Measure of the impact of future dark energy experiments based on discriminating power among quintessence models

Michael Barnard, Augusta Abrahamse, Andreas Albrecht, Brandon Bozek, and Mark Yashar

Department of Physics, One Shields Avenue, University of California, Davis, California 95616, USA

We evaluate the ability of future data sets to discriminate among *different* quintessence dark energy models. This approach gives an alternative (and complementary) measure for assessing the impact of future experiments, as compared with the large body of literature that compares experiments in abstract parameter spaces (such as the well-known $w_0 - w_a$ parameters) and more recent work that evaluates the constraining power of experiments on individual parameter spaces of specific quintessence models. We use the Dark Energy Task Force (DETF) models of future data sets and compare the discriminative power of experiments designated by the DETF as stages 2, 3, and 4 (denoting increasing capabilities). Our work reveals a minimal increase in discriminating power when comparing stage 3 to stage 2, but a very striking increase in discriminating power when going to stage 4 (including the possibility of completely eliminating some quintessence models). We also see evidence that even modest improvements over DETF stage 4 (which many believe are realistic) could result in even more dramatic discriminating power among quintessence dark energy models. We develop and demonstrate the technique of using the independently measured modes of the equation of state (derived from principle component analysis) as a common parameter space in which to compare the different quintessence models, and we argue that this technique is a powerful one. We use the PNGB, Exponential, Albrecht-Skordis, and Inverse Tracker (or inverse power law) quintessence models for this work. One of our main results is that the goal of discriminating among these models sets a concrete measure on the capabilities of future dark energy experiments. Experiments have to be somewhat better than DETF stage 4 simulated experiments to fully meet this goal.

I. INTRODUCTION

Over the last decade or so there has been mounting evidence that the energy of the universe has a large accelerating component, dubbed "dark energy"[1]. As the evidence becomes more convincing, there is growing enthusiasm for launching a major program to collect additional data that will help us better understand the nature of the dark energy [2], and indeed considerable progress is being made on this front.

To plan a strong program of dark energy studies, one needs to assess the relative impact of different possible experiments. This has most often been done by describing the dark energy in some abstract parameter space and calculating how much a given data set could constrain those abstract parameters. For example, the Dark Energy Task Force (DETF) [3], building on earlier work [4,5], used a standard two parameter model of the dark energy equation of state w as a function of cosmic scale factor a given by $w(a) = w_0 + w_a(1 - a)$ to form a figure of merit based on constraining power in the $w_0 - w_a$ parameter space. For the most part, other authors have used other abstract dark energy parametrizations [3,6,7], but more recently, we have extensively explored the impact of future experiments using the constraints produced on the actual parameters of scalar field dark energy models [8–11]. That work gives

another window on the power of future experiments, which we have argued is largely consistent with the DETF results in the $w_0 - w_a$ parameter space.

Our recent work [8–11] shows the constraining power of future experiments on specific dark energy models. However, because the natural parameter space of each quintessence model is very different from the others, we were not able to use our techniques to directly evaluate the ability of experiments to favor one dark energy model strongly over another. The one exception to this is "cosmological constant" dark energy [which has $w(a) = -1$]. Each quintessence model we considered had a part of parameter space where the quintessence closely mimicked a cosmological constant, and we used that fact systematically in [8–11] to consider discriminating power in the quintessence vs cosmological constant domain.

This paper builds on that earlier work to make a more comprehensive analysis of the ability of future data to discriminate among different quintessence models. The key new ingredient is the use of a specially chosen parameter space to represent the different quintessence models in a common and comparable form. To this end we use the "independently measured modes" of $w(a)$, which have long been appreciated for a variety of reasons [5,7,12–17]. These modes, which are different for each experiment, represent the modes or "moments" of $w(a)$ of which

uncorrelated measurements are made by that particular experiment. This feature allows us to identify the modes which are best measured and analyze them in a straightforward way (due to the lack of correlations). These modes comprise a basis which spans the space of possible functions $w(a)$. A given quintessence model with specific fixed parameters will give a specific function $w(a)$ which can then be expanded in the modes, and the expansion coefficients form the parameter space in which we work.

We use the pseudo-Nambu-Goldstone boson (PNGB) [9,18], Exponential [10,19], Albrecht-Skordis (AS) [8,20], and Inverse Tracker (IT, or inverse power law) [11,21] quintessence models for this work. This is a diverse set of models, each of which holds its own special interest among researchers (see our discussion in the introductions of [8–11] for a brief review of the motivations and [22] to place these models in a more general context). One of our results is that these four quintessence models actually occupy very different regions of the mode parameter space. This tells us that the aspects of $w(a)$ that are well measured by realistic experiments have the potential to be extremely useful in discriminating among quintessence models. How much this potential is realized is of course related to the resolution achieved by a given experiment in its mode parameter space, and that issue is the subject of much of our analysis.

One of our key results is that the goal of discriminating among these four models sets a concrete measure on the capabilities of future dark energy experiments. Experiments have to be somewhat better than DETF stage 4 simulated experiments to fully meet this goal.

Sections II, III, and IV describe our methods and Sec. V presents our detailed results, while Sec. VI provides a discussion of the relevant statistical issues. Our conclusion, Sec. VII, summarizes our results.

II. CONNECTION TO OUR EARLIER WORK

Our work builds very directly on our recent papers studying specific quintessence models [8–11]. We refer the reader to those papers to learn more about our methods. (An appendix giving the technical details about our methods that are common to all these papers can be found in [9].) One product of this earlier work is a set of Monte Carlo Markov chains (MCMC) representing the distribution of models that are consistent with specific stage 2 simulated data that are chosen to be consistent with a cosmological constant cosmology. Specifically, these chains represent the distribution of possible scalar field parameters that are consistent with a specific simulated stage 2 data set. Each quintessence model has its own scalar field parameters and its own chain representing the distribution in that space. Also, in each case we base the stage 2 data around a specific "fiducial" set of scalar field parameters that are consistent with a cosmological constant at the 1σ–2σ level. Using these different fiducial models

accounts, in a rough way, for uncertainties in the outcomes of the stage 2 experiments.

We use these chains for the work reported here by determining the equation of state function $w(a)$ for each point on the chain and projecting that function into the eigenmode-based space corresponding to a particular simulated data set from stage 3 or stage 4 (as discussed in detail below). When data from different scalar field models are analyzed using the same eigenmodes, those modes provide a common parameter space in which scalar field models with different "fundamental" scalar field parameters can be compared on a common footing. In this usage the full chain represents how scalar field models that are consistent with each other at stage 2 would be distributed in the eigenmode-based space based on data from stage 3 or 4. Any discriminating power among regions of the stage 2 chains enabled by the higher stages represents progress over stage 2 data, on which the original chains are based.

III. GENERATING EIGENMODES

For this work, eigenmodes were generated using MCMC calculations. We do this by breaking up the equation of state into nine bins linear in scale factor from $a = 0.2$ to $a = 1$, and using the value of $w + 1$ in each of these bins as
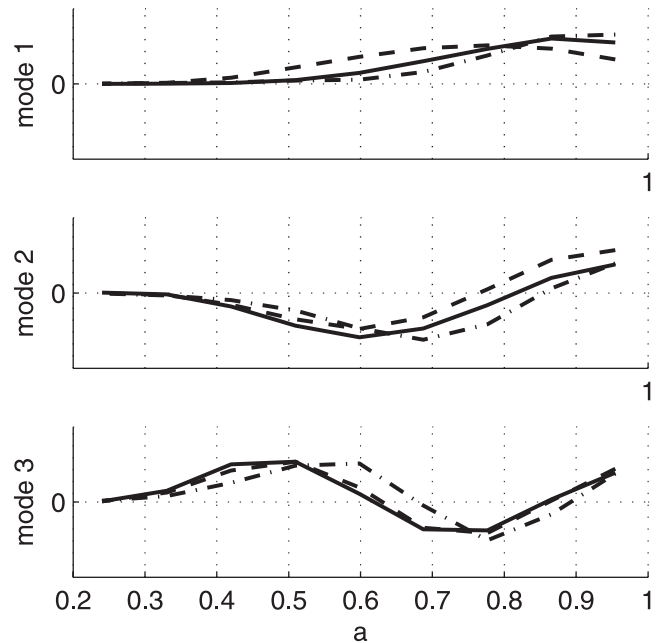

FIG. 1. The first three best-measured eigenmodes of $w(a)$. The stage 3 (dot-dashed line), stage 4 ground (dashed line), and stage 4 space (solid line) modes share a common general form at each mode level. Specific differences can be related to various differences among the experiments, including how deeply a given experiment probes in redshift. (Technically, these modes should each be represented by nine discrete bin values. The connecting lines guide the eye, and reflect a likely "continuum limit" as discussed in [7].)

parameters. We then run Markov chains with these ''bin'' parameters in addition to cosmological parameters in order to calculate a covariance matrix, from which we extract the nine-by-nine covariance matrix for said bin parameters. The eigenvectors of this matrix give us our eigenmodes, while the eigenvalues are the squares of the uncorrelated error in each mode. While these methods are slightly different from the Fisher matrix techniques of [7], the results are consistent, and our choice of binning is driven by the analysis in [7]. As with our previous work [8–11], the simulated data sets are constructed in a manner equivalent to the DETF simulated data. We do not include cluster data (due to the technical difficulty of including them discussed in [7]). There are a number of possible considerations beyond the DETF work (such as considering more carefully the impact of including cross correlations among different photometric data types [23,24]) that many expect will lead to significant improvements over the DETF projections. For this paper we use the original DETF data models for ease of comparison, except briefly in Sec. V where we consider a simpleminded extension.

Plots of the first three eigenmodes are given in Fig. 1 (ranked by how well each mode is measured). All nine eigenmodes together form an orthonormal basis, which is different for each data set. The modes pick up additional oscillations as one goes from best-measured (mode 1) to less well-measured modes.

IV. PROJECTION ONTO EIGENMODES

We use the above eigenmodes to analyze four quintessence models of dark energy: the PNGB, Exponential, AS, and IT models. Sample equation-of-state behavior of these models is illustrated in Fig. 2. In each case, we will use points pulled from MCMC as representations of each model's parameter space. The chains were run on DETF stage 2 type data generated using a fiducial point in each model, as in [8–10]; this gives us a fairly wide spread of parameter space for each model and represents, in a rough way, the uncertainties in the outcomes of the stage 2 experiments.

We use a simple algorithm to average the equation of state $w(a)$ in the nine scale factor bins, w_j. These w_j can be mapped into mode projections m_i by matrix multiplication:

$$\sum_j E_{ij}(w_j + 1) = m_i \qquad (1)$$

where E_{ij} is the jth term of the ith eigenmode. We use $w_j + 1$ to center the eigenmodes around $w(a) = -1$, so that $m_i = 0$ for all i is a cosmological constant cosmology. The actual value of m_i for a given $w(a)$ depends on the number of bins used, so it is more convenient to look at $\frac{m_i}{\sigma_i}$, where σ_i is the square root of the ith covariance matrix eigenvalue. This expresses the power in each eigenmode in units of its uncorrelated error, and should be relatively
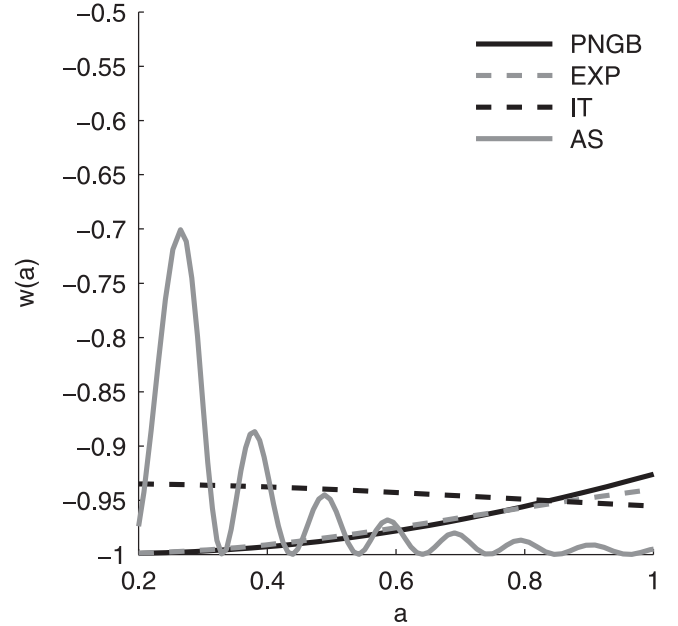

FIG. 2. A characteristic $w(a)$ function for each of the four scalar field models considered in this paper.

stable as we change the number of modes by, for example, refining the bin size.

When we consider how to display the range of quintessence models in the eigenmode space, graphing the power in the first three modes in a rotatable three-dimensional plot can be fascinating. However, as this does not lend itself to the static two dimensions of a paper, it is more enlightening to view the modes two at a time (an interested reader may contact the authors of this paper to view the three-dimensional versions).

Figure 3 shows stage 2 distributions of our four example quintessence models represented in the first four modes of combined stage 3 data (using photometric data and systematics designated ''optimistic'' by the DETF, just as was done in [8–11]). Because these points are plotted in units of the error of each mode, one can get an intuitive idea of the stage 3 resolving power by noting that stage 3 data should roughly resolve areas of unit size in these plots.

It is interesting to note how the models examined here occupy very distinct portions of the ''mode space'' except for where they meet at the origin [$w(a) = -1$]. We should mention here that these figures in many cases display only a subset of the total space accessible by the models, because the parameter spaces of the models were restricted in the MCMC. This is particularly true of the AS model, for which we expect that the kinks and wiggles in the displayed distribution will characterize a broader distribution of such features. This may also be an issue for the PNGB model, which was restricted in the MCMC analysis to the concave-down portion of its potential. The Exponential model, however, has a concave-up potential and gives us some idea of where the concave-up portion of the PNGB
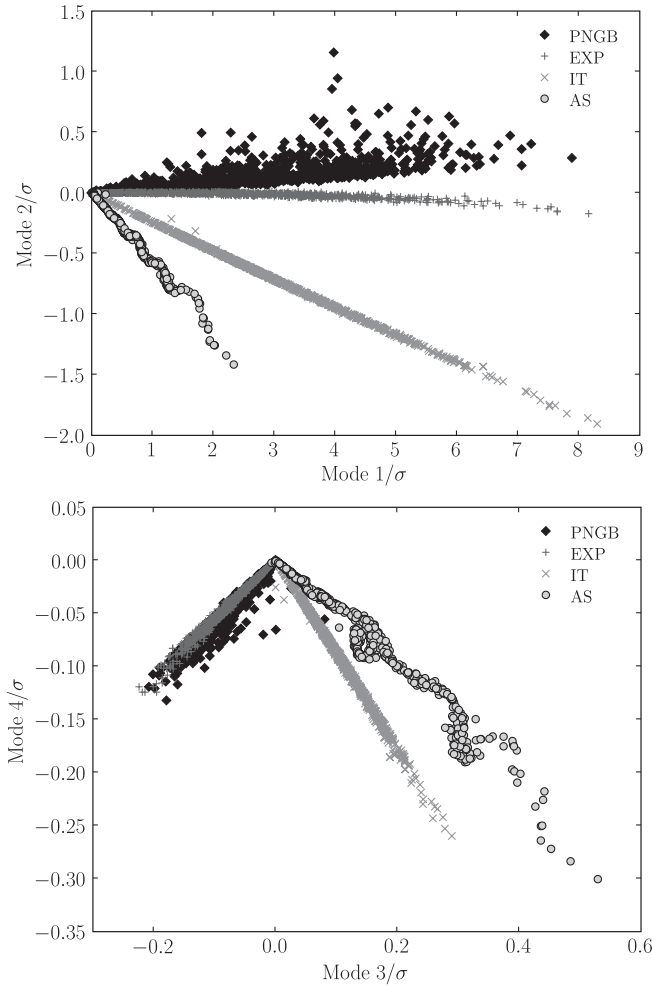
scalar field models, but only in extreme cases will it help us distinguish between them (noting that a very large value would actually rule out the AS model). It is clear that it is the higher modes that will distinguish between models, though the plot shows that stage 3 data will not do this very well.

The authors of [6] also made plots using similar modes to ours, but their approach is different. In [6] the potentials are drawn from an abstract continuous space that covers large ranges of possibilities, whereas our potentials are drawn from specific scalar field models which each only have certain classes of behaviors even as the parameters are varied fully. These differences have allowed us to discover a much more striking structure than can be seen in [6].

In Figs. 4 and 5 we show the mode projection plots for stage 4 space and ground data sets, respectively (using the
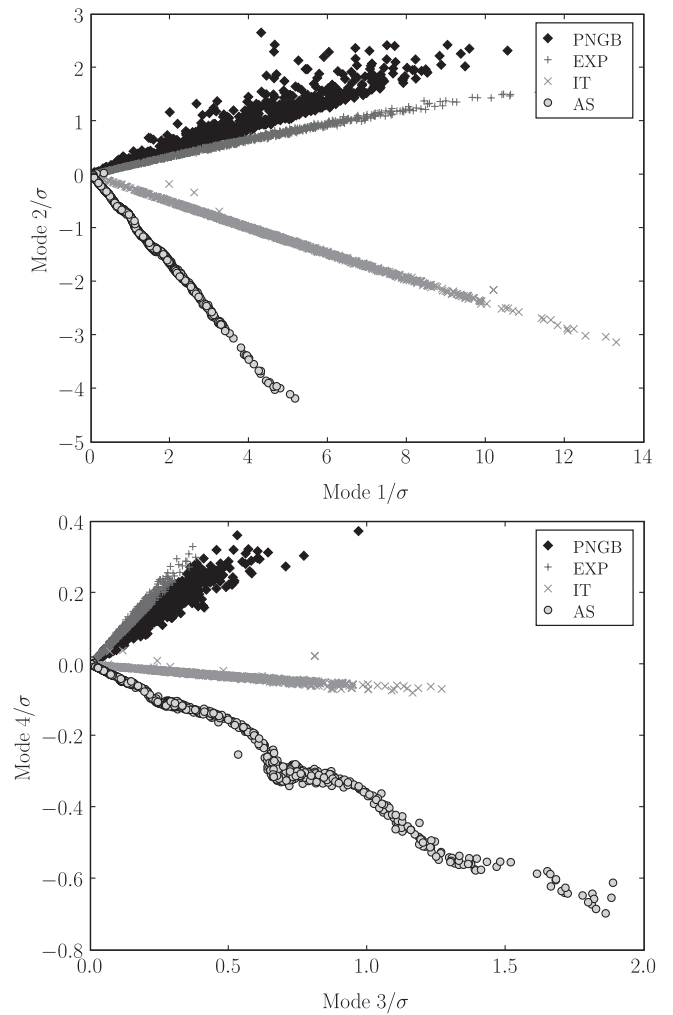

FIG. 3.   *Stage 3 photometric:* A plot of the projections of the PNGB, Exponential, AS, and Inverse Tracker models onto the first four stage 3 photometric eigenmodes. The displayed points are a sampling from MCMC for each model, and the scale of each axis is given by σ_i, the uncertainty in measurements of that mode. Note that, while a significant portion of each model has a first mode projection larger than 1σ, very few points have a second mode projection of 1σ or larger.

model may lie in the mode space. As discussed in [8–11], most of these restrictions were required to eliminate parameters that would be completely unconstrained by even the best future data.

On a similar note, there is a small fraction of the points ($< 1\%$) in the Inverse Tracker model that never display tracking behavior, but instead display the thawing [25] behavior of the PNGB and Exponential models; this can be seen in the plots as a handful of outlying points above the main concentration of Inverse Tracker model points.

It is of interest that, as we look at parametrizations of increasing distinction from the cosmological constant, we see a consistent increase in the amplitude of their first mode projections. As such, we might look at the first mode projection as a signal of the presence of one of these

FIG. 4.   *Stage 4 ground:* A plot of the projections of the PNGB, Exponential, AS, and Inverse Tracker models onto the first four stage 4 ground eigenmodes. While the first mode is measured better by this data than by the stage 3 data, it is the improvement in the measurement of the second mode that makes the models distinguishable over a wide range of their parameters.
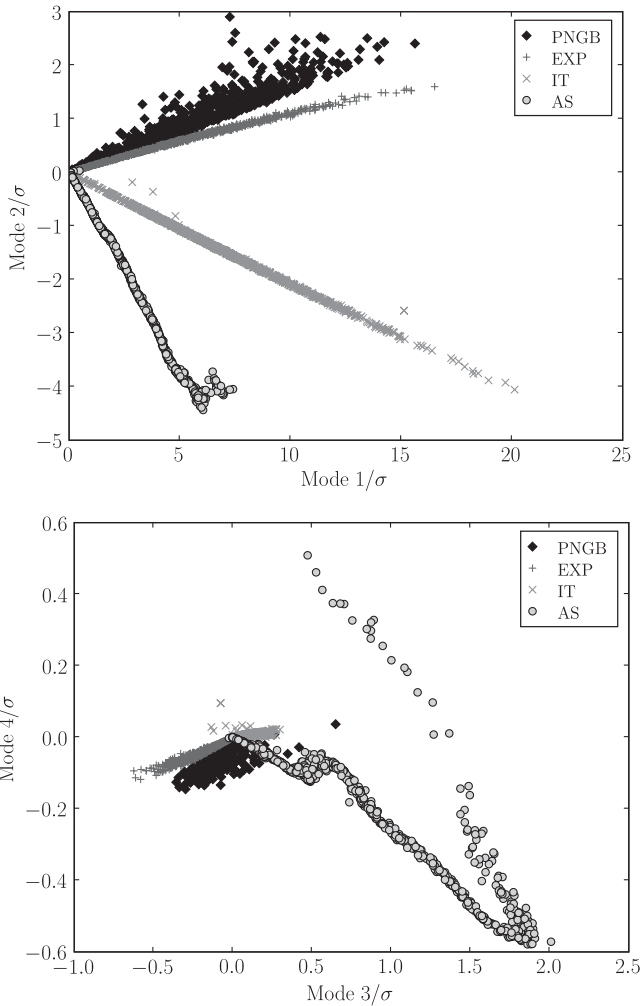
FIG. 5. *Stage 4 space:* A plot of the projections of the PNGB, Exponential, AS, and Inverse Tracker models onto the first four stage 4 space eigenmodes. As with the stage 4 ground plot, the most significant improvement, in terms of model distinction, going from the stage 3 data to the stage 4 space data is the measurement of the second mode.

same DETF optimistic combinations used in [8–11]). We can see from these plots that the overall shape of the distribution of projected models remains very similar in the first two modes. The stage 4 data sets do a better job of measuring the first mode than stage 3, a much better job of measuring the second mode, and even begin to give resolution of the third mode that is of marginal value in resolving these models. As a result, for a scenario in which the stage 3 projects detected the presence of quintessence at the level of a few sigma, the stage 4 data should be able to discriminate between this sample of models to at least that level.

V. RESOLVING THE QUINTESSENCE MODELS

With the models projected onto the eigenmode space, we then have a common space in which to compare the

different scalar field models and evaluate how well upcoming experiments will distinguish among them. We can use mode uncertainties as a metric to find how measurably different each quintessence model will be. Informed by the plots of mode projections, we can reasonably expect that the first mode will dominate this measure, but the second mode (and, to some extent, the third) will play a part in setting the minimum separation. As for the higher modes, we can expect to see these swamped by their uncertainty; a combination of the poor measurement and the lack of power in these modes by our quintessence models makes them largely irrelevant. Figure 6 gives information about the distribution of mode amplitudes for our chains. From Fig. 6 one can see that the scatter in the higher modes will be much larger than the distribution of physically interesting values. This gives us, in effect, a prior on the value of these modes that is much stronger than the data constraints from even stage 4 experiments. Therefore, this calculation will actually be done using only the first three eigenmodes.

Assuming stage 2 data as discussed above, higher stage data might come from a universe based on any of the quintessence models represented by points in our mode-space plots. Our next step is to scan all possible data outcomes and evaluate the potential discriminating power. We represent the stage 2 distribution of each quintessence model, by choosing four "test points" spread evenly across
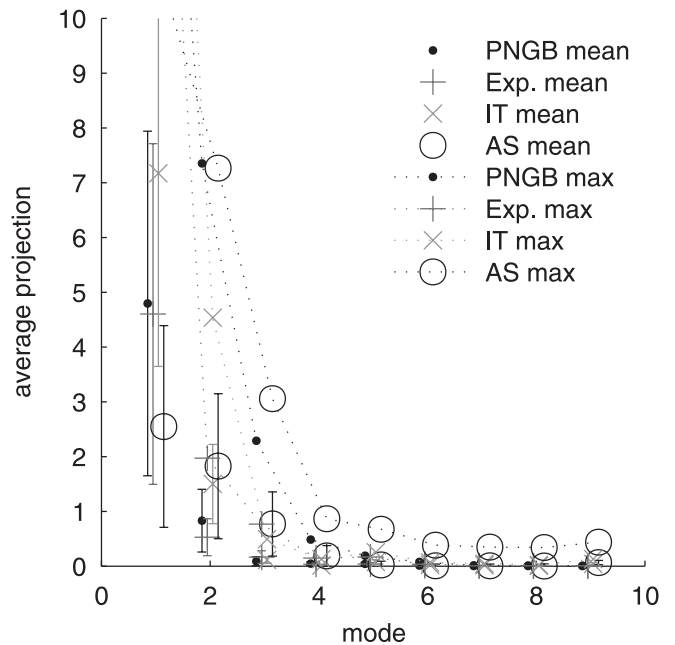

FIG. 6. A plot of the mean (with bars at 1 standard deviation) and maximum projections of the examined models into the stage 4 space mode space, in terms of the mode uncertainties. It can be seen that for nearly all of the model points, the projections onto the modes higher than 3 are negligible. Also, though the AS model has outlying points that have some power in the higher modes, the bulk of the model does not.

the mode space. Four test points each for four quintessence models gives a total of 16 "data outcomes" which are meant to represent (in a rough way) the full range of possibilities. The distribution of our test points in mode space is shown in Fig. 7. These "test points" were chosen by their first mode projections, which places them in sequence along the very nearly linear regions the models cover [26]. For each of the 16 test points, we analyze the ability of data centered on the test point to exclude other points on the chains. Note that in Fig. 7 the test points are not represented with noise, which is expected in any data set (at a level given by the σ_i's) and which is reflected in our quantitative analysis.

The curves in Figs. 8–10 show the fraction of model points with χ^2 less than a specific value, given on the x axis [27]. To read these plots, start with the labels on the left-hand side. These read PNGB, EXP (Exponential), IT, and

AS, and label which model the "test points" used in that row of plots were pulled from. Then, look to the top of the figure, where the same model labels mark the columns; these label which model the test points are being compared to in each column. In each plot, four functions are graphed: χ^2 vs the fraction of the compared model's points that have that χ^2 or lower relative to the test point in question. For example, looking at the PNGB vs PNGB plot, the curves represent how far the rest of the PNGB model is from each PNGB test point. The point where each function touches the x axis (the minimum χ^2) is, in this case, a loose measure of how densely the MCMC populates the mode space at that test point. However, if one were to look at the PNGB row and the AS column, the graph there represents how far the PNGB test points are from all of the points in the AS chain, and the minimum χ^2 of each function represents by what χ^2 each of the PNGB test points would rule out the entire AS model.

The sharp left-hand edges of curves in Figs. 8–10 imply very strong discriminating power at the level of χ^2 given by the point on the x axis where the edge is located. We have organized information about this important feature in Tables I, II, and III, which give numerical results for the minimum χ^2 that one could estimate from Figs. 8–10. In general (due to the outlying points and other factors), one expects some low level tails even on the otherwise sharp rising edges. In order to make sure we are quantifying the true edge of the figure, we ignore the closest 1% of points from the compared model. This is equivalent to finding the χ^2 where each plot crosses the 0.01 fraction mark. For completeness, we have again included the Exponential model, though we do not expect any experiment to distinguish between it and the PNGB model. The numbers reported here reinforce the intuition we gain from the mode plots: only in extreme cases will stage 3 distinguish between the models. The only examples that rise above the level of 99% confidence (which is a χ^2 of 11.36 for the three parameters used here) are due to the previously mentioned observation that the AS model has a much smaller range of amplitudes in the first mode than the other three models.

The stage 4 ground and stage 4 space χ^2 values show significant improvement over stage 3. Again, the largest χ^2 values come from the large first mode separation between the last test point for each model and the nearest AS model point.

Looking at Tables II and III it appears that the stage 4 data lie on some kind of threshold: There are quite a few entries greater than 11.36 (indicating exclusion at 99% confidence) but plenty that are lower. To explore the nature of this threshold a bit more, Table IV presents χ^2 values for a hypothetical data set that would improve the σ_i by a factor of 1.5 over stage 4 space. Because of a variety of considerations (including those discussed in [23,24]), many believe that such improvements (or even much better
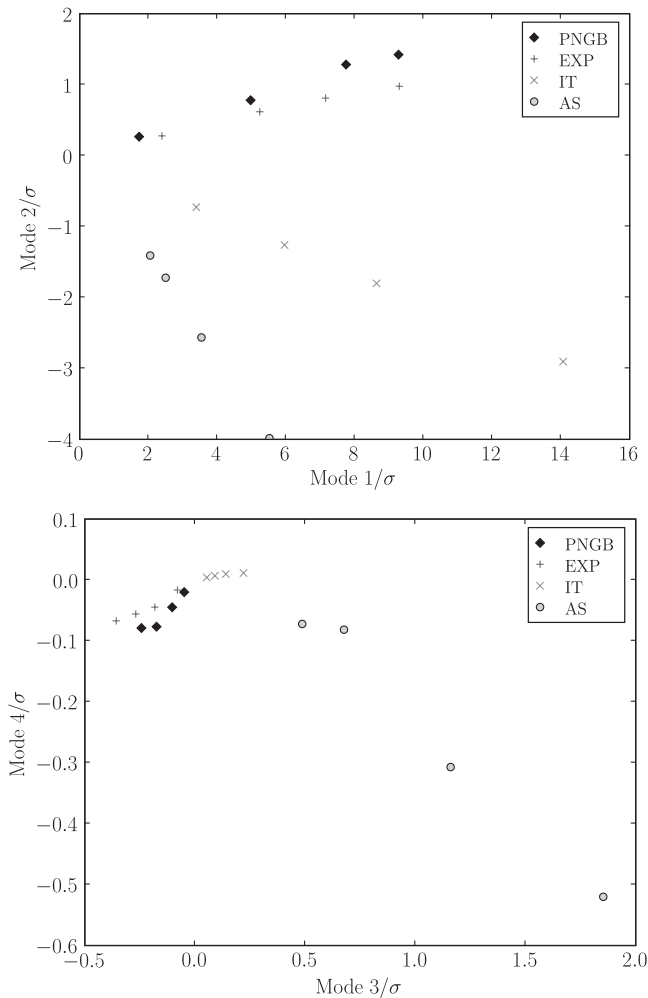

FIG. 7. The four test points for each model (16 test points in all) represent possible experimental outcomes. The test points are plotted here in the eigenmodes for stage 4 space. One can compare this figure with Fig. 5 to see the degree to which the full range of behavior of each model is represented by the test points.
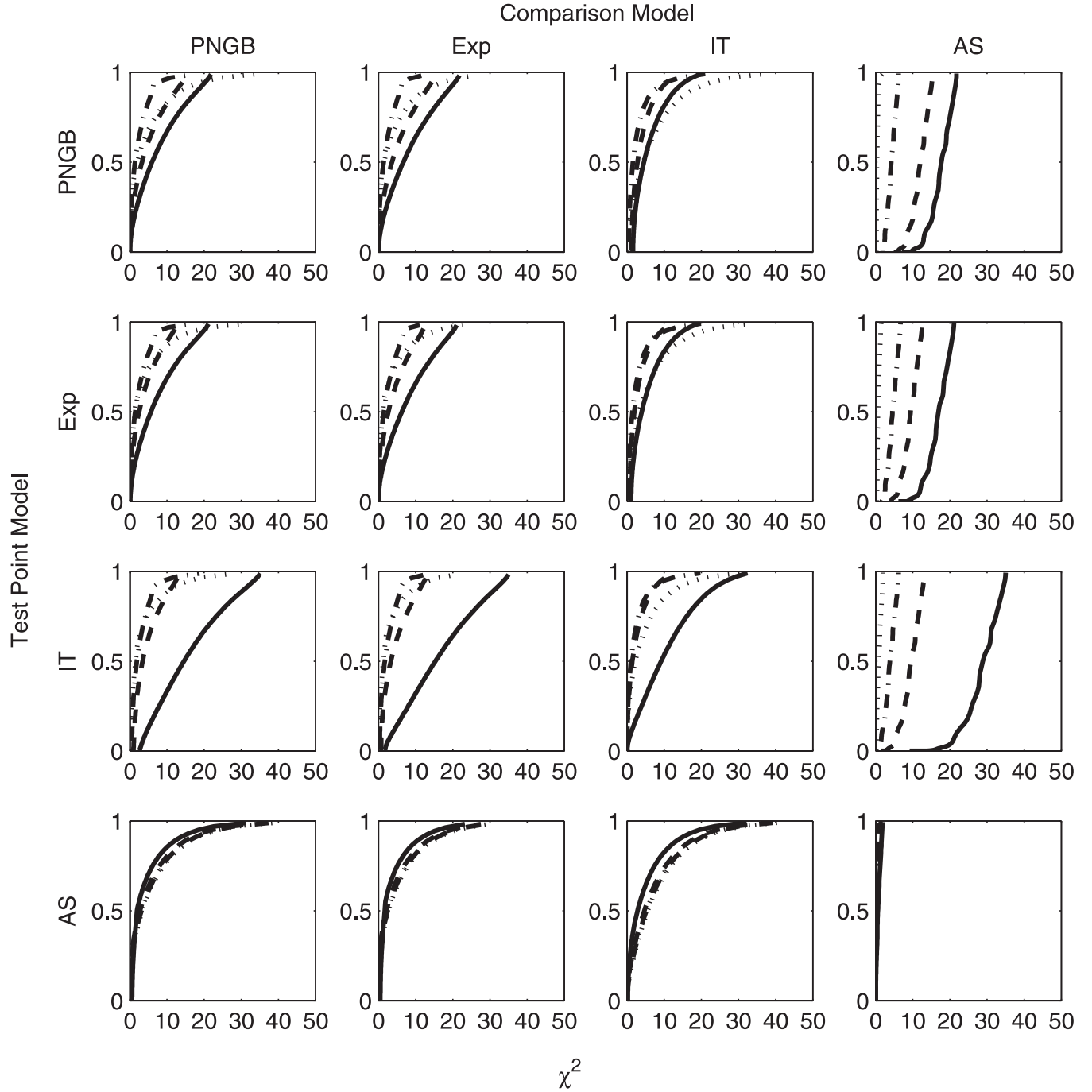
FIG. 8. *Stage 3 photometric*: These plots show the distributions of χ^2 as each test point is compared with each comparison model. Specifically, we plot the fraction of χ^2 values less than χ^2 given on the x axis. The rows of this figure correspond to which model the test points have been pulled from (as seen in Fig. 7), while the columns correspond to the model which is compared to those test points. In each plot, the test point closest to the origin of the mode space (and, thus, a ΛCDM model) is denoted by a dotted line, the next closest as a dot-dashed line, the next as a dashed line, and the farthest as a solid line. The 99% confidence interval for three parameters is $\chi^2 = 11.36$. The relatively sharp left-hand edges of these curves are an interesting feature (related to the gaps between models in mode space) which is discussed in the text and Table I.

ones) over the DETF modeling of stage 4 are realistic for some experiments. In Table IV there are indeed a great many more entries greater than 11.36, further supporting the notion of a threshold around stage 4. As a reference point, the σ_i of the stage 4 data sets are a factor of about 4 or 5 smaller than the stage 3 σ_i for most modes. To get a more complete understanding of which experiments might achieve (or beat) the level of model discrimination shown
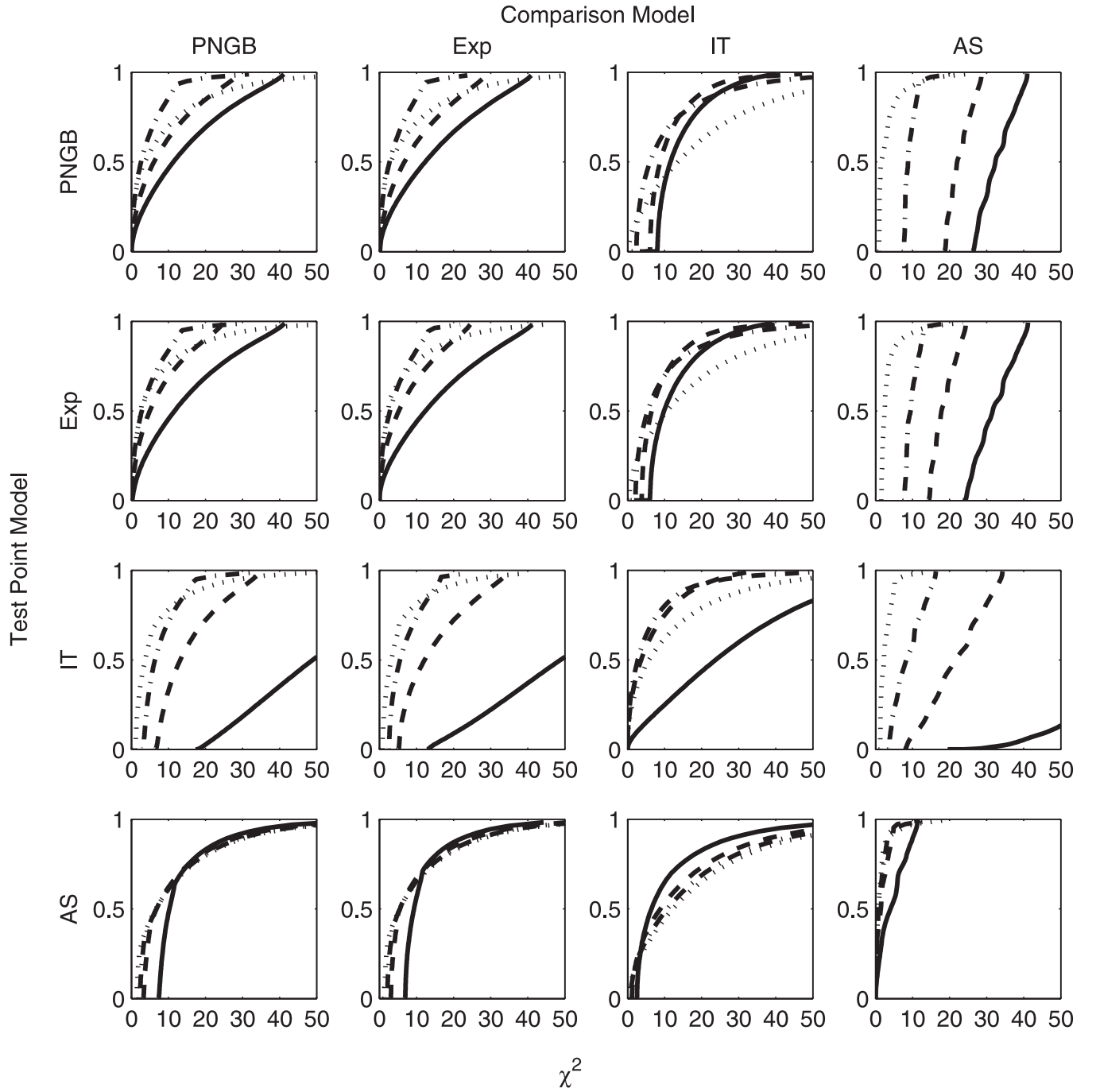
FIG. 9.   *Stage 4 ground*: These plots show the distributions of χ^2 as each test point is compared with each comparison model. Specifically, we plot the fraction of χ^2 values less than χ^2 given on the x axis. The rows of this figure correspond to which model the test points have been pulled from (as seen in Fig. 7), while the columns correspond to the model which is compared to those test points. In each plot, the test point closest to the origin of the mode space is denoted by a dotted line, the next closest as a dot-dashed line, the next as a dashed line, and the farthest as a solid line. The 99% confidence interval for three parameters is $\chi^2 = 11.36$. The relatively sharp left-hand edges of these curves are an interesting feature (related to the gaps between models in mode space) which is discussed in the text and Table II.

in Table IV, one would have to undertake detailed modeling of alternative experiments and data analysis schemes. Such a systematic analysis is not the subject of this paper, and we leave that for future work.

VI. DISCUSSION OF STATISTICS

The MCMC of the selected scalar field models in this analysis are used as a means to represent the phenomenol-
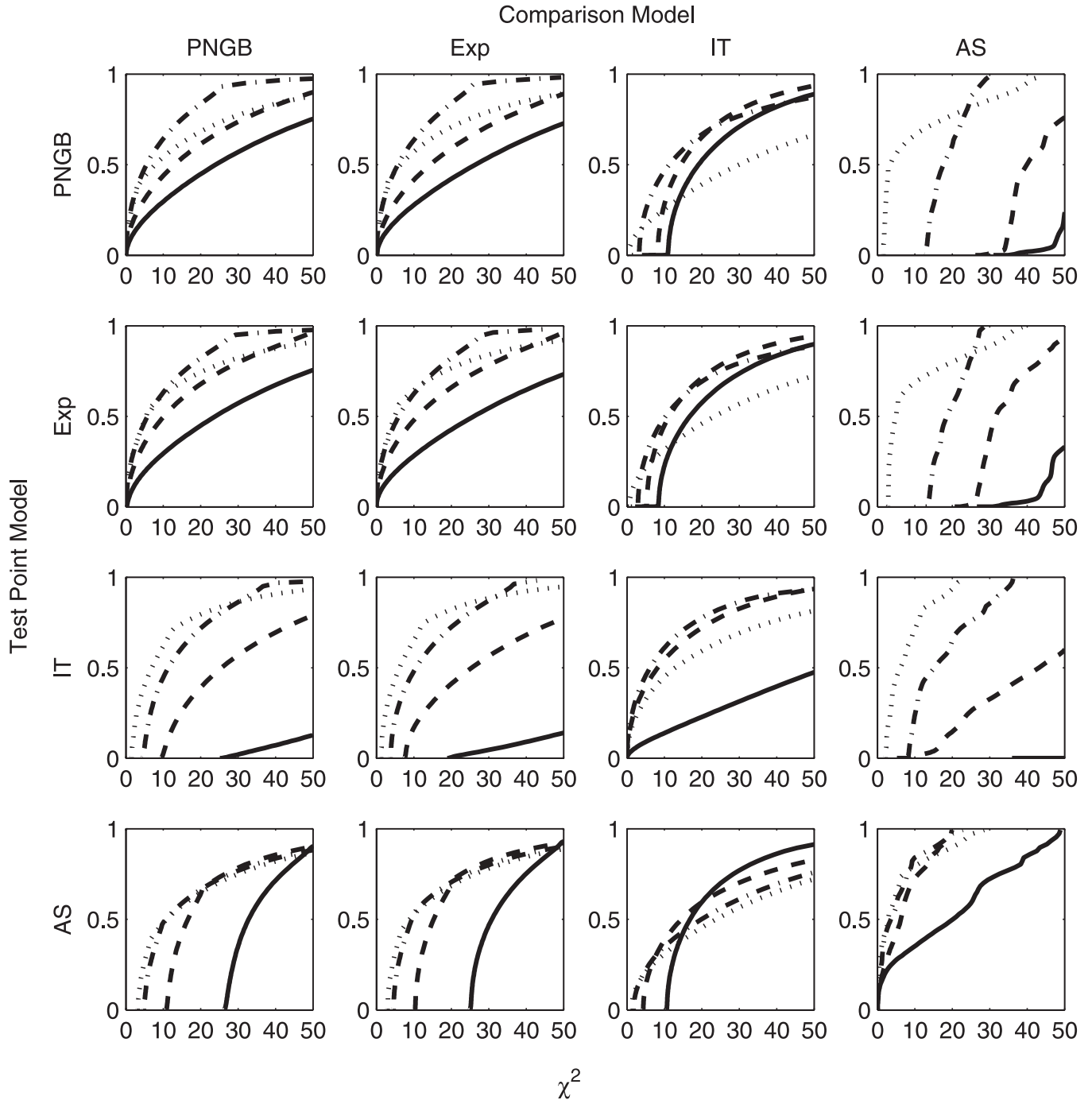
FIG. 10.   *Stage 4 space*: These plots show the distributions of χ^2 as each test point is compared with each comparison model. Specifically, we plot the fraction of χ^2 values less than χ^2 given on the x axis. The rows of this figure correspond to which model the test points have been pulled from (as seen in Fig. 7), while the columns correspond to the model which is compared to those test points. In each plot, the test point closest to the origin of the mode space is denoted by a dotted line, the next closest as a dot-dashed line, the next as a dashed line, and the farthest as a solid line. The 99% confidence interval for three parameters is $\chi^2 = 11.36$. The relatively sharp left-hand edges of these curves are an interesting feature (related to the gaps between models in mode space) which is discussed in the text and Table III.

ogy associated with each model, that is, to represent the full spread of points in mode space that the model can occupy. This is not to be confused with the set of priors that would be used in a Bayesian analysis. The question we

seek to answer with Figs. 8–10, and Tables I, II, III, and IV, is "If the universe is a single realization of a particular model, how well will a given experiment rule out other models/phenomenologies?" The method we use to answer

TABLE I. Stage 3 photometric. This is a table of the minimum χ^2 (ignoring the smallest 1% of χ^2 values). The 99% confidence level is a χ^2 of 11.36. Notably, the only comparisons that rise to that level are between the test points with large first mode projections and the AS model, telling us that this is mostly a first mode measurement. As discussed in the text, this table gives the values of χ^2 where the curves in Fig. 8 sharply approach the x axis.

PNGB	PNGB	Exponential	IT	AS
Point 1	0.001	0.001	0.1	0.2
Point 2	0.002	0.01	0.5	1.8
Point 3	0.004	0.04	1.2	6.2
Point 4	0.01	0.04	1.6	10.0
Exp.	PNGB	Exp.	IT	AS
Point 1	0.004	0.001	0.1	0.4
Point 2	0.01	0.001	0.4	1.8
Point 3	0.03	0.001	0.7	4.3
point 4	0.1	0.01	1.1	9.1
IT	PNGB	Exp.	IT	AS
Point 1	0.2	0.1	0.001	0.2
Point 2	0.5	0.4	0.0004	0.7
Point 3	1.0	0.7	0.001	3.3
Point 4	2.7	1.8	0.01	16.4
AS	PNGB	Exp.	IT	AS
Point 1	0.1	0.1	0.1	0.0001
Point 2	0.2	0.1	0.1	0.0001
Point 3	0.2	0.2	0.1	0.0002
Point 4	0.6	0.5	0.2	0.001

TABLE III. Stage 4 space. This is a table of the minimum χ^2 (ignoring the smallest 1% of χ^2 values). Again, the low first mode projections for the AS model are responsible for the highest χ^2 values, but we can also see significant separation for test points that have first mode projections in the range of the model they are being compared to. As discussed in the text, this table gives the values of χ^2 where the curves in Fig. 10 sharply approach the x axis.

PNGB	PNGB	Exponential	IT	AS
Point 1	0.01	0.01	0.4	1.6
Point 2	0.01	0.05	3.2	13.0
Point 3	0.02	0.2	8.2	30.0
Point 4	0.04	0.2	10.9	37.4
Exp.	PNGB	Exp.	IT	AS
Point 1	0.02	0.002	0.6	2.8
Point 2	0.05	0.003	2.9	13.6
Point 3	0.1	0.01	5.2	24.5
Point 4	0.3	0.02	8.4	33.2
IT	PNGB	Exp.	IT	AS
Point 1	1.5	1.3	0.005	2.2
Point 2	4.6	3.8	0.002	8.2
Point 3	9.7	7.7	0.003	9.4
Point 4	27.8	20.8	0.1	57.3
AS	PNGB	Exp.	IT	AS
Point 1	3.2	3.0	1.1	0.002
Point 2	4.9	4.6	1.8	0.003
Point 3	10.9	10.4	4.3	0.01
Point 4	26.5	25.1	10.6	0.01

TABLE II. Stage 4 ground. This is a table of the minimum χ^2 (ignoring the smallest 1% of χ^2 values). Again, the low first mode projections for the AS model are responsible for the highest χ^2 values, but we can also see significant separation for test points that have first mode projections in the range of the model they are being compared to. As discussed in the text, this table gives the values of χ^2 where the curves in Fig. 9 sharply approach the x axis.

PNGB	PNGB	Exponential	IT	AS
Point 1	0.003	0.005	0.3	0.9
Point 2	0.004	0.04	2.4	7.6
Point 3	0.01	0.2	6.0	18.8
Point 4	0.03	0.2	8.0	26.5
Exp.	PNGB	Exp.	IT	AS
Point 1	0.01	0.001	0.4	1.6
Point 2	0.04	0.002	2.1	7.8
Point 3	0.1	0.003	3.8	14.5
Point 4	0.3	0.01	6.0	24.4
IT	PNGB	Exp.	IT	AS
Point 1	1.1	0.9	0.002	1.2
Point 2	3.2	2.6	0.001	3.6
Point 3	6.7	5.2	0.002	8.3
Point 4	18.7	13.6	0.04	30.1
AS	PNGB	Exp.	IT	AS
Point 1	1.4	1.3	0.5	0.001
Point 2	2.3	2.1	0.8	0.001
Point 3	3.3	3.1	1.2	0.001
Point 4	7.4	7.0	2.6	0.001

TABLE IV. Stage 4 space, with experimental uncertainty reduced by 2/3 in each mode. This is a table of the minimum χ^2 (ignoring the smallest 1% of χ^2 values).

PNGB	PNGB	Exponential	IT	AS
Point 1	0.01	0.01	0.9	3.6
Point 2	0.01	0.1	7.3	29.1
Point 3	0.04	0.4	18.4	67.5
Point 4	0.09	0.4	24.1	84.1
Exp.	PNGB	Exp.	IT	AS
Point 1	0.04	0.01	1.4	6.4
Point 2	0.1	0.01	6.6	30.7
Point 3	0.3	0.01	11.8	55.1
Point 4	0.7	0.05	18.8	74.6
IT	PNGB	Exp.	IT	AS
Point 1	3.5	2.9	0.01	4.9
Point 2	10.4	8.5	0.01	18.4
Point 3	21.9	17.4	0.01	21.1
Point 4	62.4	46.9	0.2	129.0
AS	PNGB	Exp.	IT	AS
Point 1	7.2	6.8	2.5	0.004
Point 2	10.9	10.3	4.0	0.01
Point 3	24.6	23.3	9.8	0.01
Point 4	59.7	56.6	23.9	0.01

this is to take the probability of observing the observables generated by that realization (for a given experiment) at all other points in the eigenmode space, and then find the highest of those probabilities (smallest χ^2) out of our sets of model-based phenomenologies. We can then speak to the power of the experiments to distinguish between models.

Because we are only interested in the regions of closest approach between a model and a hypothetical realization, the exact distribution of the points in the chains used for this work is largely unimportant, so long as it does a reasonable job of exploring the possibilities of each. Our focus on the regions of closest approach is really a luxury which is enabled by the remarkable structure we have found in the eigenmode space. An outcome with more overlap (such as we do see between the Exponential and PNGB models) might draw one into a Bayesian analysis where priors on the different model distributions in mode space would play a critical role. One can loosely think of our approach as a very conservative one that allows for very unprincipled theorists who might cook up reasons to place delta function priors on the point of closest approach. We have shown that high quality stage 4 experiments could rule out entire models, even under these adverse circumstances.

There is an important assumption built into our analysis. Our choice of four test points spread out along the mode curves for each model implies that we are giving "equal weight" to the actual universe residing anywhere along the mode-space curves. In the current confused state of dark energy theory, one can expect a wide variety of individual preferences on this point. For someone with a different preference, our results about the structure in mode space remain unchanged, and will still impact any alternative analysis performed to express different prejudices about possible outcomes of future experiments. We feel we have built assumptions into our analysis which allow us to most directly address the following question: "If one of these experiments is carried out and gets a significant signal, what are the prospects for that signal completely ruling out a given scalar field model?" While we have focused our quantitative analysis on cosmologies based on one or the other of our scalar field models, one also can see (by inspecting the mode-space figures) that it would be quite possible to get an experimental signal that rules out all the scalar field models considered here. That again is a consequence of the mode-space structure demonstrated in this paper.

VII. CONCLUSIONS

We have considered the ability of future dark energy experiments to discriminate among different scalar field quintessence models of dark energy. To this end we have projected the equation-of-state functions $w(a)$ for each model into the space of best-measured eigenmodes of future experiments. We believe this approach is effective and convenient for investigating the ability of a given data set to discriminate among different quintessence models. Specifically, this approach offers a way around the fact that parameters of different quintessence models are typically not defined in the same spaces, which makes more direct comparisons of the models problematic.

The four quintessence models considered here create a distinctive structure when projected into the mode spaces. The goal of discriminating among these quintessence models gives an alternative and complementary measure of the impact of future experiments. In large part due to the structure in mode space, this measure has some striking features that are different from other measures considered previously.

We have shown that the DETF stage 3 data will have very little utility in discriminating among the four quintessence models, although it will significantly probe the possibility of noncosmological-constant-like behavior. DETF stage 4 simulated data appear to lie right at an interesting threshold in that this data show significant discriminating power among the quintessence models we considered. We also showed that modest improvements over DETF stage 4 (which many consider quite realistic for some stage 4 experiments currently proposed) allow one to cross this threshold more completely, leading to substantially greater discriminating power.

It is important to note that at our current level of theoretical understanding all quintessence models are suspect, and we are not advocating the use of the measures presented here to the exclusion of other approaches. However, as discussed in [8–11], we have chosen an interesting sampling of reasonably motivated quintessence models. Since such quintessence models are part of the current theoretical discussion of dark energy, discriminating power among these models should be part of how we evaluate the impact of dark energy experiments.

[1] J. Frieman, M. Turner, and D. Huterer, arXiv:0803.0982.
[2] A. Albrecht, AIP Conf. Proc. **957**, 3 (2007).
[3] A. Albrecht *et al.*, arXiv:astro-ph/0609591.
[4] E. V. Linder, Phys. Rev. Lett. **90**, 091301 (2003).
[5] D. Huterer and M. S. Turner, Phys. Rev. D **64**, 123527 (2001).
[6] D. Huterer and H. V. Peiris, Phys. Rev. D **75**, 083503 (2007).
[7] A. Albrecht and G. Bernstein, Phys. Rev. D **75**, 103003 (2007).
[8] M. Barnard, A. Abrahamse, A. Albrecht, B. Bozek, and M. Yashar, Phys. Rev. D **77**, 103502 (2008).
[9] A. Abrahamse, A. Albrecht, M. Barnard, and B. Bozek, Phys. Rev. D **77**, 103503 (2008).
[10] B. Bozek, A. Abrahamse, A. Albrecht, and M. Barnard, Phys. Rev. D **77**, 103504 (2008).
[11] M. Yashar, B. Bozek, A. Abrahamse, A. Albrecht, and M. Barnard (unpublished).
[12] D. Huterer and G. Starkman, Phys. Rev. Lett. **90**, 031301 (2003).
[13] L. Knox, A. Albrecht, and Y. S. Song, arXiv:astro-ph/0408141.
[14] E. V. Linder, Astropart. Phys. **24**, 391 (2005).
[15] A. G. Riess *et al.*, Astrophys. J. **659**, 98 (2007).
[16] R. de Putter and E. V. Linder, arXiv:0710.0373.
[17] S. Sullivan *et al.*, Phys. Rev. Lett. **100**, 241302 (2008).
[18] J. A. Frieman, C. T. Hill, A. Stebbins, and I. Waga, Phys. Rev. Lett. **75**, 2077 (1995).
[19] B. Ratra and P. J. E. Peebles, Phys. Rev. D **37**, 3406 (1988).
[20] A. Albrecht and C. Skordis, Phys. Rev. Lett. **84**, 2076 (2000).
[21] B. Ratra and P. J. E. Peebles, Phys. Rev. D **37**, 3406 (1988).
[22] E. J. Copeland, M. Sami, and S. Tsujikawa, Int. J. Mod. Phys. D **15**, 1753 (2006).
[23] M. Schneider, L. Knox, H. Zhan, and A. Connolly, Astrophys. J. **651**, 14 (2006).
[24] H. Zhan, J. Cosmol. Astropart. Phys. 08 (2006) 008.
[25] R. R. Caldwell and E. V. Linder, Phys. Rev. Lett. **95**, 141301 (2005).
[26] The fact that each model occupies a somewhat ''one dimensional'' region in the mode space is part of why it seems reasonable to represent the distribution of data outcomes by the four test points for each model.
[27] A technical clarification: The ''model points'' used here are the unique points taken from MCMC, so the repeated points in the original chain, where the algorithm did not step to another point, are counted only once for the purposes of these plots. As discussed in Sec. VI, we expect this omission to impact the detailed shapes of the curves shown here but not the minimum χ^2 value (where the curve hits the x axis). Since we focus on this minimum χ^2 value in our analysis, we do not consider the effect this has on the graphs to be significant or meaningful enough for our analysis to justify the increased computation including the repeated points would entail.