Cosmology with galaxy cluster phase spaces

Alejo Stark,[1,1] Christopher J. Miller,[1,2] and Dragan Huterer[2]

[1]*Department of Astronomy, University of Michigan, Ann Arbor, Michigan 48109, USA*
[2]*Department of Physics, University of Michigan, Ann Arbor, Michigan 48109, USA*
(Received 22 November 2016; published 31 July 2017)

We present a novel approach to constrain accelerating cosmologies with galaxy cluster phase spaces. With the Fisher matrix formalism we forecast constraints on the cosmological parameters that describe the cosmological expansion history. We find that our probe has the potential of providing constraints comparable to, or even stronger than, those from other cosmological probes. More specifically, with 1000 (100) clusters uniformly distributed in the redshift range $0 \leq z \leq 0.8$, after applying a conservative 80% mass scatter prior on each cluster and marginalizing over all other parameters, we forecast 1σ constraints on the dark energy equation of state w and matter density parameter Ω_M of $\sigma_w = 0.138(0.431)$ and $\sigma_{\Omega_M} = 0.007(0.025)$ in a flat universe. Assuming 40% mass scatter and adding a prior on the Hubble constant we can achieve a constraint on the Chevallier-Polarski-Linder parametrization of the dark energy equation of state parameters w_0 and w_a with 100 clusters in the same redshift range: $\sigma_{w_0} = 0.191$ and $\sigma_{w_a} = 2.712$. Dropping the assumption of flatness and assuming $w = -1$ we also attain competitive constraints on the matter and dark energy density parameters: $\sigma_{\Omega_M} = 0.101$ and $\sigma_{\Omega_\Lambda} = 0.197$ for 100 clusters uniformly distributed in the range $0 \leq z \leq 0.8$ after applying a prior on the Hubble constant. We also discuss various observational strategies for tightening constraints in both the near and far future.

I. INTRODUCTION

The relation between the dynamics of clustered galaxies and the totality of the observable Universe has long been a fruitful site of investigation for physical cosmology. One good example is Zwicky's seminal investigation of dark matter, which crystallized through an analysis of the dynamics of galaxies in the Coma cluster (then referred to as "clusters of nebulae") [1]. Three decades ago, also utilizing the Coma cluster as a cosmological laboratory, Ref. [2] demonstrated that the Universe's matter energy density was subcritical—thereby pointing to the existence of some other unknown substance and demonstrating the capacity for the infall regions of galaxy clusters to be mobilized as a powerful cosmological probe. Just a few years after this, Ref. [3] used the Coma cluster in conjunction with three other nearby clusters to argue for the existence of "caustics" in the redshift-distance space of galaxy clusters. These caustics trace an identifiable edge that would later be demonstrated to be representative of the radial escape velocity profile of galaxy clusters [4]. From this point on, the caustics would be used solely to generate mass profiles of galaxy clusters via the so-called caustic technique [5,6]. These dynamical cluster masses have been used to constrain cosmology through the cluster mass function [7]. However, the capacity of caustic edges themselves to constrain cosmological parameters has not been pursued since the work of Ref. [8].

Following Refs. [2,3,8] but also deviating significantly from their approach, Ref. [9] demonstrated the necessity to include a cosmological term to describe the escape velocity profile of galaxy clusters as inferred from their projected phase spaces. In particular, Ref. [9] presented an analytic model based on the Poisson equation that can reproduce the projected escape velocity profiles of galaxy clusters as measured from their phase spaces. The analytical escape velocity profile prediction requires a known mass profile and a known velocity anisotropy profile $\beta(r)$. Given these and the cosmological parameters, the analytical escape velocity edge has been shown to match expectations to high precision and accuracy using N-body simulations [10].

If both weak lensing (WL) mass estimates and a measurement of β can be inferred for a galaxy cluster, we can turn this around and through analytic theory constrain cosmology by measuring edges of clusters through their phase spaces (see Refs. [6,10–13] for the various methods utilized to estimate the escape velocity edge from the phase space of galaxy clusters).

However, while Ref. [9] demonstrated the necessity to include a cosmological and redshift-evolving term in the escape velocity profile of clusters, it did not quantify the precision with which one can constrain cosmological parameters with this observable. This is the task that this paper takes on.

Our paper is organized as follows: in Sec. II we describe the theoretical observable we work with—the redshift-evolving and cosmology-dependent escape velocity profile of galaxy clusters. In Sec. III we detail how we use the Fisher matrix formalism to quantify how well this observable can constrain cosmology given current and future systematic errors. We present the results of this analysis in Sec. IV. In Sec. V we present observational strategies that may be utilized to optimize cosmological constraints. In Sec. VI we

discuss our observable in relation to other probes and speculate as to how we may improve constraints in the future through joint likelihood analyses. We conclude the paper and provide ways of extending our work in Sec. VII. The appendix details the construction of our Fisher matrix.

II. THEORETICAL OBSERVABLE

The theoretical observable quantity we work with throughout this paper is the projected escape velocity radial profile of galaxy clusters. Observationally, the escape velocity profile of a cluster is inferred from the phase space (v_{los} vs r space) of the cluster. More specifically, the line of sight galaxy velocity (v_{los}) vs physical distance (r) space is constructed by measuring the redshifts of cluster galaxies (z) of a cluster at a redshift z_c and then converting them to velocities via

$$v_{los}(r) = c\frac{(z - z_c)}{(1 + z_c)}, \tag{1}$$

where c is the speed of light. The physical distance from the cluster's center (r) is inferred from the angular diameter distance (d_A) and the measured angular separation (θ),

$$r = d_A(z)\theta = \left[\frac{1}{1+z}\frac{cH_0^{-1}}{\sqrt{\Omega_K}}\sin\left(\sqrt{\Omega_K}\int_0^z\frac{dz'}{E(z')}\right)\right]\theta. \tag{2}$$

H_0 is the Hubble constant and the redshift-evolving function $E(z)$ is detailed in Eq. (10) below. Note that $\sin(x)$ (the nonflat closed universe case) becomes x for the flat universe case. Also detailed below is the parameter Ω_K which quantifies the openness or closedness of the Universe. The escape velocity profile can then be inferred from this phase space through various techniques (see Refs. [6,10–13]).

This phase-space-inferred escape velocity profile can be modeled with a function of the mass distribution of a specific cluster as specified by its gravitational potential [in our case we use the Einasto profile with three free parameters: $\alpha, r_{-2}, \rho_{-2}$; see Eq. (6) below], and the anisotropy parameter (β) of that specific cluster (which allows us to take into account projection effects). As mentioned in the introduction, the profile is also a function of redshift z and cosmology (Ω_M, h, etc.). The escape velocity radial profile is therefore given by a function of these cosmology and cluster parameters combined,

$$v_{esc}(r, z, \Omega_M, h, \ldots, \beta, \alpha, r_{-2}, \rho_{-2}). \tag{3}$$

We note that while we utilize only Einasto density profiles in this paper, in principle, any parametrized mass profile can be used in our framework, as long as the parametrization is a density-potential Poisson pair (see, e.g., Ref. [10]).

While the mass dependence and projection effects have long been considered in studies of this observable (see for instance Ref. [11]), only recently has the cosmological

dependence of the escape velocity profile been considered. More specifically, the cosmological dependence of the escape velocity has been studied in relation to both observational data and simulations of standard general relativistic cosmology as well as modified theories of gravity [9,10,14]. These investigations found the need to include a cosmological term in order to reproduce numerical results as well as observational data.

Qualitatively, it should be unsurprising that the cosmological background within which a cosmic structure is embedded at a given epoch will create the conditions for its evolution and development. For instance, whether a gravitationally bound structure today (say, a galaxy group or galaxy cluster) can become unbound at late times in an accelerating universe is a function of both the curvature of space and the mass-energy content of the background the particular structure is embedded in [15,16]. It is therefore clear that the theoretical escape velocity profile must take into account the cosmological background.

More specifically, the cosmological dependence in the escape velocity profile in Refs. [9,10,14] comes in through the "equivalence radius." In an accelerating universe, the radius out to which one has escaped a cluster's potential is a function of cosmology. This minimal radius required to escape, termed the equivalence radius (r_{eq}), decreases in an accelerating universe (see Ref. [9] and references therein). The projected escape velocity profile is then given by

$$v_{esc}(r, z) = \sqrt{\frac{1}{g(\beta)}\left[-2(\Psi(r) - \Psi(r_{eq})) - qH^2(r^2 - r_{eq}^2)\right]}. \tag{4}$$

We note that, as expected, at $r = r_{eq}$ the escape velocity is nil. The function of the equivalence radius then is to normalize the escape velocity at that point. Note that this equation is derived by integrating the acceleration equation of the effective potential that takes into account both the "negative" acceleration due to the mass of the cluster and the "positive" acceleration of the background at late times. We therefore obviate the negative acceleration due to ram pressure on galaxies given that it is negligible when compared to averaged gravitational effects [17]. The "equivalent radius," then, is named for the condition that it sets: it is the point at which the negative (inward) acceleration due to the pull of the cluster and the positive (outward) acceleration due to the acceleration of the Universe balance each other. For a derivation of Eq. (4) see Refs. [9,10]. We note also that the physical distance r in Eq. (4) is cosmology dependent via Eq. (2).

Lastly, as detailed in Ref. [9], Eq. (4) emerges from the acceleration equation derived by Ref. [18] in the context of a flat universe. However, in the cosmological regime we work with, the acceleration equations for a spatially flat

and nonflat universe converge. See Secs. 5.3.1–5.3.3 of Ref. [18] and Sec. 2.1 of Ref. [9].

We now consider the two main components of the projected escape velocity profile function [Eq. (4)]: cluster parameters (projection effects and mass profile information) and cosmological parameters.

A. Cluster parameters

The two components of the cluster parameter set are the mass profile parameters and the anisotropy parameter β encapsulated by the function $g(\beta)$.

The anisotropy parameter β is given by $\beta(r) = 1 - \sigma_t^2/\sigma_r^2$, where σ_t is a function of the azimuthal and tangential velocity dispersions and σ_r is the radial velocity dispersion. See Sec. 5.1 of Ref. [9] for details. Note that while the anisotropy profile of the cluster is actually a function of radius [$\beta(r)$], the observed and simulated profile is nearly flat within 0.3–1 virial radii [12,19,20]. Therefore, in what follows we only consider the escape velocity profile within this radial range so that we can reduce the anisotropy profile to a single value for a given cluster. Once β for a given cluster is inferred we can then use the function $g(\beta)$ to project our escape velocity profile (see Refs. [9,11]). In particular, this function is defined geometrically and given by

$$g(\beta) = \frac{3 - 2\beta}{1 - \beta}. \tag{5}$$

As can be implied from this equation, the effect of $g(\beta)$ on the escape velocity profile within the radial range we consider below is to suppress the profile by a constant value when compared to the three-dimensional case [$g(\beta) = 1$]. Quantitatively, the limits of this function are set by the limiting cases of the anisotropy parameter β: radial infall ($\beta = 1$), circular motion ($\beta = -\infty$) and isotropy ($\beta = 0$). For the radial range we are considering, on average, $g(\beta) \sim 3.3$. This entails that, on average, within the radial range considered below, nonprojected escape velocity profiles are $\sqrt{g(\beta)} \sim 1.8$ times higher than projected profiles. As a rule of thumb, the more radial the velocity anisotropy of a cluster, the more suppressed the escape velocity profile will be. For a more thorough quantitative analysis of these projection effects see Refs. [9,12].

The other component of the cluster parameters that make up our observable is the mass profile of a given cluster. Information about the mass profile of the cluster comes through the gravitational potential $\Psi(r)$. Following [9,10] we pick the Einasto representation of the potential given its capacity to trace the mass distribution of galaxy clusters beyond the virial radius. The gravitational potential $\Psi(r)$, then, is a function of three free parameters: the shape parameter α, the radius where the logarithmic slope of the density profile is equal to -2 (r_{-2}) and the density at r_{-2} (ρ_{-2}). As calculated via the Poisson equation [21], the potential as inferred from an Einasto density field is

$$\Psi(r) = -\frac{GM}{r}\left[1 - \frac{\Gamma(3/\alpha, s^\alpha)}{\Gamma(3/\alpha)} + s\frac{\Gamma(2/\alpha, s^\alpha)}{\Gamma(3/\alpha)}\right]. \tag{6}$$

$\Gamma(a, x)$ is the upper incomplete gamma function. We are also utilizing the unitless scale radius s given by

$$s = \frac{r}{r_{-2}}\left(\frac{2}{\alpha}\right)^{1/\alpha}. \tag{7}$$

And we rewrite the total mass (M) as

$$M = 4\pi\rho_{-2}r_{-2}^3 F(\alpha). \tag{8}$$

In this last equation we have defined the function of α,

$$F(\alpha) = \Gamma(3/\alpha)\left(\frac{e^2}{8}\alpha^{3-\alpha}\right)^{1/\alpha}. \tag{9}$$

Note that in general we follow the definitions and specifications used in Refs. [9,10]. We treat the mass profiles and β as observable quantities with uncertainties. In simulations, β shows no significant evolution as a function of redshift in the range we are interested in (i.e., $0 \leq z \leq 0.8$; see Fig. 2 in Ref. [22]) and is inferred independently of cosmological assumptions [13,23–26]. This is not the case for the weak lensing mass profile-inferred parameters (namely α, r_{-2}, and ρ_{-2}). We analyze how systematic errors introduced by uncertainties in cosmological parameters affect weak lensing mass errors in Sec. III C below as well as in the appendix.

B. Cosmological parameters

Having defined the first set of parameters of the escape velocity profile related to its mass content (α, r_{-2}, ρ_{-2}) and projection effects (β), let us now focus on the redshift-evolving and cosmology-dependent terms in Eq. (4), namely, the Hubble parameter H, the deceleration parameter q, and the equivalence radius r_{eq}. We describe these terms below.

Given that we work in a regime where the radiation energy density is negligible ($z \leq 0.8$), the Hubble parameter is given by

$$H^2 = H_0^2 E^2(z) = H_0^2\left[\Omega_M(1+z)^3 + \Omega_k(1+z)^2 \right.$$
$$\left. + \Omega_{\text{DE}}\exp\left\{3\int_0^z d\ln(1+x)[1+w(x)]\right\}\right], \tag{10}$$

where Ω_M and Ω_{DE} are energy densities in matter and dark energy relative to critical, $w(z)$ is the time-varying equation of state (EoS) of dark energy, and the spatial curvature density parameter is $\Omega_k \equiv 1 - \Omega_M - \Omega_{\text{DE}}$. Note that throughout this paper we work with the scaled Hubble constant h defined via $H_0 = 100h$.

Once we have the Hubble parameter as a function of redshift we can derive the corresponding deceleration
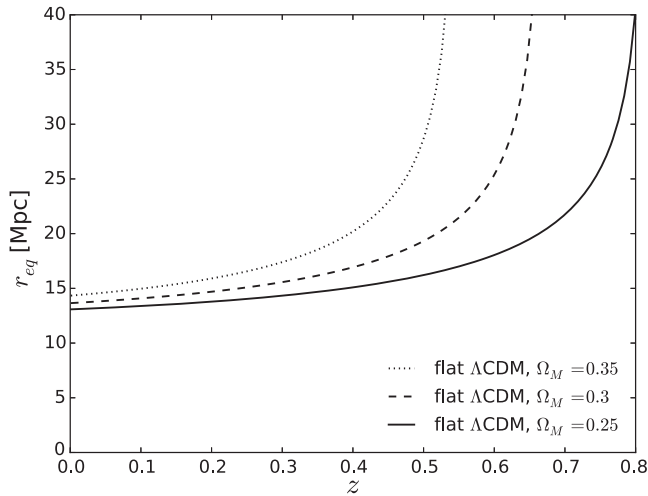
FIG. 1. Behavior of the equivalent radius (r_{eq}) as a function of redshift (z) for a cluster with mass $M_{200} = 4 \times 10^{14} \, M_\odot$ for three values of Ω_M in a flat ΛCDM universe and $h = 0.7$. At the transition redshift for each given cosmology, the radius shoots up to infinity. Note how r_{eq} becomes more and more sensitive to Ω_M at higher z.

parameter, $q \equiv -\ddot{a}a/\dot{a}^2$, where a is the scale factor, as a function of redshift via

$$q = \frac{(1+z)}{H} \frac{dH}{dz} - 1. \tag{11}$$

Lastly, the equivalent radius (r_{eq}) is the physical distance at a given cosmic epoch, for a given cosmology, where the inward pull of gravity balances the outward pull of cosmic acceleration. It is given by [16]

$$r_{eq} = \left(\frac{GM}{-qH^2} \right)^{1/3}. \tag{12}$$

This radius is what sets the normalization of the gravitational potential of our galaxy cluster, $\Psi(r_{eq})$, in Eq. (4) [9,10]. That is, beyond this equivalence radius, the effective potential of the cluster described by Eq. (4) $[-2\phi = v_{esc}^2(r,z)]$ is normalized to 0 (see Ref. [9] for a derivation). Note also that the distance at which this balance of forces occurs, that is, between cosmic acceleration and the cluster's gravitational pull, is far away enough from the cluster center that the cluster can be represented as a point mass. Lastly, this quantity can be thought of as a proxy for how sensitive $v_{esc}(r,z)$ is to cosmology. In lieu of taking analytic derivatives of Eq. (4) one can study the analytic derivatives of r_{eq}. We have confirmed this by looking at numerical derivatives of $v_{esc}(r,z)$ and compared them to the analytic derivatives of r_{eq} for the same cosmology, finding good agreement.

In Fig. 1 we show that, for a fixed cluster mass the equivalence radius is sensitive to both cosmology and
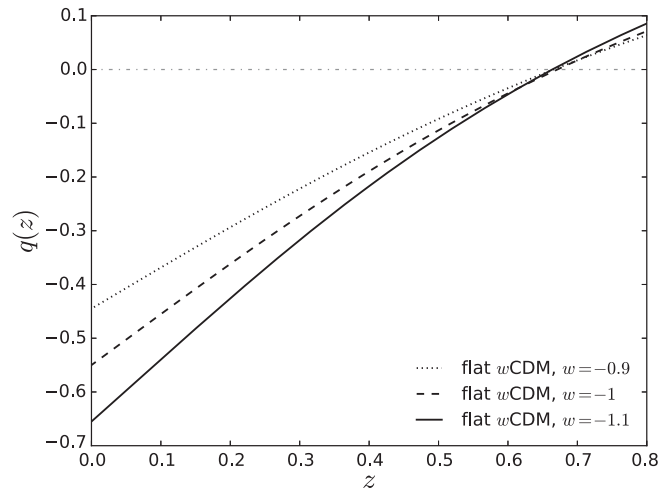

FIG. 2. Redshift evolution of the deceleration parameter in a flat universe with fixed $\Omega_M = 0.3$ and $h = 0.7$. Notice the divergence of values of q at low redshifts when the equation of state parameter is varied.

redshift. More specifically, we pick a mass of $M_{200} = 4 \times 10^{14} \, M_\odot$, where M_{200} is defined as the mass enclosed by a sphere with an average density equivalent to 200 times the mean density of the Universe. Note that as the cluster gets closer to the acceleration transition redshift [$q(z) = 0$] the equivalence radius shoots up to infinity. We consider the effects of this behavior on our observable below.

1. $v_{esc}(r,z)$ in an accelerating universe ($q < 0$)

Within the $q < 0$ regime, the escape velocity profile is described by Eq. (4). As shown by Ref. [9], in an accelerating universe the effect on the observable generated by changing the various cosmological parameters is to modify both the amplitude and the shape of the escape velocity profile. For instance, in a flat universe, a larger dark energy density makes it easier to escape a galaxy cluster, while a larger dark matter component makes it harder to escape the cluster. This is illustrated in Fig. 1. Using the equivalence radius as a proxy to gauge how $v_{esc}(r,z)$ changes, we see that the equivalence radius blows up at the given transition redshift for that cosmology [i.e. z that yields $q(z) = 0$].

It is important to emphasize that, in the regime where $q(z) < 0$, $v_{esc}(r,z)$ is a direct measure of both expansion and acceleration, since qH^2 is a function proportional to both $H(z)$ and $dH(z)/dz$. This makes our observable a powerful probe of cosmology. For example, note the sensitivity of the deceleration parameter to the dark energy equation of state w as shown in Fig. 2. In particular, the variation of $q(z)$ with respect to w increases at lower redshifts. Figure 3 also demonstrates the sensitivity of the observable to w by showing the fractional difference between the escape velocity for a flat ΛCDM universe at $z = 0$ and two dark energy models, quintessencelike dark energy (solid line) [27] and phantom dark energy (dotted
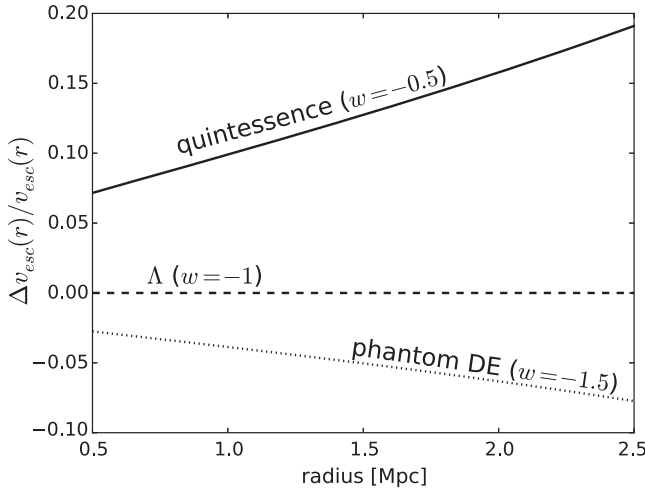
FIG. 3.   Using a single cluster with mass $M_{200} = 4 \times 10^{14}\ M_{\odot}$ at $z = 0$, we show the fractional difference between the escape velocity profile of a flat wCDM universe with dark energy equation of state $w = -1$ and two other dark energy models (quintessence in the solid line and phantom dark energy in the dotted line). More specifically, $\Delta v_{\text{esc}}(r, z)/v_{\text{esc}}(r, z) = [v_{\text{esc}}(w)/v_{\text{esc}}(w = -1)] - 1$. Quintessence therefore acts similarly to increasing the dark matter density. That is, it increases the escape velocity profile relative to the Λ case (see Fig. 1 of Ref. [9]). On the other hand, the phantom dark energy suppresses the escape velocity profile relative to the Λ case. Lastly, we highlight that the fractional difference increases with radius in both cases.

line) [28]. Note the $\sim$10% level differences between these models and Λ (dashed line). A quintessence dark energy, then, would increase the escape velocity profile, whereas a phantom dark energy would decrease the escape velocity (both relative to ΛCDM).

2. $v_{\text{esc}}(r, z)$ at the transition redshift ($q = 0$)

As implied in the previous subsection, as we approach $q = 0$ the equivalence radius blows up and Eq. (4) is reduced to the gravitational potential described by the Einasto gravitational potential $\Psi(r)$. More explicitly,

$$\lim_{q \to 0} v_{\text{esc}}(r, z) = \sqrt{\frac{1}{g(\beta)} [-2\Psi(r)]}. \tag{13}$$

One immediate consequence we derive from this behavior is that the only cosmological dependence we get beyond this point is through r [see Eq. (2)].

We now investigate what happens to our observable beyond the transition redshift or, equivalently, what happens to our observable with combinations of cosmological parameters that yield a decelerating universe (i.e. $q > 0$).

3. $v_{\text{esc}}(r, z)$ beyond the transition redshift ($q > 0$)

In the regime beyond the transition redshift, or when $q > 0$, it makes no sense to speak of an equivalence radius.

Recall that the equivalence radius is defined in the context of a balance of forces. Given that there is no balance of forces between the mass-induced pull of gravity and the repulsive acceleration as there is in the case where $q < 0$, there is no such equivalence radius.

We note that within the virial radius, the theoretical expectation embodied in Eq. (4) works to high precision up to $z \sim z_t + \delta z$, where z_t is the transition redshift and δz is small. See, for example, Fig. 4 in Ref. [10]. Beyond the transition redshift, however, the analytic theory of Eq. (4) is complicated both by cluster assembly dynamics and the theoretical ambiguity of what occurs to the escape velocity profile in a universe that is approaching the Einstein–de Sitter case. When $q > 0$, the internal dynamics of a bound system like a galaxy cluster are solely governed by the Poisson equation with gravity acting to source to accelerate the member galaxies.

Given this, in the Fisher matrix analysis that follows, we pick $z = 0.8$ as the maximum redshift out to which we can realistically push our probe. Later on in the paper we discuss the implications of defining a redshift range for the probe.

Having described both the cluster and cosmological parameter dependence of our probe, and having considered the regimes of applicability of our probe, we now quantitatively characterize, through the Fisher matrix formalism, how well our observable may be able to constrain cosmological parameters.

III. FISHER MATRIX

The Fisher matrix formalism has been vital in predicting how well a given cosmological probe can constrain any given set of cosmological parameters [29–33].

In this section we briefly go over the Fisher matrix formalism. We construct the Fisher matrix for our observable quantity, include parameters describing systematic errors in these observations, discuss how we apply priors on cluster and cosmological parameters, and calculate the marginalized errors on cosmological parameters.

A. Formalism

In general, the Fisher information matrix is a function of the derivatives of the log likelihood of our observable with respect to the observable's parameters (p),

$$F_{ij} = \left\langle -\frac{\partial^2 \ln \mathcal{L}}{\partial p_i \partial p_j} \right\rangle. \tag{14}$$

Given that the observable quantity is the escape velocity as a function of redshift and radius we have that the Fisher matrix elements are sums over clusters (index n) and radii (indices k and l),

$$F_{ij} = \sum_{nkl} \frac{\partial v_{\text{esc}}(z_n, r_k)}{\partial p_i} (C^{-1})_{kl} \frac{\partial v_{\text{esc}}(z_n, r_l)}{\partial p_j}, \tag{15}$$

where C is the covariance of escape velocity measurements at different radii. Here we assume that the escape velocity measurements in different clusters are uncorrelated, and that the measurement covariance for different radii in a cluster (matrix C) is independent of redshift.

After adding priors to our cluster parameters the Fisher information matrix is then given by

$$F = F_{ij} + F_{\mathrm{prior}}. \qquad (16)$$

From this F we can then compute the marginalized lower bound of the uncertainty on any of our N parameters via the Cramér-Rao inequality, $\sigma_{p_i} \geq \sqrt{(F^{-1})_{ii}}$. The inverse is calculated via Gaussian elimination. We ensure that the inversion of the matrices we work with is stable by calculating the condition number for each F matrix and ensuring that it is less than $\sim 10^{12}$.

As detailed in the previous section, the parameters in our observable can be split up into two sets: cluster and cosmological parameters. Consequently, the parameters from which we take the derivatives with respect to (p) can be represented by the union of the following two sets:

$$p \in p_{\mathrm{clus}} \cup p_{\mathrm{cosmo}}. \qquad (17)$$

The parameters describing the clusters 1 to N (i.e. three parameters for the mass profile plus one more describing the anisotropy parameter, per cluster) can then be encapsulated in the following set:

$$p_{\mathrm{clus}} \in \{\beta_1, \alpha_1, r_{-2,1}, \rho_{-2,1}, \ldots, \beta_N, \alpha_N, r_{-2,N}, \rho_{-2,N}\}. \quad (18)$$

The cosmological parameter set p_{cosmo} is composed of the cosmological parameters. We detail the cosmological models we study, and the corresponding sets of p_{cosmo}, in the next section.

Considering these sets, we therefore have a N_{dim} by N_{dim} dimensional Fisher matrix F given by

$$N_{\mathrm{dim}} = 4 \times N_{\mathrm{clus}} + N_{\mathrm{cosmo}}, \qquad (19)$$

where N_{clus} is the number of clusters and N_{cosmo} is the number of cosmological parameters. We provide a sketch of the Fisher matrix structure in Fig. 12 and explore its structure more thoroughly in the appendix.

B. F_{ij} matrix

In this subsection we focus our attention on the components that make up the F_{ij} matrix [Eq. (15)].

1. Fiducial cluster

As with any Fisher matrix analysis, the derivatives of Eq. (15) are calculated at a given set of fiducial values. For our fiducial cosmology we pick $p_{\mathrm{cosmo,fid}} \in \{\Omega_M = 0.3,$

$\Omega_\Lambda = 0.7, w_0 = -1, w_a = 0, h = 0.7\}$. For our fiducial cluster parameters we pick $p_{\mathrm{clus,fid}} \in \{\alpha = 0.1984, \rho_{-2} = 1.0521 \times 10^{14} \, [M_\odot/\mathrm{Mpc}^3], r_{-2} = 0.497 \, [\mathrm{Mpc}], \beta = 0.145\}$. The β fiducial value is around what has been estimated for galaxy clusters (see Refs. [13,23–26]) and the three Einasto fiducial cluster parameters are equivalent to a cluster of mass $M_{200} = 4 \times 10^{14}$. More specifically, we use the mass-concentration relation in Ref. [34] to map our fiducial mass M_{200} to the Einasto parameters at $z = 0$ by fitting the Navarro-Frenk-White (NFW) density profile to the Einasto density profile. For analytical representations of these density profiles see Ref. [35] and references therein.

Furthermore, to calculate the derivatives of Eq. (15) we place our fiducial cluster along different redshifts z_n and recalculate its angular size θ with fixed r at that given redshift via Eq. (2). We discuss this radial range in the next subsubsection.

In Fig. 4 we plot some of the derivatives for our fiducial cluster used in our Fisher matrix analysis. More specifically, we plot the radial derivatives of $v_{\mathrm{esc}}(r, z)$ with respect to various cosmological parameters for 100 clusters of the same fiducial mass uniformly distributed in the range $0 \leq z \leq 0.8$. Figure 4 is a useful way to study the sensitivity of our observable to various cosmological parameters. In particular Fig. 4 tells us that our observable is more sensitive to certain cosmological parameters at higher redshifts, such as Ω_M (the red lines, which represent high redshift clusters are higher than the blue ones, which represent low redshift clusters), and more sensitive to other parameters at lower redshift, such as the dark energy equation of state parameter w. Judging from the derivatives in Fig. 4, our probe is most sensitive to the parameter Ω_M. Note also that for all parameters, the sensitivity increases with radius. We emphasize that beyond the transition-to-acceleration redshift ($z_t = 0.671$) in the standard ΛCDM model we no longer gain much cosmological information as encapsulated by Eq. (13). Beyond this redshift we still get cosmological information via r, which is a function of the angular diameter distance [Eq. (2)].

2. Radial bins and covariance matrix C^{-1}

Having described the derivatives of Eq. (15) we now describe the covariance matrix C in that same equation. Simply put, this matrix embodies the covariance of between different measurements of our observable at a given radial bin in a cluster n.

For a perfect three-dimensional observation of the galaxy velocities, there would be no projection effects and therefore one would expect a nonzero covariance between v_{esc} at different radii. This effect is random between different radial bins, and it effectively decouples the measurements in different radial bins, reducing the bins' covariance. More specifically, this scatter arises from the fact that when observing a galaxy cluster, random galaxies along the line of sight may be included in the phase space. This drastically
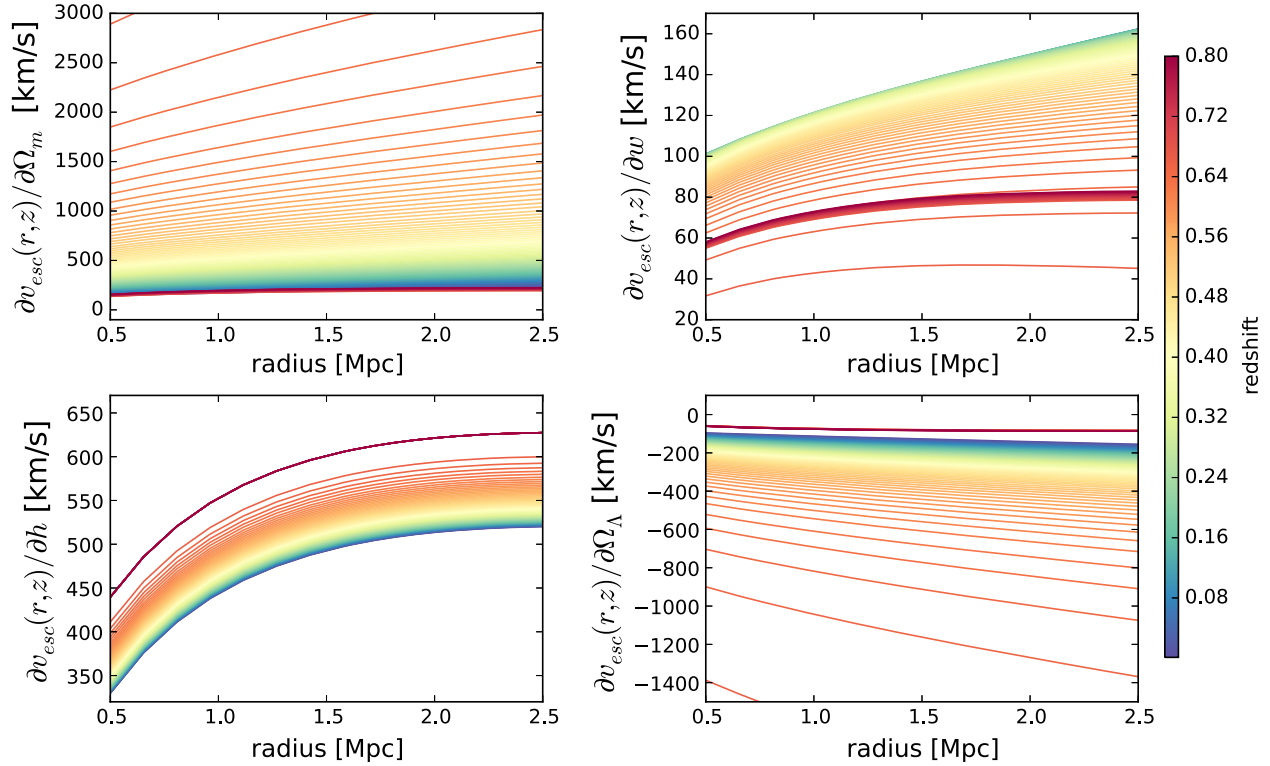
FIG. 4.   Sensitivity of the observable $v_{esc}(r,z)$ to cosmology. The partial derivatives of the escape velocity profile are calculated numerically at 100 different redshifts (z) with respect to the various cosmological parameters for our fiducial cluster. The specific redshift of a given profile can be identified through the color bar on the right-hand side of the figure. In all cases, the information grows (i.e. the observable becomes more sensitive) the farther out we go radially. Some parameters are most sensitive at higher redshifts (Ω_M, Ω_Λ and h) while others are more sensitive at lower redshifts (w). Note that beyond the transition redshift for our fiducial cosmology, the derivatives with respect to all parameters reach a limit, as implied by Sec. II B 2.

reduces the covariance between radial bins. Therefore, it is a good assumption that the radial covariance matrix is diagonal. That is, C^{-1} is reduced to $\sigma_{v_{esc}}^{-2}$.

Mathematically, at any given radius, the uncertainty in $v_{esc}(r,z)$ is given by the combination of the spectroscopic uncertainty, uncertainty due to the edge measurement, and any additional intrinsic uncertainties,

$$\sigma_{v_{esc}} = \sqrt{\sigma_{spec}^2 + \sigma_{edge}^2 + \sigma_{edge,int}^2}. \qquad (20)$$

We choose $\sigma_{spec} = 50 \text{ km s}^{-1}$ to match the redshift accuracy of modern spectroscopic surveys [36]. To calculate the uncertainty on the statistically inferred edge at a given radial bin we follow Ref. [37] who used simulations to show that when viewing along a line of sight the edge can be recovered to high statistical precision ($\sim$5% or less) when 100 or more galaxies are used in the phase spaces (see their Fig. 4—bottom right). For our fiducial cluster this corresponds to 50 km s^{-1} since the observed edge is typically around 1000 km s^{-1}. We also allow for an intrinsic scatter between the observed edge and how accurately it can recover the true underlying gravitational potential. Reference [38] used simulations to quantify the

statistical accuracy and precision of the cluster mass using the projected edge when the density, potential, and anisotropy are known exactly. They find that there is a statistical floor of 25% error in mass, which translates into a 12.5% error in the edge since mass scales as the square of the potential. For our fiducial cluster, this intrinsic error corresponds to $\sim$125 km s^{-1}.

As in Eq. (20), we sum these three components in quadrature giving $\sigma_{v_{esc}} = 143.61$ km s^{-1} for our fiducial cluster. This is about a 15% total error in the measurement of the projected phase-space escape velocity edge. Lastly, note that if we double the total error budget on v_{esc}, our constraints on the cosmological parameters increase by about 80%.

We next tackle the question of how many radial bins we should use for a given cluster in our Fisher matrix calculations.

Our one requirement here is to be able to resolve the shape of the velocity edge vs radius, $v_{esc}(r)$. In all cases, we assume densely sampled phase spaces (i.e., 100–200 galaxies within the virial radius). Since more massive clusters are also larger in size, we can maximize the number of useful radial bins by choosing the most massive clusters for our analysis.

At the moment, we can typically rely on reasonably accurate weak lensing mass estimates for the highest mass systems in the Universe. Therefore, for this paper we assume a South Pole Telescope-like sample with $M_{200} > 3 \times 10^{14} M_\odot$ [39]. More specifically, as mentioned elsewhere in the paper, we pick a fiducial mass of $M_{200} = 4 \times 10^{14} M_\odot$.

Furthermore, as detailed in Sec. II A, we can only work within the radial range of about 0.3–1 virial radii if we assume that the anisotropy profile can be reduced to a single parameter. Henceforth for our profiles and constraints we pick a radial range between $0.5 \leq r \leq 2.5$ Mpc, the outer range corresponding to about r_{200} (where the density reaches $200\times$ the mean value) and the inner range corresponding to about $0.3 \times r_{200}$ described above [40]. This radial range, given $\Delta r \sim 0.1$ Mpc, yields $N_{\text{bins}} = 14$ radial bins. This is what we use in our calculations throughout our paper [see Eq. (21) below].

Note that if we change the number of radial bins from 14 to 7 and 21, the constraints on the cosmological parameters change by $\sim 20\%$. We can therefore, in principle, get better constraints than what is presented in the next section by increasing the number of radial bins. However, binning too finely is not desirable given that it can introduce additional statistical noise in the observable and may not even be possible, given the density of galaxies in the phase spaces that are observationally viable.

Given all of this, Eq. (15) thereby becomes

$$F_{ij} = \sum_{n=1}^{N_{\text{clus}}} \sum_{k=1}^{N_{\text{bins}}} \frac{1}{\sigma_{v_{\text{esc}}}^2} \frac{\partial v_{\text{esc}}(r_k, z_n)}{\partial p_i} \frac{\partial v_{\text{esc}}(r_k, z_n)}{\partial p_j}, \quad (21)$$

and is therefore a sum over N_{clus} clusters and N_{bins} radial bins. This is the F_{ij} matrix we utilize for all of our constraints presented in Sec. IV, in conjunction with the prior matrix which is discussed in the following subsection.

C. Prior information

The Fisher information matrix formalism allows us to add additional independently measured information attained on certain model parameters (both cluster and cosmological) to our Fisher matrix. This is implemented via the prior information matrix F_{prior} in Eq. (16). The structure of this matrix is given by

$$F_{\text{prior}} = \begin{pmatrix} C_{\text{cosmo}}^{-1} & & & 0 \\ & C_{\text{cluster}}^{-1} & & \\ & & C_{\text{cluster}}^{-1} & \\ 0 & & & \ddots \end{pmatrix}. \quad (22)$$

We discuss the elements of this prior information matrix below.

1. Cluster prior information (C_{cluster}^{-1})

In our case, the mass parameters (which come from weak lensing mass estimates) and the anisotropy parameter (which comes from analysis of the phase spaces of the clusters via the Jeans equations) are known to within some precision from these independent measurements. We therefore add a prior information matrix to account for this external information on the noncosmological parameters.

More specifically, the error bars on the weak lensing mass estimates are chosen to be similar to what is reported in the literature based on recent observations. As a representative sample, see the metacatalog compiled in Ref. [34]. We particularly choose an M_{200} error range from 20% to 40% which is based on ground-based imaging. For instance, Ref. [41] reports typical statistical errors of $\sim 20\%$ using Surprime-Cam imaging for 50 clusters to a redshift of 0.7. Similarly, Ref. [42] primarily used Subaru/Suprime-Cam to obtain weak lensing errors at a level of 20%–30% based on the CLASH sample. Reference [43] also reports $\sim 40\%$ statistical errors for four clusters using science verification data from the Dark Energy Survey on the CTIO 4m Blanco DECam imager. This is why, as our baseline, we use the upper range of 40% statistical error on the mass. In what follows we also consider 5% statistical error as a floor that can be potentially achieved through the technique of stacking galaxy clusters [44,45].

However, there are also systematic errors that need to be considered. These can come from a variety of observational sources including the telescope point-spread function, the background redshift distribution, the intrinsic shape variations of the background galaxies, and more. Reference [41] reports systematic errors that are small compared to the statistical errors (7%) when the telescope optical system is well characterized.

Few researchers have allowed for cosmology to vary during the weak lensing mass estimation process. As we approach higher precision requirements, we need to incorporate variations in cosmology when calculating the cluster masses and it could play an important role in the overall error budget. Therefore, we estimate how large the mass errors would grow when allowing cosmology to vary during the mass estimation process. We do this while keeping the nominal statistical error, which represents the current conservative end of ground-based results. As explained in Appendix C, we find that the total (statistical+cosmological systematic) error, for example, increases from 20% to 40%. That is, considering cosmological systematics in the weak lensing analysis increases the mass error by a factor of 2. Also as detailed in the appendix, the cosmological systematics can be undercut by applying a prior on the Hubble parameter h. We tabulate these results in Table I.

Note that when in this paper we speak of some percentage of "mass scatter" we refer to both the statistical and cosmological systematic error as tabulated in the first

TABLE I.   Cluster parameter uncertainties that make up the F_{prior} matrix used in the various cases considered in the constraint forecasts section as well as the error on the edge $\sigma_{v_{\mathrm{esc}}}$ that makes up the F_{ij} matrix. Note that, in principle, there is no covariance between β and the three other parameters, but we simply change the uncertainty on β for the other cases with reduced weak lensing mass scatter. Furthermore, as explained in the text, we use the WL mass percent error on M_{200} as shorthand to describe uncertainties in all three Einasto parameters. In the last row we tabulate the uncertainties in the three Einasto parameters after applying a prior on h from Ref. [46]. All other cluster parameter uncertainties listed contain both a statistical error as inferred from WL analyses and a systematic error from cosmology as explained in the prior information section and Appendix C.

Cluster parameter uncertainties					
WL mass error	σ_α	$\sigma_{\rho_{-2}}\ [M_\odot/\mathrm{Mpc}^3]$	$\sigma_{r_{-2}}\ [\mathrm{Mpc}]$	σ_β	$\sigma_{v_{\mathrm{esc}}}\ [\mathrm{km\,s}^{-1}]$
5% stat + 5% cosmo sys ("stacked")	0.0024	5.887×10^{12}	0.0314	0.02	90.14
20% stat + 20% cosmo sys ("40% mass scatter")	0.0096	23.589×10^{12}	0.1342	0.5	143.61
40% stat + 40% cosmo sys ("80% mass scatter")	0.0181	43.904×10^{12}	0.2913	0.5	143.61
40% stat ("Riess *et al.* 2016 prior on h")	0.0096	23.589×10^{12}	0.1342	0.5	143.61

three rows of Table I. The last row of Table I shows the uncertainties for a case in which the Hubble prior has been implemented. In that case, the 80% uncertainty on the mass is reduced to 40% after applying the prior on h from Ref. [46] ($\sigma_h = 0.0174$). See Sec. C in the appendix for details.

Note also that when we speak of mass scatter we refer to how the statistical uncertainties in all three cluster parameters change (α, ρ_{-2} and r_{-2}). More specifically, we use the more widely reported percent scatter on M_{200} as a shorthand to describe the uncertainties in our Einasto cluster parameters. As with the fiducial mass, we attain these uncertainties by propagating the 1σ errors on the NFW density profile to the Einasto density profile parameters by fitting the profiles out to one virial radius.

Also unless otherwise stated, and as shown in Table I, for the anisotropy parameter we use an uncertainty of $\sigma_\beta = 0.5$ which is also a fairly conservative choice. For example, see uncertainties on β as derived from a Jeans' analysis in Ref. [23].

These parameter uncertainty priors come into our prior matrix only along the diagonal. This is because we are already considering covariances on the prior information between cluster and cosmological parameters through the aforementioned analysis. Note that this does not mean that we are not considering covariances between parameters. We are; these are all encapsulated in the Fisher matrix F_{ij}. We simply neglect covariance on most prior information. In this sense, our attempt to grapple with the covariances in our prior matrix (detailed in Appendix C) is only an approximation. Note, however, that some parameter uncertainties in the prior matrix are actually nil. For instance, a Jeans equation analysis inference of β is what allows us to decouple the edge uncertainty $\sigma_{v_{\mathrm{esc}}}$, as well as other cluster parameters from σ_β. In contrast, inferring β from edge measurements as done by Ref. [9] would introduce a complete covariance between edge measurements and the anisotropy parameter. The Jeans equation evades this problem given that it is a function of the derivative of the potential rather than the potential itself.

Lastly, the inverse cluster covariance matrices ($C_{\mathrm{cluster}}^{-1}$) contain the aforementioned priors on the mass parameters as well as the prior on the cluster anisotropy parameter. We add covariance between relevant parameters of the cluster prior submatrix. In particular, we model the covariance between cluster parameters r_{-2} and ρ_{-2}. We discuss the structure of the submatrices $C_{\mathrm{cluster}}^{-1}$ in the appendix.

2. Cosmological prior information (C_{cosmo}^{-1})

In relation to the cosmological prior information matrices, we note that for most cases we are not adding any cosmological priors so that in Eq. (22), C_{cosmo}^{-1} is an N_{cosmo} by N_{cosmo} null matrix. Where noted, we add a diagonal prior on the C_{cosmo}^{-1} matrix from the 2.4% determination of the Hubble constant from Ref. [46].

IV. CONSTRAINT FORECASTS

We now derive marginalized uncertainties on cosmological parameter p_{cosmo}, marginalized over the cluster nuisance parameters p_{cluster}. Note that in some cases we also marginalize over remaining cosmological parameters to generate two-dimensional likelihoods.

We consider three cosmological models. The first two assume a flat universe ($\Omega_k = 0$) and the last case assumes a nonflat universe ($\Omega_k \neq 0$).

A. Constant equation of state w

For this case, we consider the following set of cosmological parameters:

$$p_{\mathrm{cosmo}} \in \{\Omega_M, w, h\}. \tag{23}$$

Via Eq. (10) we find the Hubble parameter which yields

$$H^2 = H_0^2 E(z)^2 = H_0^2[\Omega_M(1+z)^3 + \Omega_{\mathrm{DE}}(1+z)^{3(1+w)}] \tag{24}$$
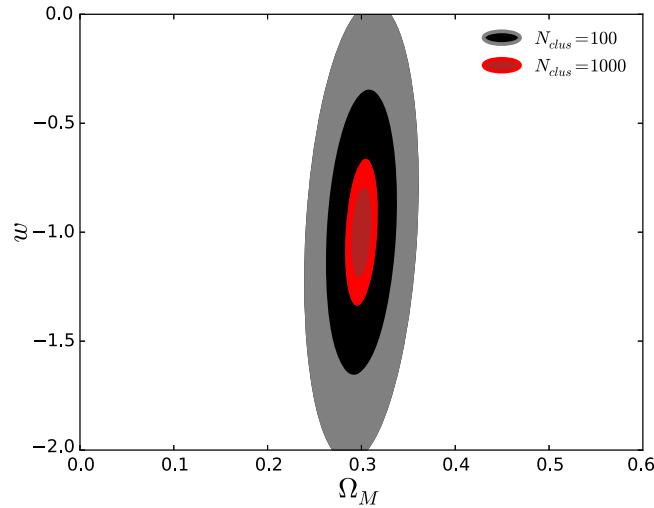
and the deceleration parameter via Eq. (11),

FIG. 5.   68% and 95% confidence constraints for the flat wCDM case after marginalizing over all other parameters. We use $N_{\mathrm{clus}} = 1000(100)$ clusters as shown in red (black) uniformly distributed in the range $0 \leq z \leq 0.8$. The priors on the Einasto parameters assume a uniform 80% weak lensing mass scatter for all redshifts (see Table I). The 1σ errors are $\sigma_{\Omega_M} = 0.007(0.025)$ and $\sigma_w = 0.138(0.431)$ for $N_{\mathrm{clus}} = 1000(100)$.

$$q = \frac{1}{2}[\Omega_M(z) + (1 + 3w)\Omega_{\mathrm{DE}}(z)]. \qquad (25)$$

Here, $\Omega_M(z) = \Omega_M(1 + z)^3 E(z)^{-2}$ and $\Omega_{\mathrm{DE}}(z) = \Omega_{\mathrm{DE}}(1 + z)^{3(1+w)}E(z)^{-2}$. The combination then of the deceleration parameter and Hubble parameter that makes up the equivalent radius and our observable is given by

$$qH^2 = \frac{H_0^2}{2}[\Omega_M(1 + z)^3 + (1 + 3w)(1 - \Omega_M)(1 + z)^{3(1+w)}]. \qquad (26)$$

Notice that the quantity $E(z)$ cancels out in this expression.

The constraints in the $\Omega_M - w$ plane, after marginalizing over h and all other cluster parameters, are shown in Fig. 5. In particular, in black/grey (dark/light red) we plot the 68% and 95% confidence level constraints for a set of $N_{\mathrm{clus}} = 1000(100)$ clusters uniformly distributed in the range $0 \leq z \leq 0.8$. The marginalized 1σ errors are $\sigma_{\Omega_M} = 0.007(0.025)$ and $\sigma_w = 0.138(0.431)$ for $N_{\mathrm{clus}} = 1000(100)$. This is an extraordinarily tight constraint considering that it is achieved by the escape velocity method alone, before adding constraints from other probes.

B. w_0 and w_a

In the previous case we considered a constant dark energy of EoS. However, in principle, the dark energy EoS can evolve with redshift. A popular way to parametrize the redshift-evolving dark energy EoS through the so-called Chevallier-Polarski-Linder (CPL) parametrization is given by (see Refs. [47,48])

$$w(z) = w_0 + w_a \frac{z}{1 + z}. \qquad (27)$$

In this case we consider the following set of cosmological parameters:

$$p_{\mathrm{cosmo}} \in \{\Omega_M, w_0, w_a, h\}. \qquad (28)$$

Again we derive the Hubble parameter for this case,

$$\begin{aligned} H^2 = H_0^2 E(z)^2 = H_0^2 [\Omega_M(1 + z)^3 \\ + \Omega_{\mathrm{DE}}(1 + z)^{3(1+w_0+w_a)} e^{-3w_a \frac{z}{1+z}}], \end{aligned} \qquad (29)$$

and the deceleration parameter

$$q = \frac{1}{2}\left[\Omega_M(z) + \Omega_{\mathrm{DE}}(z)\left(1 + 3w_0 + \frac{3w_a z}{1 + z}\right)\right]. \qquad (30)$$

The redshift-evolving mass and dark energy densities are given by $\Omega_M(z) = \Omega_M(1 + z)^3 E(z)^{-2}$ and $\Omega_{\mathrm{DE}}(z) = \Omega_{\mathrm{DE}}(1 + z)^{3(1+w_0+w_a)} e^{-\frac{3w_a z}{1+z}} E(z)^{-2}$.

Again the redshift-evolving energy density term $E(z)$ cancels and we are left with

$$\begin{aligned} qH^2 = \frac{H_0^2}{2}\Bigg[\Omega_M(1 + z)^3 + (1 - \Omega_M)(1 + z)^{3(1+w_0+w_a)} \\ \times \exp\left\{-\frac{3w_a z}{1 + z}\right\}\left(1 + 3w_0 + \frac{3w_a z}{1 + z}\right)\Bigg]. \end{aligned} \qquad (31)$$

We show the resulting two-dimensional w_0–w_a likelihood in Fig. 6 for a uniform set of clusters in the range $0 \leq z \leq 0.8$, after marginalizing over Ω_M, h, and all other cluster parameters. With $N_{\mathrm{clus}} = 1000$ clusters uniformly distributed between $0 \leq z_c \leq 0.8$ with 80% (40%) weak lensing mass scatter the turquoise (purple) contours express the following marginalized uncertainties: $\sigma_{w_0} = 0.139(0.124)$ and $\sigma_{w_a} = 0.968(0.857)$.

We then reduce the number of clusters to $N_{\mathrm{clus}} = 100$ and apply a cosmological prior on the Hubble constant from Ref. [46] ($\sigma_h = 0.0174$) to yield the following marginalized 1σ errors, $\sigma_{w_0} = 0.191$ and $\sigma_{w_a} = 2.712$ (see pink contours in Fig. 6). Note that the pink contour is made using the same redshift range ($0 \leq z \leq 0.8$) and systematic error (40% weak lensing mass scatter and $\sigma_\beta = 0.5$) as before but with only 100 clusters. Note that this constraint on w_0 is comparable to the constraint achieved with 1000 clusters of Fig. 5 (red contours).

C. Nonflat universe, Ω_M and Ω_Λ

So far we have only considered flat universes in our analysis. We now drop this assumption and assume the possibility of nonzero curvature, while fixing the dark energy equation of state to $w = -1$. We then have the following set of parameters to constrain:
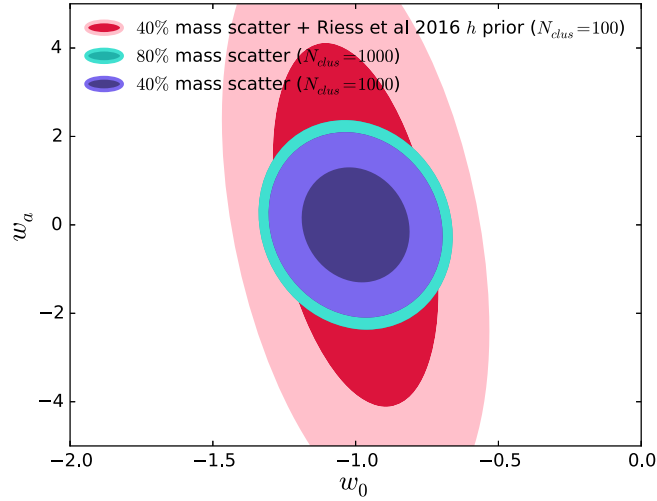
FIG. 6.   68% and 95% contours for the dynamic dark energy case using the CPL parametrization of dark energy marginalized over Ω_M and h as well as the other cluster parameters. More specifically, we use $N_{\text{clus}} = 1000$ clusters uniformly distributed between $0 \leq z_c \leq 0.8$ with 80% (40%) weak lensing mass scatter in the turquoise (purple) contour which yields $\sigma_{w_0} = 0.139(0.124)$ and $\sigma_{w_a} = 0.968(0.857)$. With $N_{\text{clus}} = 100$ uniformly distributed in the same redshift range as before and with 40% mass scatter but now adding a prior on the Hubble constant from Ref. [46] we attain $\sigma_{w_0} = 0.191$ and $\sigma_{w_a} = 2.712$ (pink contours).

$$p_{\text{cosmo}} \in \{\Omega_M, \Omega_\Lambda, h\}. \tag{32}$$

For this case the Hubble parameter is given by

$$H^2 = H_0^2 E(z)^2 = H_0^2[\Omega_M(1+z)^3 + \Omega_\Lambda + \Omega_k(1+z)^2]. \tag{33}$$

And the deceleration parameter is given by

$$q = \frac{1}{2}\Omega_M(z) - \Omega_\Lambda(z). \tag{34}$$

Multiplying these two we have

$$qH^2 = \left[\frac{1}{2}\Omega_M(1+z)^3 - \Omega_\Lambda\right]H_0^2. \tag{35}$$

The constraints in the Ω_M–Ω_Λ plane, after marginalizing over all cluster parameters, are shown in Fig. 7. With just $N_{\text{clus}} = 100$ uniformly distributed between $0 \leq z \leq 0.8$ we can achieve the following marginalized uncertainties $\sigma_{\Omega_M} = 0.101(0.185)$ and $\sigma_{\Omega_\Lambda} = 0.197(0.428)$ after applying the 2.4% level prior on H_0 from Ref. [46] (compare green to black contours).

V. OBSERVATIONAL STRATEGIES

In order for the escape velocity measurements to yield competitive cosmological constraints in both the near term and the far future, several considerations must be taken into account. In this section we particularly focus on how future
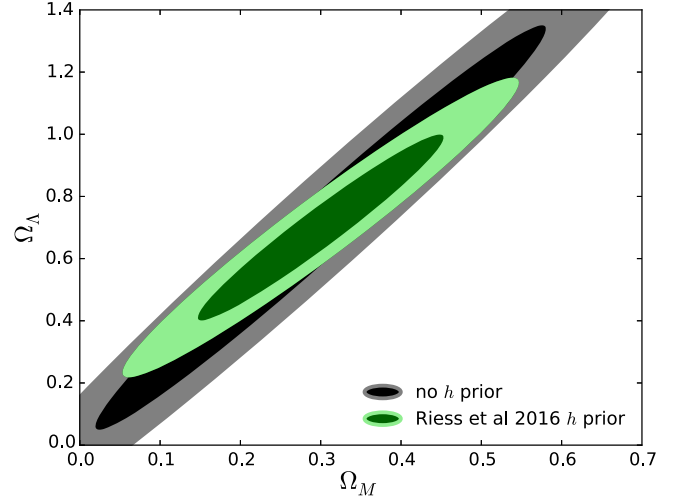

FIG. 7.   68% and 95% contours for the nonflat ΛCDM case for $N_{\text{clus}} = 100$ uniformly distributed between $0 \leq z \leq 0.8$ with 80% mass error. Applying a prior on the Hubble constant h from Ref. [46] ($\sigma_h = 0.0174$) allows us to break the degeneracy and thereby significantly improve our constraints. The marginalized 1σ constraints derived from the green (black) contours are $\sigma_{\Omega_M} = 0.101(0.185)$ and $\sigma_{\Omega_\Lambda} = 0.197(0.428)$.

surveys should target specific redshifts in order to optimize cosmological constraints. We also explore the extent to which reducing systematic uncertainties in both weak lensing mass estimates and measurements of the anisotropy parameter yields significantly better cosmological constraints when compared to simply increasing N_{clus}.

A. Redshift range

We investigate how the cosmological constraints vary with a change of the limits on this redshift distribution. In particular, note that the contours shown in Figs. 5–7 utilize a uniform redshift distribution in the range $0 \leq z \leq 0.8$ for the cluster sample. How would these contours change if we changed the cluster redshift range used?

To quantify the effect, we calculate how the area of the 1σ contours in any given two-dimensional cosmological parameter space changes as a function of the redshift distribution chosen. Mathematically, this entails taking the covariance matrix that contains the marginalized parameters we are interested in and calculating how its determinant changes as a function of the maximum and minimum redshifts for a given distribution. More explicitly, the inverse area corresponding to the 2 by 2 covariance matrix for the marginalized parameters within the 1σ bound is given by (see Refs. [32,33])

$$A^{-1}(p_i, p_j) = \frac{1}{\sqrt{|\det[\text{Cov}(p_i, p_j)]|}}. \tag{36}$$

For the w_0–w_a case the inverse area is the Dark Energy Task Force "figure of merit" [33].
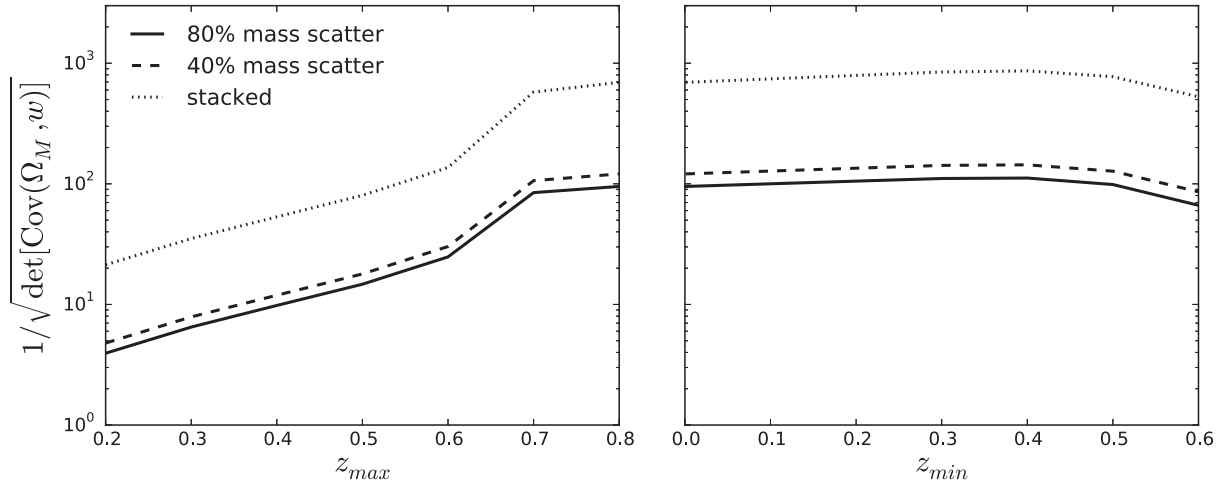
FIG. 8.   Inverse area of the Ω_M-w covariance matrix after marginalizing over all other parameters as a function of maximum (left panel) and minimum (right panel) redshift used. We assume a uniform redshift distribution of clusters in the range $0 \leq z \leq z_{\max}$ (left panel) and $z_{\min} \leq z \leq 0.8$ (right panel). Each calculation of the inverse area assumes a fixed number of clusters, $N_{\mathrm{clus}} = 100$. Note that $\Omega_M - w$ constraints can be improved by reducing the mass uncertainty from 80% (solid line) to 40% (dashed line), in which case the contour area decreases by a factor of ~ 1.3. The constraint can be further increased by utilizing stacked weak lensing mass estimates and stacked phase spaces, and this yields a 10% mass scatter and $\sigma_\beta = 0.02$ (see Table I), decreasing the contour area by a factor of ~ 7.5 (compare solid to dotted lines). The left panel also illustrates that using clusters beyond the transition redshift leads to a gradual loss of cosmological information. While a tighter constraint can be achieved by incorporating higher redshift clusters, the right panel demonstrates that we still need low redshift clusters to achieve the tightest constraints on w. See Fig. 9. We conclude that as broad as possible redshift range of clusters be used (e.g. $0 \leq z \lesssim 0.8$).

The result for the flat wCDM case as a function of redshift range used is shown in Fig. 8. It shows the inverse area for a fixed number of clusters (100 in this case) uniformly distributed in the range $0 \leq z \leq z_{\max}$ (left panel) and $z_{\min} \leq z \leq 0.8$ (right panel).

As the left panel of Fig. 8 implies, we can get the tightest constraints on this cosmological parameter subset by picking 100 clusters uniformly distributed in the range $0 \leq z \leq 0.8$. The physical reason for this can already be inferred from Fig. 4 which shows the derivatives of our observable with respect to the various cosmological parameters. In that figure we note that while our observable is sensitive to Ω_M at high redshifts, it is simultaneously more sensitive to w at low redshifts. Moreover, our probe is relatively more sensitive to Ω_M than to w (compare the absolute maximum of the derivatives). This immediately implies that the higher redshift clusters will end up contributing more to the joint constraint.

However, we emphasize that this does not mean that we should therefore only pick high redshift clusters. In Fig. 9 we plot the marginalized uncertainty for both w (top panel) and Ω_M (bottom panel) as a function of $z_{\min}$. As we pick higher redshift clusters the constraint on Ω_M improves but the constraint on w is degraded. This is also shown by the relatively flat but ultimately decreasing tendency of the right panel in Fig. 8. As such, we emphasize that we need both high and low redshift clusters if we are going to attain a tight constraint on both Ω_M and w. This applies to other constraints as well.
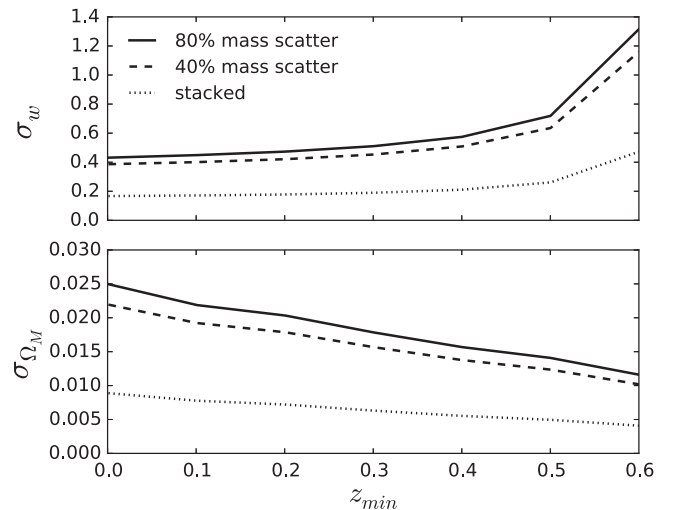

FIG. 9.   Marginalized Ω_M and w uncertainties for the flat wCDM case with 100 clusters distributed in the range $z_{\min} \leq z \leq 0.8$. Note that as the minimum redshift $z_{\min}$ increases, a factor of ~ 1.7 improvement in σ_{Ω_M} is attained (see solid line, bottom panel). The tradeoff is a significant loss of information on the dark energy equation of state parameter w at high redshift (top panel). These effects combine to make the inverse area plot as a function $z_{\min}$ relatively flat (see right-hand side panel in Fig. 8). Note also that a maximum factor of ~ 4 improvement in the uncertainty of these parameters may be achieved if cluster parameter uncertainties are reduced (compare solid to dotted lines).

Secondly, note that this particular optimized choice (i.e. picking clusters uniformly distributed in the range $0 \leq z \lesssim 0.8$) arises from our fiducial cosmology which yields $z_t = 0.671$. Reference [49] for instance finds a transition redshift of $z_t = 0.43 \pm 0.07$ based on a linear parametrization of $q(z)$. Therefore, the inferences of the transition redshift are highly model dependent [50]. What this means is that, observationally, since the optimization of our probe is loosely based on the transition redshift, we recommend that a redshift distribution as broad as possible be used. In particular, we recommend that clusters uniformly distributed in the range $0 \leq z \leq 0.8$ be used. Picking this redshift range allows us to safely take into account current uncertainties in the transition redshift.

Lastly, note also that this upper limit ($z = 0.8$) is also set by the particular processes of cluster assembly. In other words, beyond this redshift our analytic model is unable to take into account the full complexity of cluster-formation dynamics because clusters are still assembling at that redshift for acceptable cosmologies (see Fig. 2 in Ref. [51]).

B. Reducing systematics and stacked clusters

We now study the effects of reducing statistical errors on the cluster parameters. Clearly, reducing statistical errors in the weak lensing mass estimates, in the inference of the anisotropy parameter, and in the measurement of the edge, improves our cosmological constraints, but by how much? For this exercise, we consider increased precision from better measurements on individual clusters, increased cluster sample sizes, and galaxy cluster stacking techniques. We note that stacking is not necessarily equivalent to averaging over a large sample. For additional information on stacking, we refer the reader to detailed analyses of stacking weak lensing data and stacking phase spaces [45,52]. To quantify the improvements, we use Eq. (36) with a fiducial sample of $N_{\rm clus} = 100$ and focus on the $\Omega_M - w$ case.

Figure 8 shows how constraints may be improved by decreasing the scatter on the mass parameters from 80% (solid black line) to 40% (dashed black line). The difference in the inverse area size is a relatively modest factor of ~1.3. However, simultaneously reducing the uncertainty of the mass parameters to 10% as well as reducing the uncertainty of the anisotropy parameter σ_β and the uncertainty on the escape velocity edge $\sigma_{\rm vesc}$ yields an area that is ~7.5 times smaller (compare solid lines to dotted lines). For the exact values of the uncertainties used in our matrix for this stacked case see Table I. Figure 9 follows and demonstrates how each specific marginalized error (on w and Ω_M) varies with $z_{\rm min}$ as we change the error on the cluster priors. Looking at Fig. 9, an improvement of a factor of ~2–4 on both σ_M and σ_w may be achieved with decreased uncertainties (compare solid to dotted lines).

Currently, the only way to attain uncertainties in the mass parameters of the smallest order in Figs. 8 and 9 requires a stacking analysis. For example, see cosmological constraints
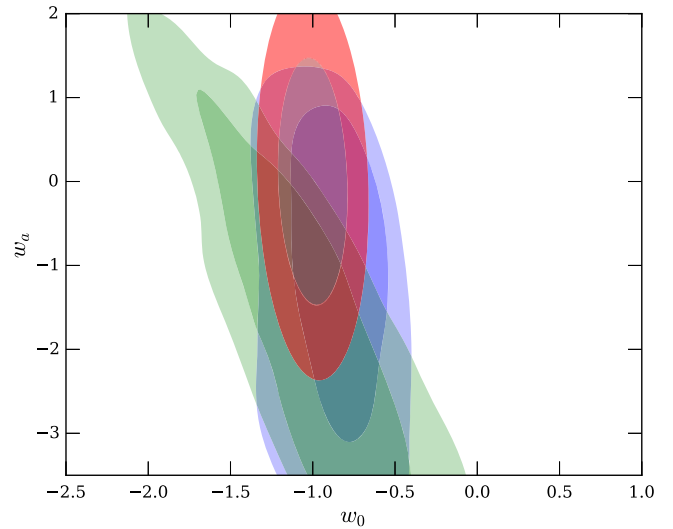

FIG. 10. 68% and 95% contours for the dynamic dark energy case using the CPL parametrization of dark energy marginalized over Ω_M and h as well as the other cluster parameters. The blue contours are reproduced from the latest JLA Type Ia supernovae (SNIa) constraints of Ref. [58]. The green contours are reproduced from the Planck 2015 TT likelihood of Ref. [59]. In red are constraints derived from a sample of $N_{\rm clus} = 1000$ clusters uniformly distributed in the redshift range $0 \leq z \leq 0.8$, after applying a conservative 80% mass scatter prior (same as the turquoise contours of Fig. 6). In all cases, no prior assumptions about the Hubble constant are used.

derived from a weak lensing analysis in Ref. [45] as well as stacked phase-space analyses in Ref. [52]. Similarly, achieving $\sigma_\beta = 0.02$ entails stacking clusters and/or developing some other approach that has not yet been fully investigated. Thus, the dotted line in Fig. 8 represents not 100 clusters, but 100 cluster ensembles with high-precision mean masses and mean β's. Each cluster ensemble is built from a number of individual noisy weak lensing cluster profiles and poorly sampled cluster phase spaces. One thing to consider in a future stacked phase-space analysis is that systematic uncertainties (e.g., cluster miscentering) must be accounted for at high precision.

From an observational perspective, it is an interesting question whether one should expend resources on increasing the sample size, or on decreasing the systematic uncertainties. Given a Planck cosmology, to $z = 0.8$ we expect to have over 40,000 clusters with $M_{200} > 4 \times 10^{14}\ M_\odot$ with respect to 200× the mean density of the Universe [53]. Thus, it seems reasonable to expect that 1000 of these clusters will eventually have both weak lensing mass estimates and well-sampled radius/velocity phase spaces. Such an effort would require photometry and spectroscopy over about 1000 square degrees of the sky. As an example, the Dark Energy Spectroscopic Instrument (DESI [54]) is targeting over 1000 square degrees of the Dark Energy Survey sky coverage [55]. Likewise, the Prime Focus Camera will provide significant multiobject spectroscopy over more than 1000
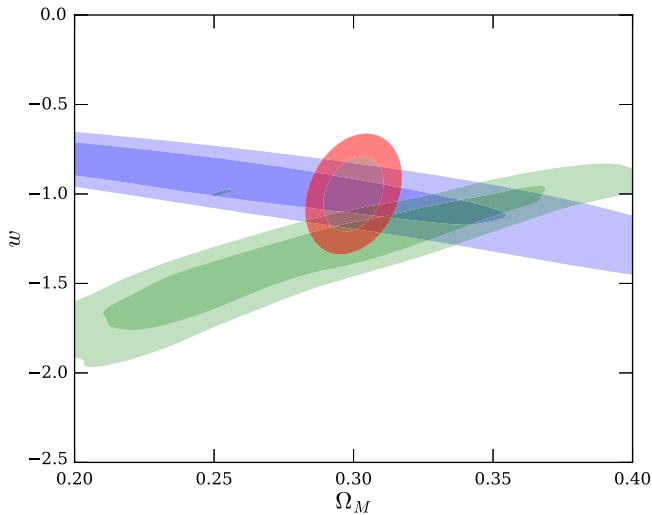
FIG. 11.   68% and 95% contours for the flat wCDM case. The red contours are derived the same as what is shown in Fig. 5. The blue contours are reproduced from the latest JLA SNIa constraints shown in Ref. [58]. The green contours are reproduced from the Planck 2015 TT likelihood [59]. In all cases, no prior assumptions about the Hubble constant are used.

square degrees of imaging taken with the Subaru Hyper Suprime-Cam [56]. There is also new Principal Investigator-based (PI-based) instrumentation, such as the Michigan-Magellan Fiber System on the Magellan observatory, which can be used to specifically and efficiently target clusters with previously measured weak lensing masses [57]. Thus, it is realistic to expect 1000 clusters with densely sampled phase spaces and redshift errors ~ 50 km s^{-1} and weak lensing mass errors of 40% or less (statistical) in the near future. The technology to collect the needed spectra either exists or is being constructed with the aim to achieve redshift errors on par with existing surveys [36] and the weak lensing mass errors of 40% or less are already being achieved with current imagers [41,43]. Therefore, in Figs. 5, 10 and 11 we show the constraints after increasing the fiducial sample size from 100 to 1000 clusters.

Lastly, it should be obvious that a combination of both more clusters and reduced systematic error would be the optimal solution which yields the tightest constraints. Our analysis in this section is premised on the assumption that both of these options may not be easily available.

VI. DISCUSSION

Other than by reducing statistical errors, increasing the number of clusters, and stacking, we may, in principle, improve the constraints through a joint likelihood analysis with other cosmological probes. In this section we discuss our constraints and their degeneracies in the context of other probes.

We note that in Fig. 7 an improvement in the forecasted constraints can be achieved after applying the prior on h from Ref. [46]. Information on the Hubble parameter

breaks numerous degeneracies in our probe. After all, our technique itself is fundamentally based on constraining qH^2 (see Sec. IV A). However, by including cosmological dependencies on the radial coordinate [Eq. (2)] the probe clearly has some power in constraining h on its own. This is evident in the derivatives shown in Fig. 4. If we drop this dependence and use a fixed (in Mpc) radial coordinate, Ω_Λ, w_0, and w_a all become entirely unconstrained.

Besides applying a prior on the Hubble constant, another way we may achieve a tight constraint of the w_0–w_a plane is shown in Fig. 10. We show both the 68% and 95% confidence constraints with $N_{\rm clus} = 1000$ uniformly distributed in the redshift range $0 \leq z \leq 0.8$, after applying a conservative 80% mass scatter prior (in red, the same as the turquoise contours in Fig. 6) as well as both the JLA SNIa constraints of Ref. [58] (in blue) and the 2015 Planck TT likelihood constraints of Ref. [59] (in green). A joint analysis with these probes then seems to have the potential of yielding similar constraints to what a combination of JLA data and cosmic microwave background (CMB) currently yields.

Now considering the flat wCDM case, in Fig. 11 we overplot the JLA constraints (in blue), the 2015 Planck TT likelihood of Ref. [59] (in green) as well as replot the red contours of Fig. 5. We find that a joint constraint of these probes alone can yield a joint $\sigma_w \sim 0.1$ and $\sigma_{\Omega_M} \sim 0.01$ constraint given that they cross through each other nearly perpendicularly.

We note that the degeneracies in our observable work out in such a way that our probe can act as a powerful cross-check of systematics in other probes. As an example, note how our constraints lie perpendicular to the JLA SNIa and CMB constraints in Fig. 11. In part, this is due to the way degeneracies work with our observable.

As shown in Sec. II, cosmology in our probe comes in through the quantity qH^2, which is both a function of the Hubble parameter and its derivative, $dH(z)/dz$. We emphasize this because our probe in this sense is a true dynamical probe of the expansion history of the Universe, similar to the redshift drift. Compare for instance our constraint degeneracies on the w_0–w_a plane to those of Ref. [60] (in particular, see Fig. 3).

Beyond these comparisons to other probes, we may ask ourselves if we can justify our observable's sensitivity to cosmology on its own basis. As discussed before, the sensitivity of our probe can be directly inferred from both Figs. 4 and 3. Recall that in Fig. 3 we plot the fractional difference of the escape velocity profile between the ΛCDM model ($w = -1$) and two other dark energy models. At about the virial radius, the fractional difference amounts to ~ 15%. So to a crude first approximation, we need the error budget in the observational parameters to drop below this limit in order to place constraints on w.

As explained before, and as tabulated in Table I, three sources of observational error are involved: the escape velocity edge error ($\sigma_{v_{\rm esc}}$), the anisotropy parameter

uncertainty (σ_β) and the error in the inferred Einasto parameters from weak lensing (α, $\sigma_{\rho_{-2}}$, $\sigma_{r_{-2}}$). If we focus just on the $\sigma_{v_{\mathrm{esc}}}$ we notice that this amounts to a $\sim$15% error on the escape velocity edge. As such, with just one cluster we are on the verge of being able to detect deviations from the ΛCDM model. Similarly, for our fiducial cluster, the uncertainty in $\sigma_\beta = 0.5$ amounts to a difference in the escape velocity profile also of $\sim$15% given that it comes in to our observable through the factor of $1/\sqrt{g(\beta)}$ [see Eq. (4)]. So again, with this systematic uncertainty we are close to being able to detect deviations from $w = -1$. Now let us consider the dominant source of error which comes in through the uncertainty in the inferred Einasto parameters from weak lensing. For 80% error on the mass, this amounts to an error on the edge of $\sim$40% given that the velocity goes as the square root of the mass. With just eight clusters we can naively decrease the weak lensing error to $\sim$15%, assuming that it scales as $1/\sqrt{N_{\mathrm{clus}}}$. Of course the above are unrealistic conditions, which is why instead we conduct a detailed Fisher matrix analysis.

Another key factor in this probe is that unlike supernovae observations, the cluster data map is a projected radial profile which increases the total amount of information per object, thereby further beating down the error. We previously addressed how the binning can affect the final predictions on the cosmological parameters. The key point is that with just a few tens of clusters, this probe becomes sensitive to 15% deviations in the dark energy equation of state exemplified in Fig. 3. While these are only forecasts, as a consistency check, we have compared our Fisher matrix constraints with the analysis of Ref. [9], which utilized $N_{\mathrm{clus}} = 20$ ($0 \leq z \leq 0.439$). We find that our Fisher matrix forecasts are consistent with variations of systematics studied in Ref. [9].

VII. CONCLUSIONS

We have presented a novel galaxy cluster-based probe of cosmology that has the potential of constraining cosmological parameters to high precision. More specifically, this cosmological probe is based on both the abstract and concrete need to include a cosmological term in the escape velocity profile of galaxy clusters as inferred from their phase spaces. This phase-space-inferred escape velocity profile is modeled by cluster-specific parameters (i.e. weak lensing mass profile information and the cluster's anisotropy parameter) as well as cosmological parameters. If the first set of parameters can be independently inferred, then cosmology can be allowed to vary to fit the observed escape velocity profiles—thereby constraining cosmological models.

To assess this probe's observational viability we used the Fisher matrix formalism and carefully considered the aforementioned systematics by marginalizing over the free parameters describing the gravitational potential of each cluster separately. While constraints can be improved if systematic errors in both the weak lensing mass estimates and inferences of the anisotropy parameter are reduced, we note that the gains are similarly improved by increasing the number of clusters N_{clus}. A combination of both of these approaches would of course be optimal. However, we note that even assuming conservative errors, competitive cosmological constraints can still be achieved in the near term.

It is also important to realize, as mentioned above, that this probe can currently constrain only accelerating cosmologies. In particular, the balance of forces argument with which we derive our theoretical expectation for the escape velocity profile is not valid in a nonaccelerating universe. More theoretical work needs to be done to be able to make the theoretical expectation sensitive to cosmology at epochs beyond the transition redshift.

Nonetheless, we have shown that this probe is not only able to yield high-precision constraints on cosmological parameters independently of other probes but that it complements other constraints as well. Furthermore, we emphasize that these constraints can be achieved in both the near and far future. For instance, Figs. 6 and 7 only use 100 clusters with 40%–80% weak lensing mass scatter which can easily be achieved in the near term; this is also the case with the black contour constraints of Fig. 5. Far future constraints ($N_{\mathrm{clus}} = 1000$) are forecasted in Fig. 5 (red contours) as well as in the future joint constraints of Figs. 10 and 11.

We also note that while throughout this paper we have assumed the general relativistic Friedmann equation, our theoretical expectations can, in principle, be generalized by reworking the term qH^2 to either reflect modified theories of gravity or other alternative parametrizations.

This work therefore presents a first step in the study of a promising new probe of cosmology. The cluster phase spaces we demonstrated here have the power to provide precision measurements of cosmological parameters in an accelerating universe, and thus provide sharp tests of the currently favored theoretical framework.

APPENDIX A: F_{ij} MATRIX STRUCTURE

In this section we detail the structure of the F_{ij} matrix. A schematic of the matrix and its submatrices is shown in Fig. 12. As indicated by the figure, there are three main component submatrices to the F_{ij} matrix: the cosmo-cosmo

$$
F_{ij} = \begin{pmatrix}
\boxed{\text{cosmo - cosmo}} & \boxed{\text{cluster 1-cosmo}}^T & \cdots & \boxed{\text{cluster N - cosmo}}^T \\
\boxed{\text{cluster 1 - cosmo}} & \boxed{\text{cluster 1- cluster 1}} & 0 & 0 \\
\vdots & 0 & \ddots & 0 \\
\boxed{\text{cluster N - cosmo}} & 0 & 0 & \boxed{\text{cluster N- cluster N}}
\end{pmatrix}
$$

FIG. 12.   Structure of the $N_{\text{dim}} \times N_{\text{dim}}$ [see Eq. (19)] F_{ij} matrix. The "cosmo-cosmo" submatrix (orange) has dimensions $N_{\text{cosmo}} \times N_{\text{cosmo}}$ and contains the information solely of the cosmological parameters and their inverse covariances. The "cosmo-cluster" submatrices (red) lie along the first row and column of the F_{ij} matrix and are composed of the cross-correlated information between cluster parameters and cosmological parameters. Note that along the first column these submatrices have dimension $N_{\text{clus}} \times N_{\text{cosmo}}$, and along the first row, the matrices are transposed and therefore have dimensions $N_{\text{cosmo}} \times N_{\text{clus}}$. Lastly, the "cluster-cluster" submatrices (green) lie along the diagonal and have dimensions $N_{\text{clus}} \times N_{\text{clus}}$. Note that 0's populate the off-diagonal spaces given that there is no correlation between clusters, so that the derivatives cancel out.

submatrix (orange), the cluster-cosmo submatrices (red), and the cluster-cluster submatrices (green). The 0's indicate that the off-diagonal terms are nil. We describe the components of these submatrices in the following three subsections.

1. Cosmo-cosmo submatrix

In the top left of the matrix (see Fig. 12) we have an N_{cosmo} by N_{cosmo} submatrix which is composed exclusively of the derivatives of our observable with respect to cosmological parameters. Let us consider the flat wCDM case, $p_{\text{cosmo}} \in \{\Omega_M, w, h\}$, and take a look at some terms. For this case, the first term in this submatrix is then

$$
F_{00} = \frac{1}{\sigma_{v_{\text{esc}}}^2} \sum_{n,k} \frac{\partial v_{\text{esc}}(z_n, r_k)}{\partial \Omega_M} \frac{\partial v_{\text{esc}}(z_n, r_k)}{\partial \Omega_M}. \quad \text{(A1)}
$$

The off-diagonal term in the first column, second row is simply

$$
F_{01} = \frac{1}{\sigma_{v_{\text{esc}}}^2} \sum_{n,k} \frac{\partial v_{\text{esc}}(z_n, r_k)}{\partial \Omega_M} \frac{\partial v_{\text{esc}}(z_n, r_k)}{\partial w}. \quad \text{(A2)}
$$

Therefore, we are adding information on some cosmological parameter (or a combination, as in the off-diagonal term) both across n clusters and k radial bins. As we detail in the next two sections, this is not the case for all other elements in the F_{ij} matrix.

2. Cosmo-cluster submatrices

Now let us take a look at the cosmo-cluster submatrices of Fig. 12 (shown in red). Staying in the first row but now looking at the fourth column, we are now looking at the cross information attained from cosmology and cluster parameters. In this case, the anisotropy parameter for cluster 1 (β_1), the matrix element, is

$$
\begin{aligned}
F_{03} &= \frac{1}{\sigma_{v_{esc}}^2} \sum_{n,k} \frac{\partial v_{esc}(z_n, r_k)}{\partial \Omega_M} \frac{\partial v_{esc}(z_n, r_k)}{\partial \beta_1} \\
&= \frac{1}{\sigma_{v_{esc}}^2} \sum_{k} \frac{\partial v_{esc}(z_1, r_k)}{\partial \Omega_M} \frac{\partial v_{esc}(z_1, r_k)}{\partial \beta_1} + \\
&\quad \frac{1}{\sigma_{v_{esc}}^2} \sum_{k} \frac{\partial v_{esc}(z_2, r_k)}{\partial \Omega_M} \cancelto{0}{\frac{\partial v_{esc}(z_2, r_k)}{\partial \beta_1}} + \cdots
\end{aligned}
$$
$$\text{(A3)}$$

Immediately we notice that the second term of the second sum (i.e. the derivative with respect to β_1 for the second cluster z_2) is nil. Therefore, unlike the cosmo-cosmo submatrices, in these submatrices we only sum over the kth radial bin of the cluster corresponding to that column. This is the case for subsequent columns and rows (by symmetry). For instance, if $N_{\text{cosmo}} = 3$ then F_{0j}, F_{1j}, F_{2j} where $j = \{3, 4, 5, \ldots, N_{\text{dim}}\}$. Symmetry yields the same for F_{i0}, F_{i1}, F_{i2} where $i = \{3, 4, 5, \ldots, N_{\text{dim}}\}$. The structure is the same for the cluster-cosmo submatrices along the first column, where the submatrices are simply transposed, as evoked by the superscript T in Fig. 12.

3. Cluster-cluster submatrices

Lastly, let us now take a look at the cluster-cluster submatrices of Fig. 12 (green). These submatrices express simply the covariance between cluster parameters for a given cluster. Taking a look at the first element of the first submatrix on the diagonal (the "cluster 1-cluster 1" submatrix of Fig. 12),

$$
\begin{aligned}
F_{33} &= \frac{1}{\sigma_{v_{esc}}^2} \sum_{n,k} \frac{\partial v_{esc}(z_n, r_k)}{\partial \beta_1} \frac{\partial v_{esc}(z_n, r_k)}{\partial \beta_1} \\
&= \frac{1}{\sigma_{v_{esc}}^2} \sum_{k} \frac{\partial v_{esc}(z_1, r_k)}{\partial \beta_1} \frac{\partial v_{esc}(z_1, r_k)}{\partial \beta_1} + \\
&\quad \frac{1}{\sigma_{v_{esc}}^2} \sum_{k} \cancelto{0}{\frac{\partial v_{esc}(z_2, r_k)}{\partial \beta_1}} \frac{\partial v_{esc}(z_2, r_k)}{\partial \beta_1} + \cdots
\end{aligned}
$$
$$\text{(A4)}$$

We see that the second term and on yield 0 given that they are the derivatives of clusters $z_{n \neq 1}$ with respect to β_1. This simply demonstrates that there is no cross correlation between clusters, as expected. Therefore, along the diagonal of F_{ij} we have 4×4 matrices of the various cluster parameters with respect to a given cluster, from cluster 1 to cluster N_{dim}.

Let us take a look now at some of the off-diagonal submatrices, say between cluster 1 and cluster 2. The first element is

$$F_{73} = \frac{1}{\sigma_{v_{esc}}^2} \sum_{n,k} \frac{\partial v_{esc}(z_n, r_k)}{\partial \beta_2} \frac{\partial v_{esc}(z_n, r_k)}{\partial \beta_1}$$

$$= \frac{1}{\sigma_{v_{esc}}^2} \sum_{k} \frac{\partial v_{esc}(z_1, r_k)}{\partial \beta_2} \frac{\partial v_{esc}(z_1, r_k)}{\partial \beta_1} +$$

$$\frac{1}{\sigma_{v_{esc}}^2} \sum_{k} \frac{\partial v_{esc}(z_2, r_k)}{\partial \beta_2} \frac{\partial v_{esc}(z_2, r_k)}{\partial \beta_1} + \dots$$

$$\text{(A5)}$$

Note that the first term in the sum over the radial bins of the first cluster ($n = 1$) is nil, and so is the second term of the sum over the radial bins of the second cluster ($n = 2$). By induction, all other terms are also nil. Therefore, these off-diagonal terms are all 0 given that there is no cross correlation between cluster parameters of different clusters. All of these terms are aptly represented by 0's in Fig. 12.

APPENDIX B: F_{prior} MATRIX SUBSTRUCTURE

In this section we describe the structure of the prior matrix found in Eq. (16). In particular, we focus on the structure of the submatrix elements of the F_{prior} matrix in Eq. (22). The covariance submatrices that lie along the diagonal of Eq. (22) are given by

$$C_{\text{cluster}} = \begin{bmatrix} \sigma_\beta^2 & 0 & 0 & 0 \\ 0 & \sigma_\alpha^2 & 0 & 0 \\ 0 & 0 & \sigma_{r_{-2}}^2 & -0.7\sigma_{\rho_{-2}}\sigma_{r_{-2}} \\ 0 & 0 & -0.7\sigma_{\rho_{-2}}\sigma_{r_{-2}} & \sigma_{\rho_{-2}}^2 \end{bmatrix}. \quad \text{(B1)}$$

Note that the only nonzero term off the diagonal is the covariance between r_{-2} and ρ_{-2}. Specific values for these matrix elements and the code used to produce the matrices from which we derive the constraints on cosmological parameters can be found online at [61].

APPENDIX C: WEAK LENSING MASS PRIOR AND COSMOLOGY

We now consider how uncertainties in the cosmological parameters affect the weak lensing mass uncertainties,
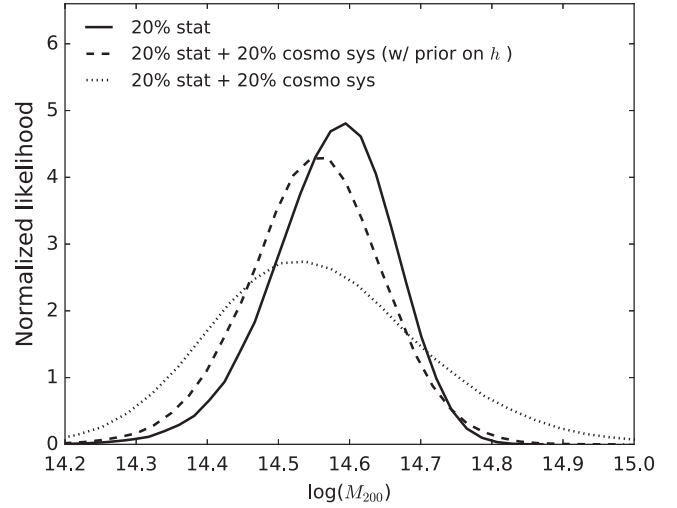

FIG. 13. Marginalized likelihood of the inferred mass of our fiducial cluster ($M_{200} = 4 \times 10^{14} M_\odot$) from the weak lensing surface density shear profile $\Sigma(r)$. The total uncertainty in the inferred mass increases by a factor of ~ 2 if no prior on the cosmological parameters is introduced; however, a reasonable prior on the Hubble parameter h from Ref. [46] recovers most of the lost information on M_{200}.

which, recall, are in turn featured in our prior information matrix C_{cluster}.

To do this, we carry out a quantitative investigation utilizing the Cluster-Lensing code of Ref. [62]. We start out by building a $\Sigma(r)$ surface density shear profile for one fiducial cluster with $M_{200} = 4 \times 10^{14} M_\odot$. We first create this profile assuming fixed, fiducial values of the cosmological parameters. We assume the profile has Gaussian errors of such size to ensure that we recover a 20% statistical error on the cluster mass after performing a simple χ^2 analysis.

We then allow the cosmological parameters to vary, and conduct a Markov chain Monte Carlo analysis with *emcee* to sample the posterior distribution and examine the likelihood of the inferred mass M_{200} [63]. The likelihood model is given by

$$\ln \mathcal{L}(\Sigma | r_k, z, \Theta) = -\frac{1}{2} \sum_k \frac{(\Sigma(r_k, z, \Theta_{\text{fid}}) - \Sigma(r_k, z, \Theta))^2}{\sigma_\Sigma^2}.$$

$$\text{(C1)}$$

Assuming the flat wCDM model, our parameter set is given by $\Theta = \{\Omega_M, w, h, M_{200}\}$. Note that the cluster profile information is reduced to a single parameter, M_{200}, because the Cluster-Lensing code uses a mass-concentration relation to create the $\Sigma(r)$ profile [62]. For a single mock $\Sigma(r)$ profile, after marginalizing over all other parameters we find that the total error in the cluster mass scale increases from 20% to 40%. That is, if cosmological parameters are allowed to vary, the weak lensing mass error increases by a factor of 2. We confirm this using statistical errors of 5% and also 40%.

We show the marginalized posterior likelihood in Fig. 13. Three different prior likelihoods are shown: a

strong prior, basically fixing cosmology (represented by the solid black line), a prior only on h (dashed line), and no prior (dotted line). Note how the 20% systematic error likelihood is broadened if cosmological parameters are allowed to vary, but that we can almost totally reduce the cosmological systematic error simply by applying a reasonable prior on the Hubble parameter

of Ref. [46] ($\sigma_h = 0.0174$). We tabulate the resulting uncertainties in the Einasto cluster parameters in Table I. As stated in the previous subsection, the code used to produce these results can be found online. Lastly, we note that since the bias in the recovered mass (dotted and dashed) is small, we do not factor it into our analysis.

[1] F. Zwicky, Astrophys. J. **86**, 217 (1937).
[2] S. A. Shectman, Astrophys. J. **262**, 9 (1982).
[3] E. Regos and M. J. Geller, Astrophys. J. **98**, 755 (1989).
[4] A. Diaferio and M. J. Geller, Astrophys. J. **481**, 633 (1997).
[5] A. Diaferio, Mon. Not. R. Astron. Soc. **309**, 610 (1999).
[6] M. J. Geller, A. Diaferio, K. J. Rines, and A. L. Serra, Astrophys. J. **764**, 58 (2013).
[7] K. Rines, A. Diaferio, and P. Natarajan, Astrophys. J. Lett. **679**, L1 (2008).
[8] E. Regös, Astron. Astrophys. **308**, 368 (1996).
[9] A. Stark, C. J. Miller, and D. Gifford, Astrophys. J. **830**, 109 (2016).
[10] C. J. Miller, A. Stark, D. Gifford, and N. Kern, Astrophys. J. **822**, 41 (2016).
[11] A. Diaferio and M. J. Geller, Astrophys. J. **481**, 633 (1997).
[12] A. L. Serra, A. Diaferio, G. Murante, and S. Borgani, Mon. Not. R. Astron. Soc. **412**, 800 (2011).
[13] D. Lemze, T. Broadhurst, Y. Rephaeli, R. Barkana, and K. Umetsu, Astrophys. J. **701**, 1336 (2009).
[14] A. Stark, C. J. Miller, N. Kern, D. Gifford, G.-B. Zhao, B. Li, K. Koyama, and R. C. Nichol, Phys. Rev. D **93**, 084036 (2016).
[15] M. T. Busha, F. C. Adams, R. H. Wechsler, and A. E. Evrard, Astrophys. J. **596**, 713 (2003).
[16] P. S. Behroozi, A. Loeb, and R. H. Wechsler, J. Cosmol. Astropart. Phys. 6 (2013) 019.
[17] A. Faltenbacher, A. V. Kravtsov, D. Nagai, and S. Gottlöber, Mon. Not. R. Astron. Soc. **358**, 139 (2005).
[18] R. Nandra, A. N. Lasenby, and M. P. Hobson, Mon. Not. R. Astron. Soc. **422**, 2931 (2012).
[19] D. Lemze et al., Astrophys. J. **776**, 91 (2013).
[20] E. Munari, A. Biviano, S. Borgani, G. Murante, and D. Fabjan, Mon. Not. R. Astron. Soc. **430**, 2638 (2013).
[21] E. Retana-Montenegro, E. van Hese, G. Gentile, M. Baes, and F. Frutos-Alfaro, Astron. Astrophys. **540**, A70 (2012).
[22] F. Iannuzzi and K. Dolag, Mon. Not. R. Astron. Soc. **427**, 1024 (2012).
[23] E. L. Łokas, R. Wojtak, S. Gottlöber, G. A. Mamon, and F. Prada, Mon. Not. R. Astron. Soc. **367**, 1463 (2006).
[24] L. Benatov, K. Rines, P. Natarajan, A. Kravtsov, and D. Nagai, Mon. Not. R. Astron. Soc. **370**, 427 (2006).
[25] R. Wojtak and E. L. Łokas, Mon. Not. R. Astron. Soc. **408**, 2442 (2010).
[26] A. Biviano et al., Astron. Astrophys. **558**, A1 (2013).
[27] S. Tsujikawa, Classical Quantum Gravity **30**, 214003 (2013).

[28] R. R. Caldwell, Phys. Lett. B **545**, 23 (2002).
[29] M. Tegmark, A. N. Taylor, and A. F. Heavens, Astrophys. J. **480**, 22 (1997).
[30] A. Heavens, Entropy **18**, 236 (2016).
[31] D. Coe, arXiv:0906.4123.
[32] D. Huterer and M. S. Turner, Phys. Rev. D **64**, 123527 (2001).
[33] A. Albrecht, G. Bernstein, R. Cahn, W. L. Freedman, J. Hewitt, W. Hu, J. Huth, M. Kamionkowski, E. W. Kolb, L. Knox, J. C. Mather, S. Staggs, and N. B. Suntzeff, arXiv:0609591.
[34] M. Sereno, Mon. Not. R. Astron. Soc. **450**, 3665 (2015).
[35] M. Sereno, C. Fedeli, and L. Moscardini, J. Cosmol. Astropart. Phys. 1 (2016) 042.
[36] A. S. Bolton et al., Astrophys. J. **144**, 144 (2012).
[37] D. Gifford, C. Miller, and N. Kern, Astrophys. J. **773**, 116 (2013).
[38] D. Gifford and C. J. Miller, Astrophys. J. Lett. **768**, L32 (2013).
[39] F. W. High et al., Astrophys. J. **723**, 1736 (2010).
[40] D. R. Semler et al., Astrophys. J. **761**, 183 (2012).
[41] D. E. Applegate, A. von der Linden, P. L. Kelly, M. T. Allen, S. W. Allen, P. R. Burchat, D. L. Burke, H. Ebeling, A. Mantz, and R. G. Morris, Mon. Not. R. Astron. Soc. **439**, 48 (2014).
[42] K. Umetsu et al., Astrophys. J. **795**, 163 (2014).
[43] P. Melchior et al., arXiv:1610.06890 [Mon. Not. R. Astron. Soc. (to be published)].
[44] M. R. Becker and A. V. Kravtsov, Astrophys. J. **740**, 25 (2011).
[45] E. Rozo, H.-Y. Wu, and F. Schmidt, Astrophys. J. **735**, 118 (2011).
[46] A. G. Riess, L. M. Macri, S. L. Hoffmann, D. Scolnic, S. Casertano, A. V. Filippenko, B. E. Tucker, M. J. Reid, D. O. Jones, J. M. Silverman, R. Chornock, P. Challis, W. Yuan, P. J. Brown, and R. J. Foley, Astrophys. J. **826**, 56 (2016).
[47] M. Chevallier and D. Polarski, Int. J. Mod. Phys. D **D10**, 213 (2001).
[48] E. V. Linder, Phys. Rev. Lett. **90**, 091301 (2003).
[49] A. G. Riess et al., Astrophys. J. **659**, 98 (2007).
[50] C. Shapiro and M. S. Turner, Astrophys. J. **649**, 563 (2006).
[51] D. E. Holz and S. Perlmutter, Astrophys. J. Lett. **755**, L36 (2012).
[52] D. Gifford, N. Kern, and C. J. Miller, Astrophys. J. **834**, 204 (2017).

[53] P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett *et al.* (Planck Collaboration), Astron. Astrophys. **594**, A13 (2016).

[54] http://desi.lbl.gov/.

[55] A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen, C. Allende Prieto, J. Annis, S. Bailey, C. Balland *et al.* (DESI Collaboration), arXiv:1611.00037.

[56] M. Takada *et al.*, Publ. Astron. Soc. Jpn. **66**, R1 (2014).

[57] M. Mateo, J. I. Bailey, J. Crane, S. Shectman, I. Thompson, I. Roederer, B. Bigelow, and S. Gunnels, Proceedings / AIAA/ASME/SAE Structures, Structural Dynamics, and Materials Conference ; sponsored by American Institute of

Aeronautics and Astronautics (AIAA), American Society of Mechanical Engi **8446**, 84464Y (2012).

[58] M. Betoule *et al.*, Astron. Astrophys. **568**, A22 (2014).

[59] P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett *et al.* (Planck Collaboration), Astron. Astrophys. **594**, A13 (2016).

[60] A. G. Kim, E. V. Linder, J. Edelstein, and D. Erskine, Astropart. Phys. **62**, 195 (2015).

[61] https://github.com/alejostark/phase_space_cosmo_fisher.

[62] J. Ford and J. VanderPlas, Astron. J. **152**, 228 (2016).

[63] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, PASP **125**, 306 (2013).