

Steven Boada

Department of Physics and Astronomy, 4242 TAMU
Texas A&M University, College Station, Texas 77843-4242

January 1, 2016

Nancy DeHaan
Rutgers University
Department of Physics and Astronomy
136 Frelinghuysen Road
Piscataway, NJ 08854

Nancy DeHaan:

Please find the enclosed application for the position of postdoctoral research fellow as advertised on the AAS Job Register, Postdoctoral Position in Galaxy Cluster Research - JRID52807. I expect to complete my Ph.D. at Texas A&M University as of May, 2016 under the supervision of Dr. Casey Papovich.

My current research is focused on using large datasets to study galaxy clusters and their application as cosmological probes. I am conducting detailed simulations of the upcoming Hobby Eberly Telescope Dark Energy Experiment (HETDEX) and the VIRUS instrument's ability to obtain accurate cluster mass estimates. I have also undertaken targeted observations of optically selected clusters with the VIRUS prototype instrument on the 2.7m telescope at McDonald Observatory. Taken together, this will allow me to provide probabilistic mass estimates for future blind spectroscopic surveys such as HETDEX.

I am excited about the opportunity to continue my research at Rutgers University, and I believe my background in galaxy cluster dynamics, cosmology, and large datasets will compliment your groups' ongoing work. I am particularly excited to work with Dr. Hughes and to be involved in the follow-up of ACT clusters.

If you have any questions or would like to discuss my application further please do not hesitate to contact me at boada@physics.tamu.edu or by phone at +1 615 200 0119. I plan on attending the Winter AAS meeting, and would be more than happy to introduce myself in person. Thank you very much for your time and consideration. I look forward to hearing from you soon.

Sincerely,

A handwritten signature in black ink, reading "Steven Boada". The signature is fluid and cursive, with the first name "Steven" and last name "Boada" clearly legible.

Steven Boada

Steven Boada

CONTACT INFORMATION

Department of Physics and Astronomy
4242 TAMU
Texas A&M University
College Station, Texas 77843-4242

Phone: +1 (615) 200-0119
E-mail: boada@physics.tamu.edu
WWW: <http://boada.github.io>

RESEARCH INTERESTS

Observation Cosmology, Large-area Sky Surveys (e.g., DES, LSST, SDSS, ACT, SPT), Galaxy Clusters, High Performance Computing (HPC), Galaxy Evolution, Interacting Galaxies and Morphology.

EDUCATION

Texas A&M University, College Station, Texas USA

Ph.D. Candidate, Physics (Astronomy focus; expected graduation date: May 2016)

- Dissertation Title: “Galaxy Cluster Dynamics in the Era of Large Spectroscopic Surveys”
- Advisor: Dr. Casey Papovich

The University of Tennessee, Knoxville, Tennessee USA

M.S., Physics (Computational Astrophysics), August, 2009

- Thesis Title: “An Automated Approach to the Study and Classification of Colliding and Interacting Galaxies”
- Advisor: Dr. Michael Guidry

The University of Tennessee, Knoxville, Tennessee USA

B.S., Physics, May, 2007

PROFESSIONAL EXPERIENCE

Texas A&M University, College Station, Texas USA

Research Assistant

August, 2010 - Present

The University of Tennessee, Knoxville, Tennessee USA

Research Assistant

August, 2007 - 2009

National Center for Computational Science, Oak Ridge National Laboratory, Oak Ridge, Tennessee USA

Visiting Scientist

May, 2007 - August, 2009

Carried out the computing projects required to complete Master’s, including modeling of interacting galaxy systems, machine learning, and other HPC tasks.

OBSERVING EXPERIENCE

Proposals

- *Measuring the Masses of X-ray-Selected, Low-Mass Galaxy Clusters and Groups with Integral Field Spectroscopy*
Co-I (PI: N. Mehrtens), McDonald Observatory, 4 nights awarded, 2013
- *Measuring the Masses of Galaxy Clusters with Integral Field Spectroscopy*
Co-I (PI: C. Papovich), McDonald Observatory, 9 nights awarded, 2012
- *Measuring the Masses of Galaxy Clusters with Integral Field Spectroscopy*
Co-I (PI: C. Papovich), McDonald Observatory, 5 nights awarded, 2012

Telescopes

- Harlan J. Smith 2.7m Telescope, Mitchell Spectrograph (formerly VIRUS-P), 20+ nights

Data Experience

- Integral Field Spectroscopy
- Hubble Space Telescope Imaging
- Sloan Digital Sky Survey Imaging and Spectroscopy

COMPUTING EXPERIENCE

Extensive experience in the processing and application of large astronomical data sets, including: the acquisition and reduction of optical integral field unit spectroscopy, querying large astronomical databases such as the Sloan Digital Sky Survey and the Millennium Simulation, analysis of multi-wavelength imaging from the Hubble Space Telescope. Key computing skills include: mastery of the Python language, and the interface with other languages and tools, considerable experience with large multiprocessor applications (e.g. Gadget-2) and high performance computing systems, supervised and unsupervised machine learning and optimization, GPGPU computing, and participation in open source and collaborative development environments, including version control. Co-author of ASTLIB python library, see <http://astlib.sourceforge.net>

TEACHING AND OUTREACH

Texas A&M University, College Station, Texas USA

Teaching Assistant

2010 - Spring, 2015

Supervised undergraduate students for weekly lab sessions, tutoring sessions, grading of homework and quizzes for Basic Astronomy, Overview of Modern Astronomy, and Survey of Astronomy.

Physics Festival

2010 - Present

Demonstrated physics and astronomy principles for students from elementary through high school and the general public.

Star Parties

2010 - Present

Discussed astronomical topics and operated telescopes for college students and the general public.

Nashville State Community College, Nashville, Tennessee USA

Adjunct Faculty

Spring, 2010

Primary instructor for introductory physics course, Conceptual Physics.

The University of Tennessee, Knoxville, Tennessee USA

Teaching Assistant

August, 2007 - 2009

Supervised laboratory experiences for undergraduate students in Introduction to Modern Physics, and Electricity and Magnetism for Engineering. Designed and taught laboratories for undergraduate Honors Astronomy.

ACADEMIC HONORS AND AWARDS

The University of Tennessee: graduated Magna Cum Laude, Phi Beta Kappa, Sigma Pi Sigma, President, Society of Physics Students 2006 thru 2007

GRANTS AND AWARDS	<ul style="list-style-type: none"> • <i>The Road to the Virgo Cluster: The DECam/IRAC Galaxy Environment Survey</i> Co-I (PI: C. Papovich), NSF Alliances for Graduate Education and the Professoriate, 2015 • <i>Graduate Student Presentation Grant</i> PI, Texas A&M University Office of Graduate and Professional Studies, 2015 • <i>Graduate Student Travel Grant</i> PI, Texas A&M University Department of Physics and Astronomy, 2015 		
POSTERS AND PRESENTATIONS	<p>Talk: CANDELS Team Meeting, University of Santa Cruz, Santa Cruz, CA July, 2015</p> <p>Talk: CANDELS Team Meeting, STScI, Baltimore, MD July, 2014</p> <p>Poster: Bashfest Symposium, University of Texas, Austin, TX October, 2013</p> <p>Talk: CANDELS Team Meeting, University of Kentucky, Lexington, KY August, 2013</p> <p>Poster: GMT Science Meeting, University of Chicago, Chicago, IL June, 2013</p> <p>Talk: CANDELS Team Meeting, University of Santa Cruz, Santa Cruz, CA September, 2012</p> <p>Poster: 219th AAS Meeting, Austin, TX January, 2012</p> <p>Poster: Bashfest Symposium, University of Texas, Austin, TX October, 2011</p> <p>Talk: Texas A&M Astronomy Symposium, Texas A&M University, College Station, TX August, 2011–13</p>		
REFERENCES	<p>Dr. Casey Papovich Dept. of Physics & Astronomy 4242 TAMU Texas A&M University College Station, Texas 77843 papovich@physics.tamu.edu</p>	<p>Dr. Vithal Tilvi School of Earth & Space Ex- ploration P.O. Box 871404 Arizona State University Tempe, Arizona 85287 tilvi@asu.edu</p>	<p>Dr. Nicholas Suntzeff Dept. of Physics & Astronomy 4242 TAMU Texas A&M University College Station, Texas 77843 nsuntzeff@tamu.edu</p>

Steven Boada

PUBLICATIONS

- **Steven Boada**, Tilvi, V., Papovich, C., Quadri, R. F., Hilton, M., Finkelstein, S., Guo, Y., Bond, N., Conselice, C., Dekel, A., Ferguson, H., Giavalisco, M., Grogin, N. A., Kocevski, D. D., Koekemoer, A. M. and Koo, D. C. *The Role of Bulge Formation in the Homogenization of Stellar Populations at $z \sim 2$ as revealed by Internal Color Dispersion in CANDELS* 2015, ApJ, 803, 104

CONFERENCE PROCEEDINGS

- Ting Li, DePoy, D. L., Marshall, Jennifer L., Nagasawa, D. Q., Carona, D. W., **Boada, S.** *Monitoring the atmospheric throughput at Cerro Tololo Inter-American Observatory with aTmCam* 2014, Proceedings of the SPIE, 9147, 91476Z

Steven Boada

Publicly available large-area sky surveys, such as the Sloan Digital Sky Survey (SDSS), have opened new avenues for astronomers to conduct research, and continue to create opportunities to answer questions which previously would have been either time or cost prohibitive using traditional ground- and space-based observations. Ongoing and upcoming surveys, such as the Dark Energy Survey (DES) and the Large Synoptic Survey Telescope (LSST), will continue to revolutionize the field by generating substantially more data than any previous survey. With such large datasets, the intersection between astrophysics and the data sciences may provide key insights in the future.

However, data science alone will not answer all of the outstanding astrophysical questions. An understanding of the observational methods used, astrophysical biases implied, and other factors is also required. My experiences in both computing at the supercomputing division of Oak Ridge National Laboratory, and observations working with ground- and space-based, optical/IR imaging, and optical integral field spectroscopy give me a unique understanding of both worlds. My research interests have been heavily influenced by these experiences and currently focuses on utilizing large-area sky surveys, machine learning and data mining techniques to answer astrophysical questions. While I am interested in many topics, at present, my research focuses on using large, ground-based, optical surveys to study galaxy clusters in a cosmological context. Below, I describe my current and previous research and give small introductions to three projects relating to galaxy clusters where data mining, machine learning, and ever-expanding datasets could make significant contributions.

PRESENT AND PAST RESEARCH

HETDEX as a Galaxy Cluster Survey: The Hobby Eberly Telescope Dark Energy eXperiment (HETDEX; [Hill et al. 2008b](#)) is a trailblazing effort to observe high- z large scale structures using a collection of wide-field integral field unit (IFU) spectrographs, VIRUS ([Hill et al. 2012](#)). HETDEX’s primary goal is to probe the evolution of the dark energy equation of state etched onto $z > 2$ galaxy populations. Over 420 degree², HETDEX expects to find 800,000 Ly α emitters, and more than 10^6 [O II] emitting galaxies at $z < 0.5$ masquerading as high- z galaxies ([Acquaviva et al. 2014](#)). It is this low- z sample that I find most interesting as it may contain up to 100 clusters with masses $\sim 10^{15} M_{\odot}$.

This low- z cluster sample is important because while we are discovering many tens of thousands of clusters with large-area sky surveys, the principal challenge to modern precision cosmology is not in the number of clusters detected, but the accurate recovery of the galaxy cluster mass (e.g., [Sehgal et al. 2011](#); [Planck Collaboration 2013](#); [Bocquet et al. 2015](#)). This study aims to investigate how well HETDEX-like observations are able to recover galaxy cluster masses over a wide range of masses and redshifts. To do this, I use two versions of mock galaxy observations to estimate the number of clusters observed and the accuracy with which cluster masses can be recovered with HETDEX. In addition to simulating the planned HETDEX observing strategy, I am simulating how targeted followup observations can improve

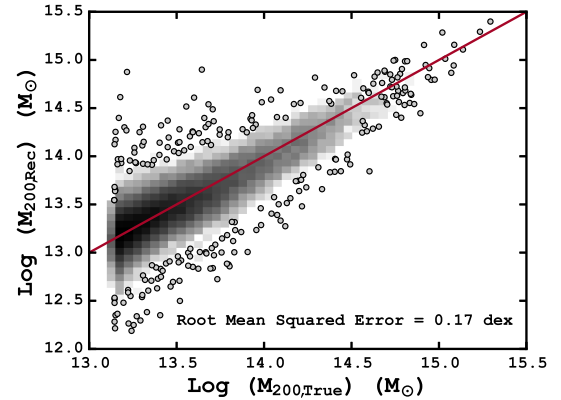


Figure 1: Tests to recover total cluster masses using mock galaxy samples. The plot shows the true halo mass ($M_{200, True}$) from the mock catalog against the dynamical mass recovered ($M_{200, Rec}$). The red line shows the 1:1 relation. The results are in good agreement at the high-mass end. At lower masses the scatter increases as there are fewer galaxies per cluster, and the clusters are less dynamically “relaxed”.

the mass estimates. Using both observing methods, I produce cluster mass probability functions, $P(M_{True}|M_{Rec}, z)$ by comparing the recovered mass to the true mass; see Figure 1. I am currently applying this methodology to real data as described below and this work will form the basis for cluster studies when HETDEX begins operation in 2016.

As a followup to the simulated survey, I have observed ten galaxy clusters selected from the SDSS DR8 and the *Chandra-XMM* X-ray Cluster Survey (Mehrtens et al. 2012) at $z = 0.2 - 0.3$. with the Mitchell IFU Spectrograph (formerly VIRUS-P; Hill et al. 2008a) at McDonald Observatory. I have measured spectroscopic redshifts and line-of-sight velocities of the galaxies in and around each cluster, derived a line-of-sight velocity dispersion, and inferred a dynamical mass for each cluster. Using the mass probability functions described above, I am able to update these cluster masses and compare them to existing literature estimates, or to the masses estimated from other observables. The goal is to estimate the ability of future, blind, spectroscopic surveys to improve the accuracy of the observable–mass relationship, and to help calibrate future large-area sky survey mass estimates, where extensive spectroscopic follow up is impractical. This work is expected to be completed in the spring, 2016.

Impact of Galaxy Bulges on Stellar Populations at $z \sim 2$: The mechanism by which galaxies in the early universe transition from blue, star-forming disks to quiescent, bulge-dominated galaxies continues to be an active area of research. To investigate possible causes, I used data from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; Koekemoer et al. 2011; Grogin et al. 2011) to study how the spatial variation in the stellar populations of galaxies relate to their morphologies and stellar masses at $1.5 < z < 3.5$ (Boada et al. 2015). I measured the variations in stellar population age (and dust attenuation) using the Internal Color Dispersion (ICD) statistic and found the ICD is related to stellar mass and galaxy morphology. Specifically, I found that galaxies with stellar masses $10 < \text{Log } M/M_{\odot} < 11$ and disk-dominated morphologies have a large variation in their stellar populations, and that as the stellar mass increases *or* the galaxy becomes more bulge-dominated the amount of this stellar diversity decreases.

DIRECTION OF FUTURE RESEARCH

Introduction: In the coming years, many large surveys, over a wide range of wavelengths, will improve our ability to use galaxy clusters to determine cosmological parameters by adding additional statistical depth to our current data sets. It is anticipated that upon their completion the South Pole Telescope and the Atacama Cosmology Telescope will find approximately one thousand clusters using Sunyaev–Zeldovich Effect (SZE) observations. Attempts are already underway to calibrate the SZE derived masses using subsamples of clusters with dynamical mass estimates or X-ray temperature measurements (Sifón et al. 2013; Bocquet et al. 2015). The eROSITA telescope on-board the *Spektrum-Roentgen-Gamma* mission will perform an all-sky, X-ray survey and detect an estimated 50,000 or more clusters. Large optical/IR surveys such as DES and LSST, PanSTARRS, the Hyper-Suprime Camera survey, the *Euclid* mission and WFIRST-AFTA will survey enormous portions of the sky and identify vast numbers of clusters at redshifts as high as $z \sim 1$ using optical selection methods (e.g., Rykoff et al. 2012; Rozo et al. 2015).

However, as many of these surveys are optimized to address a broad range of science goals, they are very rarely the best choice for any single science question. Therefore, advancement in data science techniques which are designed to overcome the limitations of the observations are required to answer specific science questions. For my future research, I am interested in using large-area sky surveys to improve the mass estimates of low mass clusters and to search for similar mass clusters at high- z which can aid in the determination of cosmological parameters. Below, I identify three

areas where I can make improvements to existing techniques and algorithms. (1) Current external mass calibrators will need to be extended to lower cluster masses to accurately calibrate the optical richness–mass relation and improve its ability to recover the low mass end of the halo mass function (HMF). (2) Correctly identifying the cluster’s center without a clear central galaxy can improve the accuracy of the optical–mass scaling relations. (3) The algorithms used to optically identify clusters assume specific things about the population of cluster galaxies, often a well-defined red sequence. This can be problematic for low mass or high- z cluster searches.

Calibration of the observable–Cluster Halo Mass Relation: Large-scale optical surveys (e.g., DES and LSST) expect to detect hundreds of thousands of galaxy clusters at $z < 1$. A major challenge for these surveys is relating a cluster observable to the total dark-matter mass. One promising mass estimator is the optical richness, λ , the weighted number of galaxies within a scale aperture (e.g., [Rozo et al. 2011](#)). Figure 2 shows that the richness correlates strongly with cluster mass on average ([Rozo et al. 2010](#)), but the absolute mass scale of the optical richness mass estimator and the scatter in cluster mass at fixed optical richness are imprecisely known ([Rykoff et al. 2012](#)). These systematics remain the major source of uncertainty in deriving cosmological constraints from cluster abundances and must be measured using independent methods to realize the full potential of these surveys.

Applying my current work to a large photometric survey such as those mentioned above is the logical next step. For example, the HETDEX observations overlap with the DES area and SDSS Stripe 82 which provides a wealth of external data. Using preexisting datasets, the observations from HETDEX, or through dedicated followup spectroscopic observations, I will be able to measure the absolute mass scale of the richness estimator and measure the scatter of the richness–mass relation, $\sigma_{M|\lambda}$. An accurate measure of this scatter can lead to a decrease on the error bar on cosmological measures (σ_8 and Ω_M) by as much as 50% ([Rozo et al. 2010](#)). An absolute calibration of the richness–mass estimation will provide a much needed tool for future imaging surveys which will identify clusters to $z = 1$ and beyond.

Investigation of Cluster Mis-centering: Misidentification of the cluster center (due to the mass distribution not being directly observable) can add significant error to the estimate of its mass. Fortunately, hierarchical growth suggests the brightest cluster galaxy (BCG) should be located at the center of the cluster’s potential well. Unfortunately, this is not always the case (e.g., [Skibba et al. 2011](#)). When there is no clear central galaxy or the brightest (most massive) galaxy is difficult to identify, the centroid of the hot intracluster medium or a weighted centroid of the member galaxy positions (e.g., [George et al. 2012](#)) can be used.

Many optical, photometry-based cluster finding algorithms (e.g., redMapper; [Rykoff et al. 2014](#)) assume a massive cD galaxy is present in each cluster, which gives a success rate of approx-

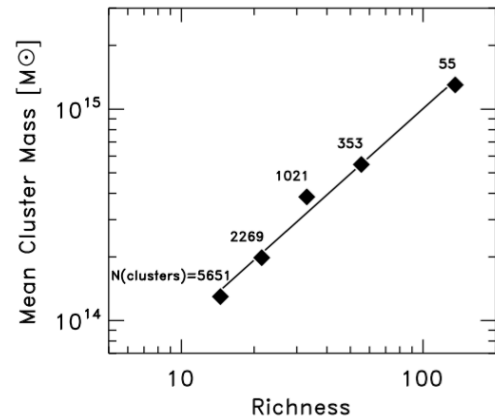


Figure 2: The relation between the mean Richness and the mean cluster mass derived from stacked weak lensing measurements. The number of clusters in each stack are given above each data point. There is a strong correlation between richness and cluster mass, however, because the data are stacked, the absolute mass scale and scatter in mass at fixed richness are imprecisely known.

imately 85% and is described as the “least bad” option for identifying the cluster center. The remaining 15% of clusters often lack a defined center because the cluster’s BCG does not match the expected color of a massive, elliptical galaxy. Specifically, cool-core clusters can have blue, star-forming central galaxies, confusing the cluster center determination.

Star-forming central galaxies are most common in clusters at the lowest end of the halo mass function, which coincide with the majority of clusters in the universe, or in clusters with increasing redshift (the Butcher-Oemler effect). To help correct this issue, I am interested in using targeted spectroscopic observations to identify the center of motion through the observed dynamics of the member galaxies. This is insensitive to the type of central galaxy and could provide additional clues to optically select central galaxies which do not match the predicted color. A second method is to use a targeted campaign or large archival dataset of *XMM-Newton* observations of low richness clusters to place further constraints on the center. Given a large enough sample, a statistical approach will be very powerful.

Search for Clusters Beyond $z = 1$: The identification of galaxy clusters in the low- z universe through optical imaging often relies on a well defined red sequence of cluster galaxies. Beyond $z = 1$ this becomes increasingly hard due to the depth required and the increasing lack of well defined red sequence. Therefore, new methods to search for and identify clusters in the early universe with optical/NIR imaging must be adopted.

The search for clusters in point data (e.g., RA, DEC, z) is not unique to astronomy and, as such, many methods (k-means, mixture models, etc.) have been developed to tackle the problem. Dehghan & Johnston-Hollitt (2014) used a data agnostic machine learning algorithm, the Density-Based Spatial Clustering of Applications with Noise (DBSCAN; Ester et al. 1996), to search for clusters, group, and filaments in the Extended Chandra Deep Field South (ECDF-S), see Figure 3. They successfully identified structures at $z < 1$ using a ground based spectroscopic survey and the MUSYC photometric catalog. Improvements could be made immediately by using modern clustering algorithms, such as Ordering Points To Identify the Clustering Structure (OPTICS; Ankerst et al. 1999) or a hierarchical version of DBSCAN (HDBSCAN; Campello et al. 2013), both of which are designed to address the weaknesses of DBSCAN.

To reach beyond $z = 1$, I will make use of deep imaging data of five fields taken as part of CANDELS with a combined area of approximately 800 arcmin². I will use machine learning techniques, mature photometric catalogs, and a wealth of auxiliary data to search for and classify structures as clusters and filaments. I am also interested in revisiting my previous CANDELS work (Boada et al. 2015) using the environmental information provided by this structure search to investigate how those results might change. This clustering search can also be extended to the large sky-area of PanSTARRS to search for dwarf galaxies or other stellar associations and will eventually be useful to the large area survey of WFIRST-AFTA. This study will provide important advances in the systematic search for large structures above $z = 1$, aid in the development and application of data agnostic machine learning to astronomy, provide improved cluster detection at all redshifts, and suggest followup targets requiring JWST.

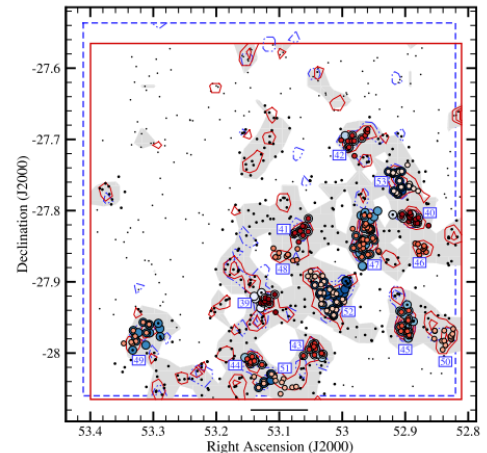


Figure 3: Potential structures detected at $0.65 < z < 0.7$ in the ECDF-S. Colored points show potential clusters whereas black points show galaxies classified as noise.

REFERENCES

- Acquaviva et al. 2014, PIAU, 10, 365
- Ankerst et al. 1999, ACM Sigmod Record, 49
- Boada et al. 2015, ApJ, 803, 104
- Bocquet et al. 2015, ApJ, 799, 214
- Campello et al. 2013 in Lec. Notes in Comp. Sci., 7819, 160-172
- Dehghan & Johnston-Hollitt 2014, AJ, 147, 52
- Ester et al. 1996, in Sec. Int. Conf. on Knowledge Discovery and Data Mining, 226–231
- George et al. 2012, ApJ, 757, 2
- Grogin et al. 2011, ApJS, 197, 35
- Hill et al. 2008a, Proc. SPIE 7014, 701470–15
- Hill et al. 2008b, Panoramic Views of Galaxy Formation and Evolution ASP Conference Series, 399
- Hill et al. 2012, Proc. SPIE 8446, 84460N
- Koekemoer et al. 2011, ApJS, 197, 36
- Mehrtens et al. 2012, MNRAS, 423, 1024
- Planck Collaboration 2013, A&A, 571, 19
- Rozo et al. 2015, MNRAS, 453, 38
- Rozo et al. 2011, ApJ, 735, 118
- Rozo et al. 2010, ApJ, 708, 645
- Rykoff et al. 2012, ApJ, 746, 178
- Rykoff et al. 2014, ApJ, 785, 104
- Sehgal et al. 2011, ApJ, 732, 44
- Sifón et al. 2013, AJ, 772, 25
- Skibba et al. 2011, MNRAS, 410, 417