Astro2020 Science White Paper

Deep Multi-object Spectroscopy to Enhance Dark Energy Science from LSST

Thematic Areas: Cosmology and Fundamental Physics

Principal (corresponding) author:
Name: Jeffrey A. Newman
Institution: University of Pittsburgh and PITT PACC
Email: janewman@pitt.edu
Phone: (412) 592-3853

Co-authors: J. Blazek (SNSF Ambizione, EPFL; CCAPP, Ohio State U.), N. E. Chisari (Oxford), D. Clowe (Ohio U.), I. Dell'Antonio (Brown), E. Gawiser (Rutgers), R. A. Hložek (Toronto), A. G. Kim (LBNL), A. von der Linden (Stony Brook University), M. Lochner (African Institute for Mathematical Sciences, South African Radio Astronomy Observatory), R. Mandelbaum (CMU), E. Medezinski (Princeton), P. Melchior (Princeton), J. A. Newman (U. Pittsburgh, PITT PACC), F. J. Sánchez (UC Irvine), S. J. Schmidt (UC Davis), and S. Singh (UC Berkeley, BCCP), for the LSST Dark Energy Science Collaboration

Abstract: Community access to deep ($i \sim 25$), highly-multiplexed optical and near-infrared multi-object spectroscopy (MOS) on 8–40m telescopes would greatly improve measurements of cosmological parameters from LSST. The largest gain would come from improvements to LSST photometric redshifts, which are employed directly or indirectly for every major LSST cosmological probe; deep spectroscopic datasets will enable reduced uncertainties in the redshifts of individual objects via optimized training. Such spectroscopy will also determine the relationship of galaxy SEDs to their environments, key observables for studies of galaxy evolution. The resulting data will also constrain the impact of blending on photo-z's. Focused spectroscopic campaigns can also improve weak lensing cosmology by constraining the intrinsic alignments between the orientations of galaxies. Galaxy cluster studies can be enhanced by measuring motions of galaxies in and around clusters and by testing photo-z performance in regions of high density. Photometric redshift and intrinsic alignment studies are best-suited to instruments on large-aperture telescopes with wider fields of view (e.g., Subaru/PFS, MSE, or GMT/MANIFEST) but cluster investigations can be pursued with smaller-field instruments (e.g., Gemini/GMOS, Keck/DEIMOS, or TMT/WFOS), so deep MOS work can be distributed amongst a variety of telescopes. However, community access to large amounts of nights for surveys will still be needed to accomplish this work. In two companion white papers we present gains from shallower, wide-area MOS and from single-target imaging and spectroscopy.

1 Introduction

The Large Synoptic Survey Telescope (LSST) will play a major role in improving our knowledge of cosmology over the years 2023–2033, constraining fundamental cosmological parameters using multiple complementary methods. However, the baseline dark energy analyses that will be carried out by the LSST Dark Energy Science Collaboration (DESC) will require additional data from other ground-based facilities to improve photometric redshift estimates, reduce systematic uncertainties, and realize the full potential of LSST [1].

In this white paper, we describe the science opportunities to enhance cosmological measurements from LSST that would be enabled by community access to **deep ($i \sim 25$), highly-multiplexed** optical and near-infrared multi-object spectroscopy on 8–40m telescopes. Every cosmological probe that we plan to apply to LSST data would benefit from these capabilities.

In two companion white papers, we describe the gains for LSST cosmology that would come from community access to wider-field multi-object spectroscopy (spanning areas > 20 $\deg^2$) and from follow-up single-target imaging and spectroscopy of supernovae and strong lens systems [2; 3].

2 Deep Spectroscopic Samples for Photo-z Training

Photometric redshifts are critical for all LSST probes of cosmology. The cosmological tests we will perform all rely on determining the behavior of some quantity with redshift, z, but we cannot measure spectroscopic redshifts (spec-z's) for the large numbers of objects detected in the LSST imaging data. Even in cases where follow-up spectroscopy of individual objects will be needed (e.g., strong lens systems and some supernova studies), photo-z's are used to identify targets of interest. However, if photometric redshift estimates are systematically biased, dark energy inference can be catastrophically biased as well (see, e.g., [4]); as a result photo-z's are both a critical tool and a major source of concern affecting all cosmological analyses. The great depth of LSST data is a compounding factor, as it exacerbates any challenges associated with follow-up spectroscopy to the depths of the samples used for LSST.

Lacking a comprehensive knowledge of galaxy evolution, the only way in which photo-z errors can be reduced and biases characterized is via galaxies with robust spectroscopic redshift measurements. We follow [5] in dividing the uses of spec-z's into two classes, "training" and "calibration." **Training** is the use of samples with known z to develop or refine algorithms, and hence to *reduce random errors on individual objects' photo-z's*. Photometric redshifts that are trained from larger and more complete spectroscopic samples greatly improve the constraining power of LSST, for example by providing sharper maps of the large-scale structure and improved clustering statistics, providing better photometric classifications for supernovae, enabling identification of lower-mass galaxy clusters at higher confidence, and yielding better intrinsic alignment mitigation for weak lensing measurements. If photo-z's are limited only by photometric errors (as with a perfect training set), LSST can deliver photo-z estimates with sub-2% uncertainties ($\sigma_z < 0.02(1+z)$), but errors in real data sets at LSST depth with our current knowledge of galaxy spectral energy distributions are closer

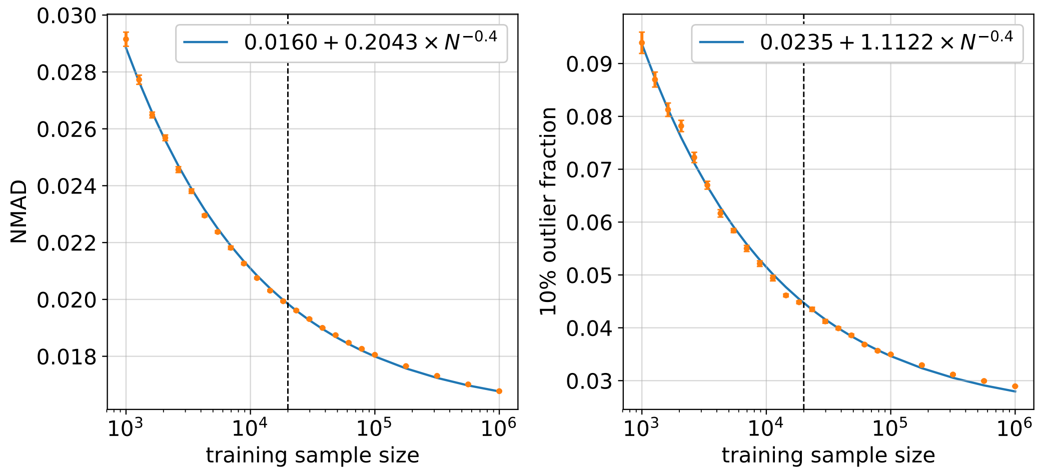

Figure 1: Orange points show photometric redshift errors and outlier rates versus the number of galaxies in the training set for galaxies with simulated LSST photometric errors. Photo-z's were calculated using a random forest regression algorithm. The left panel shows the photo-z error, quantified by the normalized median absolute deviation (NMAD) in $(z_{\rm phot} - z_{\rm spec})/(1+z_{\rm spec})$, as a function of training set size; similarly, the right panel shows the fraction of 10% outliers, i.e. objects with $|z_{\rm phot} - z_{\rm spec}|/(1+z_{\rm spec}) > 0.1$. A vertical dashed line shows the sample size for the baseline training survey from [5]. The blue curves represent simple fits to the measurements as a function of the training set size, N. This analysis uses a set of simulated galaxies from Ref. [11] that spans the redshift range of $0 < z < 4$, using a randomly-selected testing set of 10^5 galaxies for estimating errors and outlier rates.

to 5%. Achieving the ideal performance by having a large training sample spanning the properties of objects used for cosmology would improve the Dark Energy Task Force Figure of Merit from LSST lensing and Baryon Acoustic Oscillations alone by $\sim 40\%$ [6; 7].

Calibration is the problem of determining the true overall z distribution of a sample of objects; miscalibration will lead to *systematic* errors in photo-z's and hence downstream analyses [8; 7; 9; 10; 4]. For both training and calibration, we need sets of galaxies for which the true z is securely known. If spec-z's could be obtained for a large, unbiased sample of objects, both needs can be fulfilled using the same data. However, many faint galaxies fail to yield secure z's; hence other methods may be needed for calibration, as described in a companion paper [2]. Training still will benefit from incomplete samples.

In a recent paper [5], it was concluded that an effective training set of photometric redshifts for the LSST weak lensing sample would require highly-multiplexed medium-resolution ($R \sim 4000$) spectroscopy covering as much of the optical/infrared window as possible with very long exposure times on large telescopes. To enable photo-z direct calibration errors to be subdominant to other uncertainties if the training set were used for that purpose, the spec-z sample must comprise at least 20,000 galaxies spanning the full color and magnitude range used for cosmological studies, reaching $i = 25.3$; as can be seen in Fig. 1, improvements in photo-z errors and outlier rates are slow beyond this point.

A critical issue for machine learning-based photometric redshift algorithms [e.g., 12] is that sample/cosmic variance in the regions with spectroscopy can imprint on the redshift distribution over the whole sky, biasing photo-z results. To both quantify and mitigate this effect, the survey strategy described in Ref. [5] seeks to obtain spectroscopy spanning at least 15 widely-separated fields a minimum of $20'$ in diameter. Such a survey has comparable sample/cosmic variance to the Euclid C3R2 strategy of six 1 sq. deg. fields [13], but requires only $\sim$ 22% as much sky area to be covered; by spanning more fields than C3R2 it also allows more robust identification of regions that are overdense or underdense at a given z. We have calculated the amount of dark time required for such a survey (assuming one-third losses for weather and overheads) for a variety of instruments and telescopes of varying characteristics, updated from the tables in Ref. [5]; the results are summarized online at `http://d-scholarship.pitt.edu/id/eprint/36036`.

We note that such a survey has strong synergies with studies of galaxy evolution: it would determine the range of galaxy SEDs (and hence star formation histories) as a function of their local environment across the redshift and magnitude range covered by LSST cosmology samples (reaching down to $\sim L_*$ at $z = 2$, and including brighter galaxies to $z = 3$). Improved photo-z's will also enhance a variety of galaxy evolution science from LSST.

2.1 Testing the Impact of Blending on Photometric Redshifts

Due to its unprecedented depth and sensitivity to the low-surface-brightness outer regions of galaxies, the probability of two or more sources overlapping in LSST data is high. Approximately 63% of LSST sources will have at least 2% of their flux coming from other objects, in contrast to ~30% of DES sources [14; 15]. These overlaps complicate the measurement of galaxy fluxes and shapes, requiring accurate deblending [16; 17] or statistical correction techniques. Any residual light contamination will result in biases of individual-object photo-z's or potentially even in estimates of overall redshift distributions [18].

Deep multi-object, medium-resolution spectroscopy can detect the presence of blends and the redshifts of each component by identifying superimposed features in a spectrum, providing new training samples for deblenders and constraints on statistical corrections for deblending effects. While many blends can be detected as having multiple components in space-based data, some are so close that spectroscopy will provide the only definitive indication of a problem. As a result, the proposed photometric redshift training spectroscopy should also greatly enhance studies of deblending down to the depth of the LSST Gold sample, $i_{\rm lim} = 25.3$ [1].

3 Constraining Models of Intrinsic Galaxy Alignments

As discussed in the companion white paper on wide MOS, intrinsic correlations between galaxy shapes ("intrinsic alignments" or IA) induced by local/environmental effects are an important contaminant to weak lensing measurements; some constraints on this effect can be obtained via cross-correlation measurements, as described there. By enabling the 3D localization of galaxies, a deep MOS campaign would provide greatly-improved direct constraints on intrinsic alignments for typical weak lensing sources, rather than only for the bright and nearby objects which current datasets constrain [19; 14; 20]. Such data would extend our knowledge of IA to unexplored regimes, resolve the current inconsistencies between predictions of different hydrodynamical simulations [21; 22; 23], and allow better priors to be placed on IA parameters, increasing the cosmological constraining power of LSST.

Currently, the faintest magnitude-limited sample with an IA detection has $i_{\rm lim} < 19.8$ [20]. Forecasts based on recent measurements [24; 20] indicate that meaningful IA constraints require $\gtrsim 10^5$ galaxies with measured shapes and spectra. Obtaining this many spectra for a representative sample at the magnitude limit of the LSST Gold sample ($i_{\rm lim} = 25.3$) would require a significant expansion of the photo-z training program described above, and may be infeasible with currently planned facilities. However, even if we cannot reach magnitudes as faint as $i_{\rm lim} = 25.3$, extending direct IA tests closer to the LSST limits would be very valuable. Most of the IA signal will come from scales where shape noise dominates, and thus total IA constraining power (for a fixed number of galaxies) is maximized if the surface density of targets is high. Such a dense sample can be obtained by switching out targets to other, brighter galaxies during the photo-z training survey once secure redshifts are obtained; in this way, $\gtrsim 10^5$ spec-z's for objects with $i \lesssim 24$ should be obtainable during the training survey (only 9% as much observing time is needed to obtain the same S/N at $i = 24$ as at $i = 25.3$). Additional information will come from cross-correlations with shallower, wider-area surveys, as described in a companion paper [2].

4 Enhancing Cluster Cosmology via Spectroscopy

Deep MOS of galaxies in a set of fields containing galaxy clusters will improve LSST cluster cosmology in a number of ways, while simultaneously resolving open questions about galaxy evolution by determining differences between SEDs of galaxies in clusters versus the field.

Training and Testing Photometric Redshifts in Cluster Fields: Photometric redshifts are critical for weak-lensing mass calibration of galaxy clusters. They are already a leading source of systematic uncertainty in current work [25; 26], with even more stringent requirements on photo-z accuracy for LSST [1]. Photo-z's are vital for distinguishing the lensed background galaxy population from the unlensed foreground and cluster population. However, photo-z performance may degrade in higher-density regions due to the differing galaxy populations of clusters vs. the field, magnification and reddening of background sources, and severe blending due to cluster galaxies [e.g. 27].

Photo-z algorithms are generally trained and evaluated on fields selected not to contain massive structures. To ensure the robustness of cluster work, it will therefore be important to obtain additional MOS spectroscopy to weak lensing depths for a sample of ~ 20 clusters spanning a range of redshifts. This is best achieved with high-throughput, high-multiplex spectrographs with FOVs of $\sim 10'$ (wide-format IFUs may be suitable in cluster cores).

Such a program would be able to characterize the performance of photo-z probability distributions as a function of magnitude, redshift, and cluster properties. A sample of 1000 objects per cluster allows a crude binning into 3 magnitude and 3 redshift bins with $\sim$100 galaxies per bin, enough for a statistically meaningful evaluation of photo-z performance at the percent level. This should be done for a range of cluster redshifts, masses, and dynamical states, requiring ideally > 20 clusters. It is critical to achieve near-complete redshift success rates to avoid biases from target populations for which no redshift is measured in a first attempt. Such data will also yield very valuable insight into deblending performance in general.

Measurements of Cluster Kinematics and Infall Velocities: Deep MOS observations in the fields of clusters will also enable direct mass estimates for galaxy clusters via the infall method [28; 29; 30; 31], providing an additional calibration of the mass-richness relation [32] "for free" from the photometric redshift training/test spectroscopy. We hope to obtain ~ 200 redshifts at projected separation < 5 Mpc for each cluster targeted for photo-z studies, providing purely kinematic mass estimates. In conjunction with weak-lensing measurements, spectroscopy of halo infall regions also enables sensitive comparison between the dynamical and the lensing potential of gravity, which can differ from each other at measurable levels in several modified gravity theories that seek to provide an alternate explanation for cosmic acceleration [e.g., 33].

5 Recommendations

Given the large gains to LSST cosmological studies that will come from deep multi-object spectroscopy, we recommend that access to modestly-wide-field, highly-multiplexed, large aperture spectroscopic facilities be pursued during the next decade. Specifically,

- For photometric redshift training and tests of blending and intrinsic alignment effects, it is desirable to have an instrument of maximal multiplexing with a field of view of at least $20'$ diameter. Subaru/PFS, the Maunakea Spectroscopic Explorer (MSE), or GMT/GMACS with the MANIFEST fiber feed are all well-suited for this work (wider-field fiber-fed spectrographs on other >6 m telescopes could also be suitable).[1] Personnel costs are likely to be high on smaller telescopes because of the longer time required for deep surveys on them, making instruments on 4 m telescopes such as DESI unattractive for this work.

- For galaxy cluster studies, targeting more objects over smaller fields of view is desirable. For that work, suitable options may include Gemini/GMACS, Keck/DEIMOS or LRIS, GMT/GMACS (in multislit mode), TMT/WFOS, or a new, higher-multiplex spectrograph at one of these observatories; wider-field facilities such as DESI, Subaru/PFS or MSE would be less well-suited.

The telescope time required for photo-z training is substantial ($\sim$6 months of dark nights if one of the most optimal facilities is used exclusively), but it can be spread out over the ten-year span of the LSST survey, reducing the impact in any one year. Additionally, these studies are highly synergistic with studies of galaxy evolution, potentially allowing combined surveys with a large impact on multiple fields of research.

[1] See `http://d-scholarship.pitt.edu/id/eprint/36036` for survey time estimates for various instrument/telescope combinations.

Acknowledgements

The LSST Dark Energy Science Collaboration acknowledges ongoing support from the Institut National de Physique Nucléaire et de Physique des Particules in France; the Science & Technology Facilities Council in the United Kingdom; and the Department of Energy, the National Science Foundation, and the LSST Corporation in the United States. DESC uses resources of the IN2P3 Computing Center (CC-IN2P3–Lyon/Villeurbanne - France) funded by the Centre National de la Recherche Scientifique; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231; STFC DiRAC HPC Facilities, funded by UK BIS National E-infrastructure capital grants; and the UK particle physics grid, supported by the GridPP Collaboration. This work was performed in part under DOE Contract DE-AC02-76SF00515.

References

[1] The LSST Dark Energy Science Collaboration, R. Mandelbaum, T. Eifler, et al. The LSST Dark Energy Science Collaboration (DESC) Science Requirements Document. *arXiv:1809.01669*, September 2018.

[2] R. Mandelbaum *et al.* Wide-field Multi-Object Spectroscopy to Enhance Dark Energy Science from LSST. *Astro2020 White Paper*, 2019.

[3] R. Hlozek *et al.* Single-object Imaging and Spectroscopy to Enhance Dark Energy Science from LSST. *Astro2020 White Paper*, 2019.

[4] A. P. Hearin, A. R. Zentner, Z. Ma, and D. Huterer. A General Study of the Influence of Catastrophic Photometric Redshift Errors on Cosmology with Cosmic Shear Tomography. *ApJ*, 720:1351–1369, September 2010.

[5] J. A. Newman, A. Abate, F. B. Abdalla, et al. Spectroscopic needs for imaging dark energy experiments. *Astroparticle Physics*, 63:81–100, March 2015.

[6] A. Albrecht, G. Bernstein, R. Cahn, et al. Report of the Dark Energy Task Force. *arXiv Astrophysics e-prints*, pages arXiv:astro–ph/0609591, September 2006.

[7] H. Zhan. Cosmic tomographies: baryon acoustic oscillations and weak lensing. *Journal of Cosmology and Astro-Particle Physics*, 8:8–+, August 2006.

[8] H. Zhan and L. Knox. Baryon Oscillations and Consistency Tests for Photometrically Determined Redshifts of Very Faint Galaxies. *ApJ*, 644:663–670, June 2006.

[9] L. Knox, Y.-S. Song, and H. Zhan. Weighing the Universe with Photometric Redshift Surveys and the Impact on Dark Energy Forecasts. *ArXiv Astrophysics e-prints*, pages arXiv:astro–ph/0605536, May 2006.

[10] J. A. Tyson. Precision Studies of Dark Energy with LSST. *ArXiv Astrophysics e-prints*, pages arXiv:astro–ph/0609516, September 2006.

[11] Melissa L. Graham, Andrew J. Connolly, Željko Ivezić, et al. Photometric Redshifts with the LSST: Evaluating Survey Observing Strategies. *AJ*, 155:1, January 2018.

[12] S. Cavuoti, M. Brescia, C. Tortora, et al. Machine-learning-based photometric redshifts for galaxies of the ESO Kilo-Degree Survey data release 2. *MNRAS*, 452:3100–3105, Sep 2015.

[13] Daniel C. Masters, Daniel K. Stern, Judith G. Cohen, et al. The Complete Calibration of the Color-Redshift Relation (C3R2) Survey: Survey Overview and Data Release 1. *ApJ*, 841:111, Jun 2017.

[14] S. Samuroff, J. Blazek, M. A. Troxel, et al. Dark Energy Survey Year 1 Results: Constraints on Intrinsic Alignments and their Colour Dependence from Galaxy Clustering and Weak Lensing. *arXiv e-prints*, page arXiv:1811.06989, November 2018.

[15] J. Sanchez, I. Mendoza, D. Kirkby, and P. Burchat. Olber's paradox revisited – effects of overlapping sources on cosmic shear estimation: Statistical sensitivity and pixel-noise bias, in prep.

[16] P. Melchior, F. Moolekamp, M. Jerdee, et al. SCARLET: Source separation in multi-band images by Constrained Matrix Factorization. *Astronomy and Computing*, 24:129–142, July 2018.

[17] Euclid Collaboration, N. Martinet, T. Schrabback, et al. Euclid Preparation IV. Impact of undetected galaxies on weak lensing shear measurements. *arXiv e-prints*, page arXiv:1902.00044, January 2019.

[18] D. Gruen, Y. Zhang, A. Palmese, et al. Dark Energy Survey Year 1 Results: The effect of intra-cluster light on photometric redshifts for weak gravitational lensing. *arXiv e-prints*, page arXiv:1809.04599, September 2018.

[19] B. Joachimi, R. Mandelbaum, F. B. Abdalla, and S. L. Bridle. Constraints on intrinsic alignment contamination of weak lensing surveys using the MegaZ-LRG sampl e. *A&A*, 527:A26, March 2011.

[20] Harry Johnston, Christos Georgiou, Benjamin Joachimi, et al. KiDS+GAMA: Intrinsic alignment model constraints for current and future weak lensing cosmology. *arXiv e-prints*, page arXiv:1811.09598, November 2018.

[21] A. Tenneti, R. Mandelbaum, and T. Di Matteo. Intrinsic alignments of disk and elliptical galaxies in the MassiveBlack-II and Illustris simulati ons. *ArXiv e-prints*, page arXiv:1510.07024, October 2015.

[22] N. Chisari, S. Codis, C. Laigle, et al. Intrinsic alignments of galaxies in the Horizon-AGN cosmological hydrodynamical simulation. *MNRAS*, 454:2736–2753, December 2015.

[23] Nora Elisa Chisari, Cora Dvorkin, Fabian Schmidt, and David N. Spergel. Multitracing anisotropic non-Gaussianity with galaxy shapes. *Phys. Rev. D*, 94:123507, December 2016.

[24] S. Singh, R. Mandelbaum, and S. More. Intrinsic alignments of SDSS-III BOSS LOWZ sample galaxies. *MNRAS*, 450:2195–2216, June 2015.

[25] D. E. Applegate, A. von der Linden, P. L. Kelly, et al. Weighing the Giants - III. Methods and measurements of accurate galaxy cluster weak-lensing masses. *MNRAS*, 439:48–72, March 2014.

[26] T. McClintock, T. N. Varga, D. Gruen, et al. Dark Energy Survey Year 1 results: weak lensing mass calibration of redMaPPer galaxy clusters. *MNRAS*, 482:1352–1378, January 2019.

[27] A. Molino, N. Benítez, B. Ascaso, et al. CLASH: accurate photometric redshifts with 14 HST bands in massive galaxy cluster cores. *MNRAS*, 470:95–113, September 2017.

[28] A. Diaferio. Mass estimation in the outer regions of galaxy clusters. *MNRAS*, 309:610–622, November 1999.

[29] K. Rines, M. J. Geller, M. J. Kurtz, and A. Diaferio. CAIRNS: The Cluster and Infall Region Nearby Survey. I. Redshifts and Mass Profiles. *AJ*, 126:2152–2170, November 2003.

[30] M. Falco, G. A. Mamon, R. Wojtak, S. H. Hansen, and S. Gottlöber. Dynamical signatures of infall around galaxy clusters: a generalized Jeans equation. *MNRAS*, 436:2639–2649, December 2013.

[31] J. Arthur, F. R. Pearce, M. E. Gray, et al. nIFTy galaxy cluster simulations - V. Investigation of the cluster infall region. *MNRAS*, 464:2027–2038, January 2017.

[32] K. J. Rines, M. J. Geller, A. Diaferio, H. S. Hwang, and J. Sohn. HeCS-red: Dense Hectospec Surveys of redMaPPer-selected Clusters. *ApJ*, 862:172, August 2018.

[33] Y. Zu, D. H. Weinberg, E. Jennings, B. Li, and M. Wyman. Galaxy infall kinematics as a test of modified gravity. *MNRAS*, 445:1885–1897, December 2014.