statistics, however, have been suggested as more powerful means to extract information from weak lensing (WL): cross-power spectra of multiple source populations with distinct redshift distributions (also known as “tomography”; Hu 1999); the correlation of shear with foreground galaxy clusters (Jain & Taylor 2003), or more generally the cross-correlation of lensing shear with the galaxy distribution (Bernstein & Jain 2004; Zhang et al. 2005); joint analyses of density—density, density—shear, and shear—shear correlations in an imaging survey (Hu & Jain 2004); cross-correlation of magnification as well as shear (Jain 2002); use of the cosmic microwave background (CMB; Hu & Okamoto 2002; Hirata & Seljak 2003) or recombination era 21 cm signals (Metcalfe & White 2007; Pen 2004; Zahn & Zaldarriaga 2006) as source planes; cross-correlation of source density or shear with a distinct spectros- copic galaxy survey population (Newman 2008; Schneider et al. 2006); and the use of three-point statistics (Takada & Jain 2004) or statistics such as peak counts (Hennawi & Spergel 2005; Wang et al. 2004; Marian & Bernstein 2006) to move beyond two-point information. Each of these potential innovations has been individually analyzed and shown to improve cosmological constraints. The first goal of this paper is to consider the simultaneous use of all of these observable statistics: can we forecast the cosmological information that they will yield collectively in future surveys? Can we start to develop a framework in which all these signals could be analyzed simultaneously in a real experiment?

In parallel with the increasing variety of proposed WL signals, the community has identified a series of potential astrophysical and instrumental nonidealities in WL data which, if ignored, would lead to substantial systematic errors in the inferred cosmology. These include: finite accuracy in our ability to predict the deflecting mass power spectrum due to nonlinearities (Jain & Seljak 1997) and baryonic physics (Zhan 2006; Jing et al. 2006); intrinsic alignments (IA) between galaxy shapes (Croft & Metzler 2000) and between galaxy shapes and the local mass distribution (Hirata & Seljak 2004) that are not induced by lensing; multiplicative “shear calibration” errors in the derivation of lensing shear from galaxy images (Ishak et al. 2004; Huterer et al. 2006); additive “spurious shear” due to uncorrected point-spread function (PSF) ellipticity or other imaging systematics (Huterer et al. 2006; Amara & Refregier 2007); and errors in the assignment of redshifts to the source populations (Ma et al. 2006). The impact of these systematic error sources on cosmological inferences has been analyzed by different means, but a second goal of this paper is to produce a comprehensive forecast that considers their presence simultaneously.

Previous work has shown that these multiple sources of information and systematic error in WL surveys can interact in interesting ways. For example, in the presence of tomographic data, many systematics are readily distinguishable from cosmological signals and can hence be diagnosed and corrected internally to a survey; this approach is called self-calibration (Huterer et al. 2006). It has also been shown that combining galaxy density and lensing correlations can lead to self-calibration of shear calibration errors (Bernstein & Jain 2004) and the uncertainties in galaxy biasing (Hu & Jain 2004; Zhan 2006). Intrinsic alignments of galaxies can be diagnosed and corrected if tomographic information is available (King & Schneider 2003); however, this places substantially greater demands on the precision and accuracy of redshift assignment than would otherwise be needed.
modifications that might cause deviations from \( \Lambda \) models of nearly all proposed systematic errors. In Section 3, we discuss in detail the two-point statistics and very general treatment of WL survey observables that allows the incorporation of all suggested systematic errors needed to turn the lensing analysis into a tractable finite-dimensional problem. Further application of the framework to two-point analyses of WL survey forecasts will be done in future papers.

The third goal of this work is to describe the constraints by WL in a language that is not tied to a specific cosmological model. Most forecasts for WL survey constraints are done within the context of a universe that has homogeneous dark energy with equation of state \( w = w_0 + w_a(1 - a) \). Projecting the WL experiment onto this model gives concrete predictions, but obscures what the WL is really measuring. So the analysis framework presented here will be dark-energy agnostic, meaning that no specific model is assumed. We will be very explicit about the assumptions made in the analysis and try to keep them to a minimum. In fact, a great strength of WL experiments is their ability to test general relativity itself, so we seek an analysis method that is general enough to incorporate such tests. Similar to the approach of Knox et al. (2006), our analysis results in constraints on the distance and growth functions \( D(z) \) and \( g_a(z) \) without reference to the particular dark-energy or gravity modifications that might cause deviations from \( \Lambda \)CDM.

In the following section, we describe a “kitchen-sink” formalism for WL survey observables that allows the incorporation of all suggested two-point statistics and very general treatments of nearly all proposed systematic errors. In Section 3, we give a likelihood function and Fisher matrix for an unbiased spectroscopic redshift survey of source galaxies. Then we briefly describe a software implementation of the lensing and spectroscopy likelihood calculations. We describe our model for the evolution of the lensing potential power spectrum in Section 5, and in Section 6 we describe generic models used for the nuisance functions required in the lensing survey analysis. In Section 7, we use the implementation of these methods to investigate the proper choices for the bin sizes and grid spacings needed to turn the lensing analysis into a tractable finite-dimensional problem. Further application of the framework to survey forecasting will be done in future papers.

An earlier version of this WL analysis formalism was used to generate forecasts for the Dark Energy Task Force (Albrecht et al. 2006) and is described in an appendix to that report.

2. THE WEAK LENSING TWO-POINT LIKELIHOOD

2.1. Observables

We make the assumption that the universe has only weak scalar perturbations to a homogeneous and isotropic four-dimensional metric. In this case, the metric can be written in the Newtonian gauge as a perturbed Robertson–Walker metric

\[
 ds^2 = (1 + 2\Psi)dt^2 - a^2(t)(1 + 2\Phi)\left[dq^2 + \chi_0^2S_2^2(\chi/\chi_0)(\phi^2 + \sin^2 \theta d\phi^2)\right].
\]

(1)

We assign all mass and sources in the universe to a series of narrow spherical shells centered at redshifts \( a_i = (1 + z_i)^{-1} \) for \( i \in \{1, \ldots, N_s\} \). There is a comoving angular diameter distance \( D_i \) to each shell, and the comoving radial extent of each shell is \( \Delta \chi_i \). Note that the Robertson–Walker metric formula for angular-diameter distance is \( D = \chi_0 S_2(\chi/\chi_0) \), where \( \chi_0 \) is the comoving radius of curvature of the universe. For small values of the curvature \( \omega_k \equiv -k/\chi_0^2 \), we have

\[
 \Delta \chi_i \approx \Delta D \left( 1 - \omega_k D_i^2 / 2 \right) = \frac{D_i + 1 - D_i - 1}{2} \left( 1 - \omega_k D_i^2 / 2 \right). \quad (2)
\]

The Robertson–Walker metric also requires \( \Delta D \Delta z_i / h(z_i) = \Delta D_i / a_i^2 h(a_i) \). In this paper the Hubble parameter will be written as \( H(z) = h(z)H_{100} \), \( H_{100} = 100 \text{ km s}^{-1} \text{ Mpc}^{-1} \), and all distances will be in units of \( c/H_{100} = 2998 \text{ Mpc} \).

We assume that the photon sources in a survey will be divided into a series of sets \( \alpha \in \{1, 2, \ldots, N_s\} \). Note the use of latin indices for redshift shells and greek for source sets. We follow Hu & Jain (2004) by assigning each source set up to two observables: first its sky-plane density fluctuations \( g_{\alpha}(\theta, \phi) \), and second a lensing convergence \( \kappa_{\alpha}(\theta, \phi) \). The convergence \( \kappa \) might be inferred from the shear or flexion (Goldberg & Bacon 2005) of galaxies, by a quadratic estimator on the CMB or 21 cm radiation fluctuations, or by any other observable except the source density. The sources can be assigned to sets by photometric or spectroscopic redshift, or even cruder color criteria (Jain et al. 2007), but there could be other criteria such as galaxy type, or perhaps observation by different instruments. We demand only that the criteria for division of the sources be spatially homogeneous, and that the division be invariant under application of gravitational lensing distortion. For notational convenience, we assign each set a nominal redshift \( z_{\alpha} \), but a set can span a broad redshift range. If the sources are discrete objects such as galaxies, then the mean density on the sky of members of each set is denoted \( n_{\alpha} \).

A source in set \( \alpha \) has a probability \( p_{\alpha i} \) of lying on redshift shell \( i \). The collection of galaxies in set \( \alpha \) on shell \( i \) will be called the subset \( a_i \). The survey is assumed to tell us only which set any individual galaxy belongs to, but not which subset. The \( p_{\alpha i} \) are parameters which must be constrained by the lensing survey data or by additional observations, e.g., a spectroscopic redshift survey.

When the lensing sources are drawn from a spectroscopic survey (or when the source is the CMB), then the redshift probability is known a priori, and in particular the sets are probably divided by redshift so that \( p_{\alpha i} \) is essentially the identity matrix. The formalism can obviously accommodate the simultaneous analysis of WL samples with varying modes of redshift assignment.

Both the source-density fluctuation \( g_{\alpha} \) and convergence \( \kappa_{\alpha} \) have a component due to intrinsic fluctuations plus a component due to gravitational lensing. Both are also measured as weighted sums over their respective subsets. We have

\[
 1 + g_{\alpha}(\theta, \phi) = \sum_i p_{\alpha i} \left[ 1 + g_{\alpha i}^{\text{int}}(\theta, \phi) \right] \left[ 1 + q_{\alpha i} \kappa_{\alpha i}^{\text{len}}(\theta, \phi) \right].
\]

(3)

\[
 1 + \kappa_{\alpha}(\theta, \phi) = \sum_i p_{\alpha i} \left[ 1 + f_{\alpha i} \kappa_{\alpha i}^{\text{len}}(\theta, \phi) \right].
\]

(4)

Here, we have assigned each subset a magnification bias factor \( q_{\alpha i} \) and a shear calibration factor \( f_{\alpha i} \). In a simple flux-limited selection, the magnification bias factor will be determined by the logarithmic slope of the counts versus flux, and is typically of order unity. The shear calibration factor allows for the possibility that the inferred lensing convergence is mismeasured by some factor \( 1 + f_{\alpha i} \) due to multiplicative errors in the lensing methodology, e.g., as investigated by Heymans et al. (2006).
In the limit \( g^{\text{int}} \ll 1 \) and \( \kappa^{\text{int}} \ll 1 \), we can drop the second-order term in Equation (3) and write

\[
g_a = \sum_i p_{ai} \left[ g_{a\alpha i}^{\text{int}} + q_{ai} \kappa_i^{\text{int}} \right],
\]

\[
\kappa_a = \sum_i p_{ai} \left[ k_{a\alpha i}^{\text{int}} + (1 + f_{ai}) \kappa_i^{\text{int}} \right].
\]

In this case the equations are linear in all the angular functions \( g \) and \( \kappa \), so we can decompose them into spherical harmonic coefficients \( g_{a\ell m}^{\text{int}}, \kappa_{i\ell m}^{\text{int}}, \) etc, and Equations (5) and (6) hold independently for every harmonic \( \ell \) and \( m \). We henceforth assume that the spherical-harmonic decomposition has been executed for all the angular functions \( g, \kappa \), and suppress the \( \ell m \) indices for brevity.

We note that while \( \kappa^{\text{int}} \ll 1 \) is a good approximation over most of the sky, \( g^{\text{int}} \ll 1 \) is a poor approximation for thin density slices on smaller angular scales. We will forge ahead nonetheless most of the sky, recognizing that a real analysis of data with magnification bias may require inclusion of the nonlinearity coupling between spherical harmonics that is induced by magnification bias on highly structured density fields.

The lensing convergence is determined entirely by the metric if we make the assumption that light rays are following its null geodesics. The paths of null geodesics are determined by the lensing potential

\[
\phi \equiv \frac{1}{2}(\Psi - \Phi).
\]

For each of our redshift shells, we derive a field \( \psi \) from the projected lensing potential via

\[
\psi_i \equiv 2\nu^2 \int_{\Delta x} \phi a \, d\chi,
\]

where the derivatives are taken with respect to angles on the sky. We will generally assume that \( \psi \), like the observables, has been decomposed into spherical harmonics, and we will take the flat-sky approximation.

With the definition in Equation (8), and the adoption of the weak-lensing limit and Born approximation, the lensing convergence is

\[
\kappa_i^{\text{int}} = \sum_j A_{ij} \frac{\psi_j}{2\alpha_j D_j},
\]

\[
A_{ij} \equiv \begin{cases} \frac{D_{ij}}{D_i} \approx (1 - D_j/D_i)(1 - \omega_k D_i D_j/2), & z_i > z_j, \\ 0, & z_i \leq z_j. \end{cases}
\]

Here, \( D_{ij} \) is the comoving angular diameter distance to \( z_j \) as viewed from \( z_i \). In summary, the observables from the survey are, for each spherical harmonic,

\[
g_a = \sum_i p_{ai} \left[ q_{ai} \sum_j A_{ij} \frac{\psi_j}{2\alpha_j D_j} + g_{a\alpha i}^{\text{int}} \right],
\]

\[
\kappa_a = \sum_i p_{ai} \left[ (1 + f_{ai}) \sum_j A_{ij} \frac{\psi_j}{2\alpha_j D_j} + \kappa_{a\alpha i}^{\text{int}} \right].
\]

We reiterate that these equations depend only upon the assumption of a Robertson–Walker metric with scalar perturbations plus the approximation that magnification bias and intrinsic density fluctuations are additive.

The equations for the two observables are symmetric under the interchange of \( g \leftrightarrow \kappa \) and \( q \leftrightarrow (1 + f) \). Since \( q \sim 1 + f \), the lensing effects are similar. However, the intrinsic density fluctuations \( g^{\text{int}} \) are \( 300 \) times stronger than \( \kappa^{\text{int}} \), breaking the symmetry. Density-field observations are dominated by the intrinsic signal, while convergence (shear) observations are dominated by lensing effects.

### 2.2. Degeneracies

Equations (9)–(11) reveal a family of degeneracies present in lensing observations as described in Bernstein (2006). The transformations

\[
\begin{align*}
D_j &\to D_j(1 + \alpha_0 + \alpha_1 D_j + \alpha_2 D_j^2), \\
\psi_i &\to \psi_i(1 + \alpha_0 + 2\alpha_1 D_i + 2\alpha_2 D_i^2), \\
\omega_k &\to \omega_k + \alpha_2
\end{align*}
\]

leave the observables unchanged to first order in \( \{\alpha_0, \alpha_1 D, \alpha_2 D^2, \omega_k D^2\} \). It will hence be impossible for lensing+density surveys to constrain \( \omega_k \) or any quadratic (in \( D \)) deviations to \( \ln D \) unless there are prior constraints on these variables or on \( \psi, g^{\text{int}}, \) or \( \kappa^{\text{int}} \). Constraint on these three degeneracies is unlikely to arise from models of intrinsic clustering or alignment since it is unlikely that a priori models of the redshift dependence of galaxy bias could reach high precision. We hence expect that these degeneracies are going to be broken by theoretical models of the potential fluctuation power spectrum or by other distance indicators such as supernovae or BAO.

### 2.3. Limber Approximation

To forecast the constraints on the parameters of this model, we require a likelihood expression for the observables. The two fundamental assumptions we make are the following.

1. The distributions of the lensing potential, intrinsic galaxy density fluctuations, and intrinsic shape correlations \( \psi_i, \kappa_{a\alpha i}^{\text{int}}, \) and \( \kappa_{\alpha i}^{\text{int}} \) are described by a multivariate Gaussian with zero mean.
2. The Limber approximation is valid, and there is no correlation between these variables on distinct redshift shells or between different spherical harmonics,

\[
\begin{align*}
\langle X_{\ell m} Y_{\ell' m'} \rangle &= \delta_{\ell \ell'} \delta_{m m'} (D_i^2 \Delta \chi_i)^{-1} P_{X Y}(\ell/D_i), \\
\langle X_{\ell m} \psi_{\ell' m'} \rangle &= -2\delta_{\ell \ell'} \delta_{m m'} a_i (\ell/D_i)^2 P_{X \psi}(\ell/D_i), \\
\langle \psi_{\ell m} \psi_{\ell' m'} \rangle &= \delta_{\ell \ell'} \delta_{m m'} 4a_i^2 D_i^2 \Delta \chi_i (\ell/D_i)^2 p_{\psi \psi}(\ell/D_i),
\end{align*}
\]

where \( X, Y \in \{g^{\text{int}}, \kappa^{\text{int}}\} \), and \( P_{XY}(k) \) is the three-dimensional cross-spectrum of variables \( X \) and \( Y \) at epoch \( a_i \).

The first assumption ensures that the likelihood of an observation is fully specified by the expected covariance matrix of the observables. The second assumption implies that this covariance matrix can be expressed in terms of the three-dimensional cross-power spectra of the lensing potential, subset densities, and subset intrinsic correlations \( \{\phi_i, \delta_{a\alpha i}^{\text{int}}, \kappa^{\text{int}}_{a\alpha i}\} \) at each redshift shell.
2.4. Biases and Correlations

A typical convention is to express the galaxy density power \( P^{gg} \) as a bias-scaled version of the mass spectrum \( P^m \), plus a Poisson shot-noise contribution, and then describe the mass-galaxy covariance \( P^{mg} \) with a correlation coefficient

\[
p^{gg} = (b^g)^2 P^m + \frac{1}{\rho},
\]

\[
p^{mg} = b^g r^g P^m.
\]

The comoving volume number density \( \rho \) of the sources determines the shot noise for a Poisson process, but there is no guarantee that the galaxies are distributed in the mass distribution by a Poisson process. Even when the galaxies do not have Poissonian shot noise, we can still usually write the power in this way for some bias parameter \( b \); we just might keep in mind that \( r^g > 1 \) is formally allowed if the sources are not Poisson distributed.

Most generally, both the bias and correlation coefficients are different for each source subset as well as being functions of comoving wavenumber \( k \). Each set \( \alpha \) has a nominal redshift \( z_\alpha \), and each subset has a redshift deviation \( \Delta z_{\alpha i} = z_i - z_\alpha \). Galaxies with bad photo-z errors could easily have different bias from those with good photo-zs; for example, highly biased early types tend to have better photo-zs. So our analysis methods should allow for this complication.

We will adopt the bias/correlation notation for the intrinsic galaxy density fluctuations and for the intrinsic convergence \( \kappa^{int} \), except that we will parameterize the bias and covariance with respect to the lensing potential rather than mass distribution. If \( P^{gg}_{\alpha i} \) is the three-dimensional cross-power between density fluctuations in subsets \( \alpha i \) and \( \beta j \) at wavenumber \( k \), and we write \( P^\phi \) for the lensing potential three-dimensional power spectrum, then we express

\[
P^{gg}_{\alpha i} = b^g_{\alpha i} b^g_{\beta j} r^{gg}_{\alpha i} \left( \frac{2a_i}{5\Delta \omega_m} \right)^2 k^4 P^\phi + \frac{\delta_{\alpha \beta}}{\rho_{\alpha i}},
\]

\[
P^{gg}_{\alpha i} = b^g_{\alpha i} b^g_{\beta j} r^{gg}_{\alpha i} \left( \frac{2a_j}{5\Delta \omega_m} \right)^2 k^4 P^\phi + \frac{\delta_{\alpha \beta} \sigma^2}{\rho_{\alpha i}},
\]

\[
P^{gg}_{\alpha i} = b^g_{\alpha i} b^g_{\beta j} r^{gg}_{\alpha i} \left( \frac{2a_i}{5\Delta \omega_m} \right)^2 k^4 P^\phi,
\]

where \( \rho_{\alpha i} \) is a comoving volume density of the galaxy subset in the shell, and \( \sigma^2 \) is a measure of the shear noise per galaxy. These describe the normal “shape noise” term in the shear power spectrum and the shot noise in the density field. If flexions or other observables are used to infer the convergence, then the shape noise term may have a different form.

Moreover, if \( P^{\phi \phi}_{\alpha i} \) is the cross-power between the density of subset \( \alpha i \) and lensing potential, we express

\[
P^{\phi \phi}_{\alpha i} = -b^\phi_{\alpha i} r^\phi_{\alpha i} \frac{2a_i}{3\Delta \omega_m} k^3 P^\phi,
\]

\[
P^{\phi \phi}_{\alpha i} = -b^\phi_{\alpha i} r^\phi_{\alpha i} \frac{2a_i}{3\Delta \omega_m} k^3 P^\phi.
\]

Note that specifying the bias and correlation \( b^g \) and \( r^g \) of the intrinsic convergence with the lensing potential is equivalent to giving the “GI” and “II” intrinsic alignment information, in the notation of Hirata & Seljak (2004).

The lensing power \( P^\phi \) is a function of \( z \) and \( k \). The biases and correlation coefficients \( b^g, r^g, b^\phi, r^\phi \) are functions of \( k \), the nominal redshift \( z_\phi \) of the source set, and \( \Delta z_{\phi i} \), the difference between the subset redshift and the nominal set redshift.

Most complicated are the cross-correlation coefficients such as \( r^{gg}_{\alpha i} \), which are, most generally, functions of \( k, z_i, \) and both redshift errors \( \Delta z_{\alpha i} \) and \( \Delta z_{\beta j} \). In order for the covariance matrix of all these fields to be symmetric, we require \( r^{gg}_{\alpha i} = r^{gg}_{\beta j} \) and the symmetry \( r^{gg}_{\alpha i} = r^{gg}_{\beta j} \). Otherwise, the correlation coefficients are free to vary subject to the constraint that the overall correlation matrix of the potential and all fluctuations must remain nonnegative.

This parameterization of the fluctuations of the potential and the intrinsic fluctuations is completely general—we have not introduced any further assumptions into the model as long as all the \( b_s, r_s \), and \( P^\phi \) are free parameters (non-negative in the last case).

2.5. The Two-point Statistics

Combining the formula for observables in Equation (11), the Limber formulae in Equation (13), and the bias notations in Equations (16) and (17), the covariance matrix for the observables \( \{g_\alpha, \kappa_\alpha\} \) at a given multipole can be broken into three submatrices,

\[
C^{gg}_{\alpha \beta} = \langle g_\alpha g_\beta \rangle = \sum_{ij} p_{\alpha i} p_{\beta j} \left[ q_{\alpha i} q_{\beta j} \right.
\]

\[
\times \sum_n A_{\alpha i} A_{\beta j} \Delta \kappa_{\alpha i}^n k^4 P^\phi(k_n)
\]

\[
+ q_{\alpha i} A_{\beta j} \frac{2a_i}{3\Delta \omega_m} D_i^{-1} b^\phi_{\alpha i} r^\phi_{\beta j} k^3 P^\phi(k_i)
\]

\[
+ q_{\beta j} A_{\alpha i} \frac{2a_j}{3\Delta \omega_m} D_j^{-1} b^\phi_{\beta j} r^\phi_{\alpha i} k^3 P^\phi(k_j)
\]

\[
+ \delta_{ij} \frac{2a_i}{3\Delta \omega_m} D_i^{-2} \Delta \kappa_i^{-1} b^g_{\alpha i} b^g_{\beta j} r^{gg}_{\alpha i} r^{gg}_{\beta j} k^4 P^\phi(k_i)
\]

\[
\times \frac{\delta_{\alpha \beta}}{n_a},
\]

\[
C^{\kappa g}_{\alpha \beta} = \langle \kappa_\alpha g_\beta \rangle = \sum_{ij} p_{\alpha i} p_{\beta j} \left[ (1 + f_{\alpha i}) q_{\beta j} \right.
\]

\[
\times \sum_n A_{\alpha i} A_{\beta j} \Delta \kappa_{\alpha i}^n P^\phi(k_n)
\]

\[
+ (1 + f_{\alpha i}) A_{\beta j} \frac{2a_i}{3\Delta \omega_m} D_i^{-1} b^\phi_{\alpha i} r^\phi_{\beta j} k^3 P^\phi(k_i)
\]

\[
+ q_{\beta j} A_{\alpha i} \frac{2a_j}{3\Delta \omega_m} D_j^{-1} b^\phi_{\beta j} r^\phi_{\alpha i} k^3 P^\phi(k_j)
\]

\[
+ \delta_{ij} \frac{2a_i}{3\Delta \omega_m} D_i^{-2} \Delta \kappa_i^{-1} b^g_{\alpha i} b^g_{\beta j} r^{gg}_{\alpha i} r^{gg}_{\beta j} k^4 P^\phi(k_i)
\]

\[
\]
\[+(1 + f_{\alpha i})A_{ij} \frac{2a_i}{\delta \omega_m} D_j^{-1} b_{\alpha j} r_{\alpha j}^* k_i^3 P^\phi(k_i)\]
\[+(1 + f_{\beta j})A_{ij} \frac{2a_i}{\delta \omega_m} D_j^{-1} b_{\alpha i} r_{\alpha i}^* k_j^3 P^\phi(k_j)\]
\[+ \delta_{ij} \left( \frac{2a_i}{\delta \omega_m} \right)^2 D_j^{-2} \Delta \chi_i^{-1} b_{\alpha i}^* b_{\beta j}^* r_{\alpha i} r_{\beta j}^* k_i^4 P^\phi(k_i)\]
\[+ \delta_{\alpha \beta} \left( \frac{\sigma^2_{\alpha i}}{n} \right) . \] (20)

The comoving wavevector is \( k_i = \ell / D_i \). In each equation, note that only one of the last three terms is nonzero depending on whether \( j < i \), \( j > i \), or \( j = i \), respectively. For the shear–shear correlation \( C^{\kappa \kappa} \), the \( i = j \) term is recognizable as the “II” intrinsic-correlation effect of Hirata & Seljak (2004), while the \( i < j \) and \( j > i \) terms are their “GI” effect.

Note that the last term in the density–shear expression \( C^{\kappa \rho} \) is an additional intrinsic-correlation term between the galaxy density and the intrinsic shapes, which is distinct from the covariance between the lensing potential and shear. This galaxy–shear correlation has been constrained in the context of the covariance between the lensing potential and shear. This is above. Under our Gaussian assumption, the total likelihood for the observables is independent at each multipole, and \( C_\ell \) to be the covariance matrix derived above. Under our Gaussian assumption, the total likelihood for the survey is
\[-2 \ln L = \sum_{\ell m} \left[ \mathbf{d}_\ell^T \mathbf{C}^{-1}_{\ell m} \mathbf{d}_{\ell m} + \ln |\mathbf{C}_\ell| \right] . \] (21)

Forecasts of survey performance are made using the Fisher matrix. The usual formula for zero-mean Gaussian distributions applies (Tegmark et al. 1997). We reduce the mode sum to a series of \( N_\ell \) bins centered on multipoles \( \ell_i \), then the Fisher matrix element for parameters \( p \) and \( q \) is
\[F_{pq} = \sum_{i=1}^{N_\ell} \frac{(2\ell_i + 1) \Delta \chi_i f_{sky}}{2} \left| \mathbf{C}^{-1}_{\ell_i} \right| \text{Tr} \left[ \mathbf{C}^{-1}_{\ell_i} \frac{\partial \mathbf{C}_{\ell_i}}{\partial p} \frac{\partial \mathbf{C}_{\ell_i}}{\partial q} \right] . \] (22)

Examination of Equations (18)–(20) shows that all derivatives of \( \mathbf{C} \) with respect to parameters are very simple. The calculation of the Fisher matrix is reduced to rapid linear algebra, significantly accelerated by exploiting the very sparse nature of most of the derivative matrices.

We have thus succeeded in producing a likelihood function for the most general joint lensing+density survey for the case of Gaussian likelihoods limited to two-point statistics. Given a likelihood we can, of course, form a Fisher matrix for forecasting, or we can execute a maximum-likelihood analysis of real data. Since this likelihood function makes no mention of a particular dark-energy model, we see that the parameterization chosen here permits a highly flexible analysis. Indeed, no theory of gravity or initial conditions of the universe has been assumed either, just the existence of a Newtonian gauge metric on an RW background cosmology. The lensing potential power spectrum \( P^\phi(k, z) \) appears as a series of free parameters, as do the bias and correlation coefficients of the galaxy density and intrinsic alignments.

We have variables that describe in the most general possible fashion the important systematic errors, excepting additive shear contamination.

1. Uncertainty in power-spectrum theory will be expressed through prior distributions on the \( P^\phi_\ell \) parameters.
2. Shear calibration errors arise through finite prior uncertainty on the \( f_{\alpha i} \).
3. Magnification bias calibration errors arise through finite prior uncertainty on the \( q_{\alpha i} \).
4. Intrinsic alignments are embodied through the \( b^\alpha, r^\alpha, r^{\kappa \rho} \), and \( r^{\kappa \rho} \) coefficients.
5. Redshift distribution errors are manifested through the uncertainties in the \( P_{\alpha i} \) probabilities.

The cost of this great generality is that there are a huge number of nuisance parameters, enough to make us doubt whether the maximum-likelihood analysis—or even the Fisher matrix analysis—is feasible.

2.7. Parameter Inventory

The WL survey covariance matrix has a horrendously large number of parameters. The cosmological treatise lies in the following:

1. the two global cosmological parameters \( \omega_m \) and \( \omega_k \);
2. the distances \( D_i \), which encode the expansion history of the universe in \( N_\ell \) steps. The \( \Delta \chi_i \) and the Hubble parameters \( h(z_i) \) can be expressed in terms of these and \( \omega_k \).
3. the metric-potential power spectra \( P^\phi_\ell \), which describe the growth of dark-matter structure. For \( N_\ell \) bins in \( \ell \), there will be \( N_\ell N_z \) distinct matter-power parameters in the model. A prediction for the growth of potential fluctuations will typically be an important element of any cosmological scenario under test, so the \( P^\phi_\ell \) can be replaced as parameters by a much smaller number of cosmological parameters.

There are then a large number of nuisance parameters. If there are \( N_{\alpha i} \) nonempty source subsets, the nuisance parameters are the following:

1. the redshift distribution parameters \( p_{\alpha i} \) with \( N_{ss} - N_z \) degrees of freedom;
2. the shear-calibration errors \( f_{\alpha i} \), another \( N_{ss} \) degrees of freedom;
3. the magnification bias coefficients \( q_{\alpha i} \), another \( N_{ss} \) degrees of freedom;
4. the source-density biases \( b_\alpha^\rho \) and correlation coefficients \( r_\alpha^\rho \), with respect to \( \phi \), which may be scale dependent, yielding \( 2 N_\ell N_{ss} \) degrees of freedom;
5. the intrinsic alignment power and correlations with the mass, \( b^\alpha_\ell \) and \( r^\alpha_\ell \), another \( 2 N_\ell N_{ss} \) parameters; and
6. the correlation coefficients \( r^{\kappa \rho}_{\alpha i} \), \( r^{\kappa \rho}_{\alpha j} \), and \( r^{\kappa \rho}_{\alpha j} \), which may also be scale dependent. The number of such parameters is \( \approx 3 N_\ell N_{ss}^2 / 2 N_z \).
The number of nuisance parameters for a nonparametric analysis is enormous. If we are analyzing a photo-z survey with typical errors $\Delta z \approx 0.05(1+z)$, then we would typically want to space the redshift shells logarithmically in $1+z$ with $\Delta \ln a \approx 0.02$ so that we resolve the redshift distribution of each photo-z bin. In this case, $N_z \approx 100$ bins span $0 < z < 5$, and we will require $N_{\text{tot}} \gtrsim 1000$ if we track all subsets out to $\pm 3\sigma$ of the photo-z distribution.

To reduce the dimensionality of the likelihood function, we can replace many of the discrete nuisance parameters by the values of parameterized functions for the nuisance variables. Table 1 lists the variables in the WL likelihood function that can be replaced by parametric functions. The nuisance variables are functions of: wavevector $k$; redshift $z$; and redshift difference $\Delta z = z_a - z_i$ between the nominal and true redshifts of a source subset. In later sections, we will describe the parametric functions that we have implemented to reduce the number of degrees of freedom in the model. Each time we introduce a parametric function, we need to choose a fiducial parameter set and a prior distribution for the parameters.

### 3. SPECTROSCOPIC REDSHIFT LIKELIHOOD

If we draw a single member from source set $\alpha$ and measure its spectroscopic redshift in an unbiased fashion, then by definition the likelihood of the spectroscopic redshift being on shell $i$ is $p_{ai}$. If we measure $N_{\text{spec}}$ redshifts, and find that $N_{\text{spec}}^\alpha$ are on shell $i$, then the likelihood is

$$\ln L = \sum_i N_{\text{spec}}^\alpha \ln p_{ai}. \quad (23)$$

This is true if the redshifts are statistically independent and can be dispersed across the sky to eliminate source correlations. We assume this limit.

Following Ma & Bernstein (2008), the Fisher matrix for the parameters $\{p_{ai}\}$ resulting from the unbiased spectroscopic observations is

$$F_{ai \beta j} = \left( \frac{\partial^2 (-\ln L)}{\partial p_{ai} \partial p_{aj}} \right) = N_{\text{spec}}^\alpha \frac{\delta_{ij} \delta_{a \beta}}{p_{ai}}. \quad (24)$$

We add this Fisher information to the density–lensing Fisher matrix in Equation (22) when considering the constraints offered by a WL survey that is combined with an unbiased spectroscopic redshift survey drawn from one or more of the source population sets.

We do not, in general, presume any functional form for the $p_{ai}$ redshift distributions when the sets are assigned from photo-z’s.

### 4. AN IMPLEMENTATION

A package of C++ classes implements the Fisher matrix calculation for Gaussian lensing+density observations, the spectroscopic survey Fisher matrix in Equation (24) if appropriate to the planned experiment. Note that the cross-correlations in the WL survey data offer constraints on the redshift distribution even if there is no unbiased spectroscopic survey ($N_{\text{spec}}^\alpha = 0$).

### 5. POWER SPECTRUM MODELS

In most cosmological models, theory will offer strong guidance to the form of the lensing potential power spectrum $P^\phi(k, z)$. In most forecasting or data-reduction codes, this is a fully deterministic function of a small number of cosmological parameters. In our analysis, however, the theoretical prediction is taken as the mean $P^\phi$ value of a prior distribution of finite uncertainty.
5.1. Central Model

The WL likelihood given above can be calculated for any model that predicts $P^\phi$. We have chosen to implement a model that allows for failure of general relativity in describing growth of structure; but other models are possible if one wishes to test other tenets of general relativity. Under the following conditions:

1. the potential and the mass-energy density are related by the Newtonian Poisson equation, i.e., the field equation for nonrelativistic matter in general relativity;
2. nonrelativistic matter is the only significant inhomogeneous component of the universe, i.e., there is no dark energy clustering;
3. matter is conserved, $\rho_m \propto a^{-3}$, and
4. $\Phi = -\Psi$, as in the absence of anisotropic stress for general relativity,

then the potential power spectrum is related to the matter-density fluctuation spectrum $P^m$ via

$$k^3 P^\phi(k, a) = \left( \frac{3\omega_m}{2a} \right)^2 P^m(k, a).$$

Under these conditions, the linearized perturbations to the metric grow in a scale-free manner, so we can write

$$P^\phi_{lin}(k, a) = g^2_\phi(a) P^\phi_{prim}(k) T^2(k).$$

In our current code, the primordial power spectrum is a power law

$$\Delta^2_{prim}(k) \equiv \frac{k^3}{2\pi^2} P^\phi_{prim}(k) = \left( \frac{3\Delta_\zeta}{5} \right)^2 \left( k/k_0 \right)^{n_s-1}. \quad (27)$$

The curvature variation $\Delta_\zeta$ and spectral index $n_s$ are free parameters. A running of the slope could easily be added. The normalization wavenumber $k_0$ must be set by some convention. We typically adopt the five-year WMAP parameters as fiducial values (Komatsu et al. 2009).

The transfer function $T(k)$ is taken from Eisenstein & Hu (1999). It is a function of the matter and baryon densities $\omega_m$ and $\omega_b$. The impact of massive neutrinos could be added to the transfer function if desired. We ignore the baryon acoustic oscillations; experiments that try to exploit them will generate a distinct Fisher matrix for them.

The map from $P^\phi_{lin}$ to the nonlinear $P^\phi_m$ is derived using the prescription of Smith et al. (2003) for nonlinear $P^m$ combined with the Poisson equation (25). The Smith et al. (2003) formula also requires knowledge of $\Omega_m$ at the desired epoch, but it can be expressed in terms of other quantities that are already in our model: $\Omega_m(z_i) = \omega_m a_i^{-3} h^{-2}(z_i)$. We do not expect the Smith et al. (2003) formula to describe nonlinear growth to high accuracy for all (or any) cosmologies. It does, however, capture the dependence of nonlinear power on cosmological parameters to a level that suffices for forecasting purposes.

In general relativity, the growth function $g_\phi(a)$ is determined by the expansion history $H(z)$. Defining $F = \ln(a g_\phi)$, the growth equation is

$$F'' + \left( F' \right)^2 + F' \left( 2 + \frac{d \ln h}{d \ln a} \right) = \frac{3\omega_m}{2h^2 a^3}. \quad (28)$$

where a prime denotes differentiation with respect to $\ln a$. Note that $F'$ is the quantity of $d \ln g_m / d \ln a$ that appears in the peculiar velocity power spectrum for a tracer of mass.

If GR holds, then the above relations fully specify the model for $P^\phi$ given $\{\omega_m, \omega_b, n_s, \ln \Delta_\zeta \}$ plus the expansion history, which in turn is given by $\{D_e, \omega_k \}$. As a test of GR, we allow the growth function arbitrary deviations from the $F_{fid}$ that solves the GR growth equation for the fiducial expansion history,

$$\ln a g_\phi(a) = F_{fid}(a) + \delta F_i.$$  \quad (29)

The $\delta F_i$ become parameters of the likelihood function.

5.2. Model Errors

An important WL systematic is the expected finite accuracy in theoretical modeling of the power spectrum. We hence introduce an error function to describe the (logarithmic) difference between the power $P^\phi$ and the value predicted by the parametric model described in the previous paragraphs:

$$\ln P^\phi(k, a) = \ln P^\phi_{fid}(k, a) + \delta \ln P(k, a).$$

The nuisance function $\delta \ln P$ will be modeled by the “$k_z$” parametric form described in the Appendix. We parameterize the $\delta \ln P$ function by its values $\delta P_{ij}$ at a grid of points $(k_i, a_j)$ regularly spaced in $\ln k$ and $\ln a$. The $\delta P_{ij}$ is linearly interpolated between grid points. The $\delta P_{ij}$ become free parameters of the model, and hence parameters in the likelihood function. We then place an independent Gaussian prior on each $\delta P_{ij}$ which has mean of zero and a standard deviation of

$$\sqrt{\text{Var}(\delta P_{ij})} = 0.012 f_{Zhan} \begin{cases} 1 + 5 \ln(10 (k_i/k_1)), & k_i > k_1, \\ (k_i/k_1)^{\omega_m}, & k_i < k_1, \end{cases} \quad (31)$$

$$k_1 \equiv a^{-2.6} \text{Mpc}^{-1}. \quad (32)$$

This function is a fit to an estimate supplied by Hu Zhan of the impact of baryonic physics on the mass power spectrum (Zhan & Knox 2004; Jing et al. 2006). We scale the overall size of the theory-error systematic with the control scalar $f_{Zhan}$. We can also adjust the density $\Delta \ln k$ and $\Delta \ln a$ at which the $\delta P_{ij}$ grid points are spaced. This corresponds to setting some coherence length for theory errors in this space. In Section 7, we investigate the choice of these grid spacings. The procedure above means that we replace the $P^\phi_{fid}$ as parameters in our likelihood with a new (and hopefully smaller) set as below:

1. the small set $\{\omega_m, \omega_b, \Delta_\zeta, n_s\}$ that controls the linear power spectrum;
2. the $\{\delta F_i\}$ which defines the growth function versus redshift; and
3. a grid of theory error values $\delta P$, which are nuisance parameters to marginalize after construction of a Fisher matrix or likelihood. We have physically based priors to apply to these before marginalization.

Table 3 lists the input fields for the part of our forecasting code which constructs the power-spectrum model. The WMAP5 flat $\Lambda$CDM cosmology provides the fiducial values of all lensing power values; the program inputs define the behavior of the deviations from the theoretical model and the prior expectations on the size of such deviations.

6. NONPARAMETRIC NUISANCE MODELING

6.1. General Comments on Nuisance Functions

The WL likelihood contains many nuisance parameters that are discretized representations of nuisance functions. It is
common in the literature to assign some parametric form to a nuisance function, then marginalize over the parameters of the nuisance function to recover a purely cosmological likelihood. This can be a very dangerous approach: if the nuisance function does not in actuality follow the assumed form, then the process is invalid, and we may have greatly overestimated the power of the experiment to remove the systematic error from the signal. When marginalizing over a systematic, we must be sure that the assumed parametric form is sufficiently flexible to include any expected manifestation of the systematic. For example, we should not assume that systematics scale linearly with redshift unless there is a physical reason to expect this.

It is unfortunately not possible to model a completely free function with a finite number of free parameters. This becomes possible, however, if we limit the bandwidth of variation in the function. As an example, consider our power-spectrum theory error function $\delta \ln P(k, z)$. We could decompose $\delta \ln P$ into Fourier modes or polynomial terms over its finite $(k, z)$ domain. Retaining a finite number of modes or terms leads to a tractable parameterization, albeit with a maximum frequency or polynomial order that defines a coherence length for the reconstructed function. For $\delta \ln P$, we choose to limit the bandwidth using linear interpolation between a two-dimensional grid of specified values. In the Appendix, we describe the family of functions that we use to model the nuisance functions of $(k, z, \Delta z)$ that are common in the WL likelihood analysis (see Table 1).\(^1\) The Appendix describes both the functional form and the prior likelihoods on the parameters that are used to give the nuisance function the desired RMS uncertainty.

These models of nuisance functions are nonparametric in the sense of being able to reproduce very general types of behavior once the bandwidth is specified. The question remains: what is the proper choice of bandwidth to allow the nuisance function? Our approach is to find the bandwidth which causes the most damage to cosmological constraints under a prior that specifies the expected RMS fluctuations in the nuisance function. This is the most conservative approach. Typically, one finds the following: if the nuisance function is given a highly coherent, low-order functional form, then it is easily distinguished from cosmological signals and can be marginalized away with little damage to cosmological constraints. On the other hand, if the nuisance-function bandwidth is very high, then the broad WL kernel tends to average away the nuisance signal, leaving little trace in the cosmology. There is an intermediate point where the systematic error is most easily confused with cosmology. The conservative approach is to find this regime and use it for modeling the systematic error. In Section 7, we will find coherence lengths in $z$ and $k$ at which our systematics are most damaging.

6.2. Redshift Distributions

We specify the fiducial values of the photo-$z$ distribution $n_a$ and error probabilities $p_{ai}$ either with analytic formulae (e.g., photo-$z$ errors Gaussian in $\ln a$) or by taking the output of a simulation of galaxy detection and photo-$z$ assignment for the chosen survey. For spectroscopic samples or the CMB source plane, there are no photo-$z$ errors at all.

We do not place any parametric form or prior assumption on the $p_{ai}$. All redshift constraints arise either from the lensing survey data itself or from additional spectroscopic data. The likelihood arising from spectroscopic redshift samples is described in Section 3.

6.3. Shear Calibration and Magnification Bias

The shear calibration factors $f_{ai}$ and magnification bias coefficients $q_{ai}$ are, most generally, distinct in every subset. We use the “$z\Delta z$” functional form described in the Appendix to generate the $f_{ai}$ and $q_{ai}$ from a smaller set of function parameters. Each function is specified by a polynomial function of $\Delta z$: polynomial coefficients are interpolated between grid points equally spaced in $\ln a$ at intervals $\Delta \ln a$.

We set the fiducial functions to be $f_{ai} = 0$, and $q_{ai} = q_{fid}$ independent of $(k, z)$. The priors on the polynomial coefficients are chosen to yield a chosen RMS variation of $f$ or of $q$. As detailed in the Appendix, we also specify whether the nuisance function varies mostly along the $z$ direction or along the $\Delta z$ direction of its domain.

Table 4 lists the program inputs necessary to specify the model for $f$ or $q$: their functional form, fiducial values, and priors on deviations from the fiducial.

6.4. Galaxy Correlation Coefficients

The correlation coefficient $r^q_{\ell}$ is, most generally, different at each subset and at each multipole $\ell$. We model $r^q_{\ell}$ using the “$kz\Delta z$” function form described in the Appendix. These functions are polynomial in $\Delta z$, with the polynomial coefficients

\(^1\) In practice, we use $\ln a$ and $\Delta \ln a$ to specify each subset’s nominal redshift and redshift error, but in the text we will stick with $(z, \Delta z)$ to reduce the clutter.
linearly interpolated from a grid in \((\ln k, \ln \alpha)\) space. This grid of polynomial coefficients replaces the \(r^g_{\alpha i}\) as parameters in the likelihood.

For the fiducial correlation coefficient, we interpolate smoothly between a linear and nonlinear limit according to the value of \(\Delta \ln (k, z) = k^3 P_{\text{lin}}^{m}(k, z)/2\pi^2\),

\[
    r^g_{\alpha i}(k, z) = \frac{r^g_{\alpha i}^{\text{NL}} + r^g_{\alpha i}^{\text{L}} + (r^g_{\alpha i}^{\text{NL}} - r^g_{\alpha i}^{\text{L}}) \tan^{-1} \left( \frac{\ln \Delta \ln (k, z)}{W} \right)}{2}. \tag{33}
\]

The constant \(W\) sets the width of the transition from the linear to nonlinear regime. We set \(W = 1\) unless otherwise noted.

Each polynomial coefficient at each grid point is assigned an independent Gaussian prior. These are chosen to yield a RMS prior variance that is interpolated between linear and nonlinear limits \(b^g_{\text{rms}, \text{NL}}\) and \(b^g_{\text{rms}, \text{L}}\), just as for \(r^g_{\alpha i}\).

Table 5 lists the program inputs needed to specify the galaxy bias and correlation models.

### 6.5. Galaxy Bias

We expect \(b^g\) to vary quite strongly with \(z\) as the more distant source galaxies are likely intrinsically very bright and highly biased. There may also be a strong dependence of \(b^g\) on \(\Delta \alpha\) because both \(\Delta \alpha\) and the bias may couple strongly to galaxy spectral type. Variation with \(k\) should be weaker. We hence define \(b^g\) to be the sum of two functions,

\[
    b^g = b^g_{\text{coarse}}(z, \Delta \alpha) + b^g_{\text{line}}(k, z, \Delta \alpha). \tag{34}
\]

The fiducial values are \(b^g_{\text{coarse}} = b^g_{\text{line}} = 0\).

The coarse contribution is given a very weak prior, but can only vary slowly with \(z\): \(\Delta \ln \alpha = 0.5\) by default for the \(b^g_{\text{coarse}}\) grid nodes.

The constant \(W\) is modeled using the "\(k\Delta \alpha\)" form described in the Appendix, just as for \(r^g_{\alpha i}\). The fiducial \(b^g_{\text{coarse}}\) and \(r^g_{\alpha i}\) are taken as constant over the entire domain. The RMS variations \(b^g_{\text{RMS}, \text{NL}}\) and \(r^g_{\alpha \text{RMS}}\) in the priors of these functions are also taken to be constant over the domain. We take \(b^g_{\text{RMS}} = |b^g_{\text{line}}|\) because we expect that the best constraints on IA to always arise from self-calibration of WL surveys rather than through any external modeling or prior. Roughly speaking, the IA measured by Mandelbaum et al. (2006) for the SDSS population corresponds to \(b^g \approx -0.003\) (Bridge & King 2007), which we will normally take as our fiducial model for IA. Setting \(b^g_{\text{line}} = 0\) turns off the IA systematic entirely.

Table 6 lists the program inputs needed to specify the intrinsic alignment functions.

### 6.6. Intrinsic Alignments

The strength of intrinsic alignments is specified by the \(b_{\text{ai}}^g\) and \(r_{\alpha i}^g\) values at each multipole \(\ell\). As for the galaxy density, the free-parameter count can be reduced by specifying parametric functions \(b^g_{\text{ai}}\) and \(r^g_{\alpha i}\) of \((k, z, \Delta \alpha)\) instead. Each of these two functions is modeled using the "\(k\Delta \alpha\)" form described in the Appendix.

We expect the \(b^g_{\text{ai}}\) and \(r_{\alpha i}^g\) functions to be defined on the same \((\ln k, \ln \alpha)\) grid as the galaxy bias and covariance functions.

### 6.7. Cross-correlation Coefficients

The cross-correlation coefficients \(r_{\alpha i j}^g\) are even more complex because each depends on \(k, \alpha\), plus two subsets \(\Delta \alpha_{\text{ai}}\) and \(\Delta \alpha_{\text{bi}}\). We find it infeasible to construct nuisance-function templates spanning four dimensions. We, therefore,
simplify by first writing
\[ r_{a|\beta}^{s\kappa} = r_{a|\tilde{\beta}}^{s\kappa} + s_{a|\beta}^{s\kappa} \sqrt{[1 - (r_{a|\tilde{\beta}}^{s\kappa})^2][1 - (r_{\beta|\tilde{\beta}}^{s\kappa})^2]} \]. \quad (35)

A value \( |s_{a|\beta}^{s\kappa}| \leq 1 \) is necessary (but not sufficient) to keep the mass-galaxy covariance matrix from acquiring nonphysical negative eigenvalues. In principle, the functional form of \( s_{a|\beta}^{s\kappa} \) must vary over four dimensions, but we make the gross simplification that it is constant for the survey since we expect this type of cross-correlation to have minimal effect on cosmological constraints. The program thus requires simply a fiducial scalar \( s_{a|\beta}^{s\kappa} \) and an RMS for its Gaussian prior.

The density–density cross-correlation \( \alpha\betai \) may have substantial impact on cosmological constraints, so we model it with more freedom, though not full four-dimensional behavior. We set
\[ s_{a|\betai}^{\alpha\betai} = \max \left[ 0, 1 - K_{\alpha\betai}(k, z_i) \frac{|z_a - z_{\betai}|}{2\Delta_{\max}} \right]. \quad (36) \]

This functional form for \( s \) gives the most closely related subsets the highest covariance. Here, \( \Delta_{\max} \) is the width of the redshift distribution within a set. The function \( K_{\alpha\betai}(k, z_i) \) adjusts how quickly the subsets decorrelate as their photo-

7. Tuning the Forecast Parameters

In this section, we determine the values of bin widths and nuisance-function bandwidths that are needed for reliable extraction of maximum information from lensing surveys. Unless otherwise noted, we will derive these parameters for a canonical survey with \( f_{\text{sky}} = 0.5 \); an effective source density of 60 galaxies per arcmin\(^2\) with median redshift of 1.2; \( \sigma_z = 0.24 \); and Gaussian-distributed fiducial photo-

7.1. Multipole Bin Size

Since we have excluded baryon oscillations from our transfer function, we expect to find little information in the detailed shape of the lensing or density power spectra. The broad lensing kernel in redshift also smooths away fine structure in the convergence. So we expect the information content in the Fisher matrix to be independent of the multipole bin width \( \Delta \log_{10} \ell \) below some modest value. Larger values of \( \Delta \log_{10} \ell \) reduce the complexity and execution time of the calculations, so we seek the maximum \( \Delta \log_{10} \ell \) at which nearly all the lensing information is present.

Figure 1 plots the DETF FoM of the lensing+density survey (plus spectroscopic redshift survey and Planck prior) versus the multipole bin size \( \Delta \log_{10} \ell \) for several candidate surveys. The top line is for a very optimistic survey: \( n_{\text{eff}} = 100 \text{ arcmin}^{-2} \), \( N_{\text{spec}} = 10^7 \), \( q_{\text{RMS}} = 10^{-3} \), \( b_{\text{RMS}}^s = 10^{-4} \), \( b_{\text{RMS}}^w = r_{\text{RMS}}^g = 0.01 \), and \( b_{\text{RMS}}^q = 10^{-3} \). By reducing the systematics and the shot noise to (unrealistically) low levels, we give the lensing survey the chance to extract maximal information. We find that the FoM gains only 2% for \( \Delta \log_{10} \ell < 0.3 \).

Other lines in the plot are FoM versus \( \Delta \log_{10} \ell \) for weaker surveys with \( N_{\text{spec}} = 10^4 \) and/or \( n_{\text{eff}} = 60 \text{ arcmin}^{-2} \), \( q_{\text{RMS}} = 0.1 \), \( b_{\text{RMS}}^s = 0.1 \), and \( b_{\text{RMS}}^w = -0.003 \). In these cases, we also find that the DETF FoM increases by < 2%–3% for \( \Delta \log_{10} \ell < 0.3 \). We adopt \( \Delta \log_{10} \ell = 0.3 \) for all future use.

7.2. Multipole Range

On the right-hand side of Figure 1 we plot the DETF FoM for various ranges of \( \ell \). In this study, we assume a space-based survey obtaining \( n_{\text{eff}} = 60 \text{ arcmin}^{-2} \) over \( f_{\text{sky}} = 0.5 \). The photo-

Unfortunately, our assumption of Gaussian statistics will fail by \( \ell = 10^4 \) (Cooray & Hu 2001; Lee & Pen 2008), rendering
the Fisher calculation less reliable. We will restrict our analysis to $\ell < 10^{3.5}$, but additional study of the effect of non-Gaussian statistics is clearly needed.

The flat-sky and Limber approximations will fail at low $\ell$, but the choice of $\ell_{\text{max}}$ appears less critical to the $w_0/w_a$ information content; so we will retain the $\ell > 10$ bound in our analyses.

7.3. Scale Resolution for Nuisance Functions

We require choice of node spacing $\Delta \ln k$ in the nuisance functions for the power-spectrum theory errors, the calibration errors $f$ and $q$, and the bias/correlation parameters $b_{\text{line}}, r^g, r^e, K^{gg}$, and $K^{ee}$. We set $\Delta \ln k$ to be equal for all nuisance functions, and find the value which minimizes the DETF FoM as the “most damaging” scale of variation. We examine the default case described above for several values of the shear calibration prior $f_{\text{lim}}$ and photo-z calibration size $N_{\text{spec}}$.

Figure 2 shows $\Delta \ln k \approx 0.7–1$ yields minimum information for fixed RMS priors, but the dependence is very weak. The FoM varies by only 7% over the range $0.5 < \Delta \ln k < 1.5$. We henceforth adopt $\Delta \ln k = 1$. Perhaps not surprisingly, this makes the nuisance functions have $\approx 1$ independent node in each multipole bin of $\Delta \log_{10} \ell = 0.3$.

7.4. Redshift Resolution for Nuisance Functions

All of the nuisance functions are dependent on $z$. We next investigate the redshift node spacing $\Delta \ln a$ at which the nuisance functions are most damaging to the DETF FoM. We find that the FoM is insensitive to the $\Delta \ln a$ of the power-spectrum theory errors. The value of $\Delta \ln a$ for the calibration functions $f$ and $q$ that minimizes the FoM depends upon the strength of the prior. The choice $\Delta \ln a = 0.5$ produces an FoM that is within 2% of the minimum, so we fix this value for the theory-error and calibration nuisance functions.

The redshift freedom given to the bias and intrinsic alignment nuisance functions has a strong impact on the DETF information content. Figure 2 illustrates that, at fixed RMS prior variation, models with freedom to vary on rather fine scales $\Delta \ln a = 0.1$ are most damaging to cosmological information.

8. CONCLUSION

The core of this paper are Equations (18)–(20) for the two-point correlation matrix of the lensing and density observable multipoles produced by a typical lensing survey. This was derived under a very limited set of assumptions: a homogeneous and isotropic four-dimensional metric universe with scalar perturbations; plus the weak lensing limit, the Limber and Born approximations, and an approximation that lensing magnification bias and intrinsic density fluctuations are additive. The last four assumptions could be relaxed at the expense of computational complexity. We thus hope that data analyses based on this framework could be used to constrain a wide variety of potential explanations for the acceleration phenomenon, including gravity modifications as well as new fields in the universe. In the limit of Gaussian fluctuation fields, the two-point information is a complete description of the likelihood, and hence can be used to construct Fisher matrices or analyze data. Currently configured, the analysis yields the survey’s ability to constrain the distance function $D(z)$ and linear growth function $g_\delta(z)$ without reference to particular dark-energy models. It would be straightforward to implement scale-dependent linear-growth functions.

This framework subsumes all of the information (up to two-point level) that is likely to be obtained from lensing observations: density–density, lensing–density, and lensing–lensing correlations, plus redshift distributions from unbiased spectroscopic surveys (Section 3). Furthermore, it allows for the most important expected forms of systematic error: photo-z calibration errors, shear and magnification-bias calibration errors, intrinsic alignments, and inaccuracies in power-spectrum theory. Systematics that are additive to shear (e.g., uncorrected PSF ellipticity) or to density (e.g., uncorrected foreground extinction) have not been included. We have not done so since the additive errors could, in principle, exhibit almost any arbitrary signature in the covariance matrix of the observables. Hence, a completely general model for additive errors would be degenerate with almost all other signals. For the additive systematics, it is better to determine the level at which they...
would bias the cosmological results than to attempt to fit a model. Amara & Refregier (2007) is a good example of this approach.

Since the analysis framework is independent of models for dark energy, gravity, power-spectrum evolution, or galaxy bias, we get a stripped-down look at what parameters are truly constrained by the data, and what nuisance functions must be modeled in order to extract the cosmological information. There is a substantial suite of biases and correlation functions involved in understanding the full survey data. In other work, these have been ignored or have been quantified by reference to halo occupation models (Hu & Jain 2004; Cacciato et al. 2009). Here, we introduce generic functions for bias and calibration nuisance functions that are not based on any particular physical model.

We implement one possible model for the evolution of the lensing potential power spectrum based on general relativity but allowing for failure of the growth equation. It is straightforward to implement other potential deviations from general relativity. In the current implementation, the result of the Fisher analysis is a forecast of the ability to constrain the functions $D_A(z)$ and $g_0(z)$.

Since the analysis must be discretized in redshift and angular scale in order to be feasible, we investigated the bin sizes or bandwidths of nuisance functions that should be chosen. We find that $\approx 3$ bins per decade of angular scale suffice to extract all information (apart from baryon acoustic oscillations), and that nuisance functions should be specified no finer than this. Nuisance functions for power-spectrum theory errors and shear and magnification-bias calibration errors can be specified coarsely in redshift space ($\Delta \ln a \approx 1$), but the galaxy biases, correlations, and intrinsic alignments must be modeled with potentially finer structure in redshift ($\Delta \ln a \approx 0.1$) to immunize against potential astrophysical systematics.

In future papers, we will use this framework and its implementation to investigate the requirements for spectroscopic calibration of photo-$z$s in large lensing surveys and other practical issues. As a simple first application of our framework, we have shown here that power-spectrum theory uncertainty does not significantly degrade the cosmological power of a nominal lensing survey at $10 < \ell < 10^4$. Non-Gaussian statistics are a much more important factor to consider.

The C++ code to implement Fisher forecasting using this framework runs quickly on desktop computers despite the large number of free parameters in these general models. Interested parties should contact the author for access to the code.

This work is supported by National Science Foundation grant AST-0607667, Department of Energy grant DOE-DE-FG02-95ER40893, and NASA grant BEFS-04-0014-0018. I thank Bhuvnesh Jain, Zhaoming Ma, Chris Hirata, Ravi Sheth, Hu Zhan, and the members of the Dark Energy Task Force for helpful conversations during the long gestation of this work.

APPENDIX

PARAMETRIC FUNCTIONAL FORMS FOR NUISANCE VARIABLES

In modeling an experiment, we often encounter some systematic error associated with a nuisance variable $f$ about which we have little a priori knowledge. We would like to fit some parametric form to this variable, but would like a form that is flexible enough to describe any “reasonable” behavior the function might exhibit. We also want to conveniently relate the number and prior probabilities for the parameters to the kind of variation that $f$ might exhibit. A parametric description of some nuisance function $f$ defined over a variable $x \in [-1, 1]$ would ideally have the following properties.

1. $f(x)$ has a variable number $N$ of controlling parameters $\{a_0, a_2, \ldots, a_{N-1}\}$ such that any continuous differentiable function $F(x)$ can be approximated to any desired accuracy with a sufficiently large choice of $N$.

2. We can draw $\{a_j\}$ from independent Gaussian distributions of zero mean and widths $\{\sigma_j\}$ with the result that $\text{Var}[f(x)]$
is independent of \( x \). In other words, the nuisance value \( f \) has a uniform and well determined variance when we apply a simple diagonal Gaussian prior to the parameter set \( \{a_i\} \).

A Fourier decomposition, \( f = \sum (a_j \sin j \pi x + b_j \cos j \pi x) \), exhibits these qualities, but converges poorly when \( f(-1) \neq f(+1) \).

A.1. Linearly Interpolated Functions

Another approach is linear interpolation: choosing a spacing \( \Delta x = 2/(N-1) \), we define \( a_i \) as the value of \( f \) at \( x_i = i \Delta x - 1 \). At some other \( x_i < x < x_{i+1} \), we define

\[
    f(x) = wa_i + (1-w)a_{i+1}, \quad w = \frac{x_{i+1} - x}{\Delta x}. \quad (A1)
\]

If we assign an independent Gaussian prior of width \( \sigma_a \) to each \( a_i \), then by definition we have \( \text{Var}(f(x)) = \sigma_a^2 \) if \( x \) coincides with a node. But, the variance of \( f \) is not quite homogeneous: it drops to \( \sigma[f(x)] = \sigma_a^2 / 2 \) when \( x \) is halfway between two nodes. If we want the mean variance of \( f(x) \) over the interval \( x \in [-1,1] \) to equal \( \sigma_f^2 \), then the variance of the prior on each node needs to be \( \sigma_a^2 = 3\sigma_f^2 / 2 \).

A.2. Legendre Polynomials

Polynomial expansions are also commonly used to model nuisance functions. The simplistic form \( f(x) = \sum a_i x^i \) results in extremely nonuniform variance for \( f \) with diagonal prior on \( \{a_i\} \), and hence is inappropriate for our purpose. A better choice is to expand in Legendre polynomials \( P_n(x) \), which are orthogonal over \([-1,1]\). We define

\[
    f(x) = \sum_{i=0}^{N-1} a_i P_i(\nu x). \quad (A2)
\]

Recall that the Legendre polynomials satisfy \( P_0 = 1, P_1 = x, (n+1)P_{n+1} = (2n+1)xP_n - nP_{n-1} \).

We wish to choose priors \( \{\sigma_i\} \) on \( \{a_i\} \) that cause the variance of \( f(x) \) to be as uniform as possible for \( x \in [-1,1] \). We have not found a way to attain perfect uniformity in \( x \) with polynomial interpolation; however, the following scheme gets usefully close. We discover numerically that

\[
    \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N-1} P_i^2(\nu x)(2i+1) = \frac{2}{\pi} (1 - \nu^2 x^2)^{-1/2}. \quad (A3)
\]

This implies that if we set \( \sigma_i = \sqrt{(2i+1)\pi / 2N} \), then in the limit of large \( N \) we will obtain \( \text{Var}(f(x)) \approx (1 - \nu^2 x^2)^{-1/2} \). For \( \nu = 1 \), this would diverge at the ends of our nuisance function’s interval. If, however, we choose \( \nu = 0.9 \), the RMS is only \( \approx 1.5 \) times larger at the endpoints than at \( x = 0 \). Over the \([-1,1]\) interval, the mean variance is \( (\sin^{-1} \nu) / \nu \). Hence, if we wish to have a function with \( \text{Var}(f(x)) \approx \sigma_f^2 \), we set the priors on the Legendre coefficients to be

\[
    \sigma_i = \sigma_f \sqrt{\frac{(2i+1)\pi \nu}{2N \sin^{-1} \nu}}. \quad (A4)
\]

A.3. Standard Multidimensional Functions

The Fourier, linear interpolation, and Legendre polynomial functional forms can each be extended to > 1 dimensions in a straightforward fashion. In the lens modeling, we need functions of the following dimensions:

1. comoving wavenumber or physical scale: \( x_1 = \ln k \);
2. redshift \( z \): more precisely, we will use the variable \( x_2 = \ln(1 + z) \); and
3. photometric redshift error \( \Delta z \): more precisely, our code uses the variable \( x_3 = \Delta \ln(1+z) = \ln[(1+z_\alpha)/(1+z)] \) for subset \( a_i \).

A.3.1. The \( \Delta z \) Form

For functions over the \((x_1, x_2)\) space, we use a simple two-dimensional version of interpolation between values on a rectangular grid. The power-spectrum theory error \( \delta \ln P \) uses this form as do the \( K^{(k, z)} \) and \( K^{(k, z)} \) nuisance functions. The only complication of note is that the nodal point \( a_i \) should have a prior with variance \( \sigma_a^2 = (3/2)^2 \sigma_f^2 \) if the output function is to have variance \( \sigma_f^2 \). The factor of \((3/2)^2 \) is needed to counteract the reduced variance when interpolating between grid points.

The \( \Delta z \) nuisance function is specified by the following:

1. the spacing \( \Delta x_1 = \Delta \ln k \) of the nodes for linear interpolation in \( x_1 \);
2. the spacing \( \Delta x_2 = \Delta \ln(1 + z) \) of the nodes for linear interpolation in \( x_2 \);
3. the RMS variation \( \sigma_f \) of the function allowed under the prior, which can depend on \( x_1 \) and \( x_2 \); and
4. the fiducial dependence of \( f \) on \( x_1 \) and \( x_2 \).

A.3.2. The \( \Delta z \) Form

For nuisance variables over the \((x_2, x_3)\) space (\( z \) and \( \Delta z \)), we adopt the following strategy: we choose to have variation over \( x_3 \) be described by polynomials since we usually define the range of noncatastrophic photo-z errors to be bounded to some range \([x_3] \leq \Delta_{\text{max}} \). We define the “\( \Delta z \)” functional form as follows:

\[
    f(x_2, x_3) = \sum_{i=0}^{N-1} a_i(x_2)P_i(\nu x_3 / \Delta_{\text{max}}). \quad (A5)
\]

The Legendre coefficients are in turn defined to be linearly interpolated between a series of values \( a_i \) at redshift nodes \([\ln(1+z)] \). The \( a_i \) become parameters of the model.

The fiducial and prior values for the \( i = 0 \) terms (constant in \( x_3 \)) are treated differently than the \( i > 0 \) terms. We might, for example, expect some nuisance functions to vary strongly with \( x_2 \) (nominal redshift) but only slightly with \( x_3 \) (photo-z error) at fixed \( x_2 \).

We specify the total RMS fluctuation in \( f \) allowed by the prior to be \( \sigma_f \). But, we also specify the fraction \( \text{VarFracDZ} \) of the variance that is due to dependence on \( x_3 \). If we define

\[
    f(x_2) = \int_{-\Delta_{\text{max}}}^{\Delta_{\text{max}}} dx_3 f(x_2, x_3) / 2\Delta_{\text{max}}, \quad (A6)
\]

then we aim to achieve

\[
    \text{Var}[f(x_2)] = \sigma_f^2 (1 - \text{VarFracDZ}), \quad (A7)
\]

\[
    \text{Var}[f(x_2, x_3) - f(x_2)] = \sigma_f^2 (\text{VarFracDZ}). \quad (A8)
\]
This is achieved approximately by setting the priors on the constant terms as

\[ \sigma_{ij}^2 = \frac{3}{2} \sigma_j^2 (1 - \text{VarFracDZ}) \]  \hspace{1cm} (A9)

and the \( x_3 \) dependent terms \((i > 0)\) as

\[ \sigma_{ij}^2 = \frac{3}{2} \sigma_j^2 \text{VarFracDZ} \sqrt{\frac{(2i + 1) \pi v}{2N \sin^{-1} v}}. \]  \hspace{1cm} (A10)

The RMS prior variation \( \sigma_f \) can be made a function of \( x_2 \) without loss of generality. We typically take fiducial values of constant terms as

\[ a_{ij} = 0 \]  \hspace{1cm} for \( i > 0 \) in our nuisance functions, i.e., no fiducial dependence upon \( \Delta z \).

To summarize, the \( z\Delta z \) nuisance function is specified by

1. the order \( N \) of the polynomial in \( x_3 \);
2. the maximum range \( \Delta_{\text{max}} \) of applicability in the \( x_3 \) axis;
3. the spacing \( \Delta x_2 = \Delta \ln (1 + z) \) of the nodes for linearly interpolation in \( x_2 \);
4. the RMS variation \( \sigma_f \) of the function allowed under the prior, which can depend on \( x_2 \);
5. the fraction \( \text{VarFracDZ} \) of this variance that is due to \( x_3 \) (\( \Delta z \)) dependence; and
6. the fiducial dependence of \( f \) on \( x_2 \).

A.3.3. The \( k\sigma\Delta z \) Form

The bias and correlation coefficients can, most generally, depend on scale \((x_1)\) as well as subset \((x_2, x_3)\), so we generalize to the “\( k\sigma\Delta z \)” functional form

\[ f(x_1, x_2, x_3) = \sum a_i(x_1, x_2) \mathcal{P}(x_3). \]  \hspace{1cm} (A11)

The coefficients \( a_i \) are linearly interpolated between nodes in the two-dimensional space \((x_1, x_2)\). Thus, the free parameters of this model become the Legendre coefficients \( a_{ijk} \). As for the \( \sigma_{ij} \) function, we specify the prior by the overall mean RMS variation \( \sigma_f \) (which can be a function of \( x_2 \) and \( x_3 \)), plus the \( \text{VarFracDZ} \) specifying how much of the variance is manifested as dependence on \( \Delta z \). The formulae for the priors \( a_{ij} \) on the nodal coefficients are derived exactly as in Equations (A9) and (A10), except that we now need factors of \((3/2)^2\) to account for the reduced variance when interpolating in two dimensions.

The \( k\sigma\Delta z \) function thus requires all of the specifications as the \( \sigma\Delta z \) function plus a spacing \( \Delta \ln k \) for nodes in the \( x_1 \) axis.