Fermilab

THE DARK ENERGY SURVEY

Physical Interpretation of Cluster Pressure Profile in CMB Observation

Giovanni Volta - Università degli Studi di Ferrara

Abstract

The aim of this project is to study Galaxy Clusters using cosmic microwaves of background as probe and make data analysis with software CosmoSIS. I want to test the dependability of our analysis software with new adding of scatter relation on Log-Normal distribution of richness, using theoretical data generated from the physical model. The main quantity is the y-Compton parameter of Sunyaev-Zel'Dovich effect, which is given by two contribution: one halo term and two halo term. I took care, most of all, on one halo term in order to see it possible to use only this contribution for the next works since that the two halo term increase a lot the analysis time. Moreover, with the galaxy cluster catalog of from the Dark Energy Survey (DES) and Sunyaev-Zel'Dovich maps of Planck Large Telescope, I tried to put some constraint on the amplitude of galaxy pressure profile, namely A_p, and on the parameter that governs the behavior of scatter relation, namely σ_0. From this parameter estimation, I recovered interested results.

Academic Year 2016 - 2017

1 Introduction

The Dark Energy Survey (DES) is a international collaboration with the goals of mapping hundreds of millions of galaxies, detect thousands of supernovae, and find patterns of cosmic structure that will reveal the nature of the mysterious dark energy that is accelerating the expansion of our Universe. The survey is imaging 5000 square degrees of the southern sky in five optical filters to obtain detailed information about each galaxy. The survey was designed to cover areas of the sky that have been observed by complementary experiments.

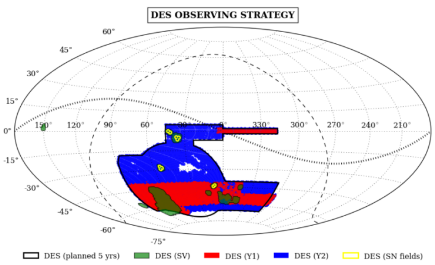

Figure 1: Footprint of the wide-area survey on the sky (colored region) in celestial coordinates; the dashed curve shows the approximate location of the Milky Way disk in these coordinates. Image Credit: Lahav et al., 2016

Here, I present my work on galaxy clusters done by stacking of Sunyaev-Zel'Dovich effect maps around known galaxy clusters and groups. Stacking analysis allows to detect the average SZ signal around a given mass halos, where the amplitude of its depends critically on the average thermal pressure profile of halos. From these kind of analysis we get a powerful probe of the distribution of hot gas in these system.
The outline of this report is as follow. In the second chapter, I describe the tools for extract useful information from the available data. In the third chapter I talk about the physics for understanding the analyses and results. As the last chapter, I present my work and I discuss how to interpret the results.

2 CosmoSIS

After completing all the acquisition, verification and reduction of the data from an survey, we have to transform compressed data sets into constraints on cosmological model parameters by comparing them to theoretical predictions. This is the last step of cosmological analysis and it's called *Cosmological parameter estimation* (CPE).
Usually the cosmology analysis applies a Bayesian approach to CPE. A likelihood function is used to assess the probability of the data, that were actually observed, given a proposed theory and values of that theory's parameters [1]. The way which the parameters changes is regulated by a sampling process: *Markov Chain Monte-Carlo*. The result is a distribution that describes the posterior probability of the theory's parameters.
The software that I used for analysis is called **CosmoSIS**: a relative new tool designed in order to incentivize the connection, the sharing and the development of cosmological analysis in the community. *Modularity* is the heart of the CPE software. In a modular approach to the cosmological likelihood, the separate physical calculations are divided into discrete pieces of code that do not have direct read and write access to the data each other. All data is then passed around via a single mechanism - the loading and saving of information in a single place. A likelihood function then becomes a sequence of modular processes, run one-by-one to form a pipeline. The last module(s) generates the final likelihood numbers. Any module in the sequence can be replaced at runtime by another calculation of the same step without affecting the others [1].

3 Cosmology

The cosmology is the science that study the origin an evolution of our Universe. The most current theory accept by the scientific community is the *Standard Model Cosmology* or ΛCDM, which is a parametrization of Universe's history.

3.1 Galaxy Clusters

A Galaxy Cluster is a cosmological structure that is made of hundreds to thousands of galaxies that are bound together by gravity with typical masses ranging from $10^{14}/10^{15}$ solar mass. They are one of the largest known gravitationally bound structures in the universe.
From dynamical study and x-ray observation, we know the relative proportion between the component of clusters: the baryonic matter is dominated by the Hot Gas component but the most dominant is the Dark Matter, approximately it is 80% of total mass. One way to obtain informations about the distribution of mass in a galaxy cluster is combing Sunyaev-Zel'Dovich effect data with the Galaxy Clusters maps.

3.2 Sunyaev-Zel'Dovich effect

Our present understanding of the primordial Universe is based, among other things, on the study of *cosmic wave of Background*. The CMB radiation comes out from the recombination epoch, when the first nuclei was formed, then it's one of the most important probe of early universe.
Theory of formation of large-scale structure predicts the existence of slight inhomogeneities in the first instances of the universe. The information of these inhomogeneities arrive to us by the study of CMB features. For example, the density inhomogeneities lead to temperature anisotropies in the CMB spectrum.
The cosmic microwave background's anisotropy is divided into two types: primary anisotropy, due to effects that occur at the last scattering surface and before; and secondary anisotropy, due to effects such as interactions of the background radiation with hot gas or gravitational potentials. Here, we are interested in the interaction between γ of CMB radiation and electrons belong to hot gas looked in the galaxy clusters: *Sunyaev-Zel'Dovich effect.*
When γ goes through intra-clusters medium, it interacts with high energy e^- by inverse Compton scattering: the electrons give energy to the CMB radiation and then we observe small spectral distortion. The SZ spectral distortion of the CMB could be expressed in terms of relative variation of temperature:

$$\frac{\Delta T_{SZE}}{T_{CMB}} = f(x) \int n_e \frac{k_B T_e}{m_e c^2} \sigma_T dl \tag{1}$$

where the integral along the line of sight usually is indicated as y, called *y-Compton parameter*, which for an isothermal cluster equals the optical depth, τ_e, times the fractional energy gain per scattering, σ_T is the Thomson cross-section, n_e is the electron number density, T_e is the electron temperature, k_B is the Boltzmann constant, $m_e c^2$ is the electron rest mass energy. As we can see, the magnitudes of SZ effect is proportional to the y-Compton parameter, which roughly speaking is a measure of the gas pressure integrated along the line-of-sight.
Since Compton-y is proportional to the depth of galaxy cluster's gravitational field, if we integrate the SZ signal around all the halos we obtain a quantity strong correlated with the mass of halo it-self.

3.3 Halo model of the Halo-SZ Cross Correletion

Stacking cosmic microwaves background maps around known galaxy clusters and groups provides a powerful probe of the mass distribution trough Sunyaev-Zel'Dovich effect. From now to on, I use *halo model framework* developed by Cooray and Sheth [2].
The key quantity of interest is the average y-Compton parameter at the projected co-moving distance r from a halo of mass M. More precisely, the significant quantity is the *excess* y-Compton parameter, with respect to the global-average y, around a certain halo. This can be written as an integral over the halo-pressure cross-correlation function:

$$\xi_{y,g}(r|M) = \frac{\sigma_T}{m_e c^2} \int_{-\infty}^{+\infty} \frac{d\chi}{1+z} \xi_{h,P}(\sqrt{\chi^2 + r^2}|M) \tag{2}$$

Here σ_T is the Thomson scattering cross section, $m_e c^2$ is the rest mass energy of an electron, z is the redshifth of galaxy cluster, $\xi_{h,P}$ is the halo-pressure cross-correlation function and the integral is done over co-moving coordinates along the line-of sight χ to the

halo. For simplicity, I omitted redshift labels although the quantity depends on the redshift of group/halo [3].
The halo-pressure cross correlation function could be calculated started from the gas/electron thermal pressure as a function of distance from the halo center - *the pressure profile* - for halos of various mass [4].

$$P_{fit}(x) = P_0 \left(\frac{x}{x_c}\right)^{\gamma} \left[1 + \left(\frac{x}{x_c}\right)^{\alpha}\right]^{-\beta} \ , \ x \equiv \frac{r}{r_{200}} \tag{3}$$

Here P_0, x_c, γ, α, and β are fitting parameters, with r_{200} denoting the radius at which the average matter density within the halo reaches 200 times the critical density. It is possible to treat each parameter as a separable function of mass and redshift. For a generic parameter A, it is described by the following formula:

$$A = A_0 \left(\frac{M_{200}}{10^{14} M_{\odot}}\right)^{\alpha_m} (1+z)^{\alpha_z} \tag{4}$$

Battaglia profile (Eq.4) is a generalization of Navarro, Frenk and White (NFW) profile. In the formula is implied the overall normalization P_{200}: self-similar amplitude of pressure profile. This is useful in order to compare different profile and to facilitated cluster stacking.
The halo-pressure correlation function describes the average excess pressure around halos as a function of the distance from their center. It is done by two contribution: one-halo term describes the pressure from the hot gas in the halo itself; two-halo term corresponds to the contribution from correlated neighboring halos.
The one-halo term is precisely the electron thermal pressure profile $P_e(r|M)$ which is proportional to the total thermal pressure profile $P(r|M)$. For one-halo term playing an important role the primordial helium fraction since the $P_e(r|M)$ depends on it.

$$\xi_{h,P}^{one-halo}(r|M) = P_e(r|M) \tag{5}$$

With r denoting the co-moving coordinate separation from halo center.
The two-halo term is more complicate respect the previous one, it is calculated starting from halo-pressure power spectrum and then using Fourier transform:

$$\xi_{h,P}^{two-halo}(r|M) = \int_0^{+\infty} \frac{dk}{2\pi^2} k^2 \frac{sin(kr)}{kr} P_{h,p}(k) \tag{6}$$

The total correlation function is given by the sum of these two term. Finally, by the integral in Eq.2, you get the average Compton-y parameter around a halo of mass M.

4 Methodology

In these analysis I used the y-maps of the thermal Sunyaev-Zel'Dovich effect, made by Planck collaboration, combined with Galaxy Clusters location comes from *DES*. The data were divided in redshifts bins and in richness bins, in order to see if it is possible to deduce some dependence from them:

$$z = [0.20, 0.35, 0.50, 0.65] \tag{7}$$

$$\lambda = [20, 30, 45, 60, 1000] \tag{8}$$

Before starting with the results, is important introduce the methodology which I used for valuating the cross-correlation function or the average y-Compton parameter.

4.1 The expected y(r) signal in lambda bins

Selecting clusters according to λ ranges means that the galaxy clusters in different richness bins will have different mass distributions. For clusters in the λ range of λ_1 to λ_2 at redshift z, we denote their mass distribution as $P(M|z, \lambda_1, \lambda_2)$. Assuming that the mass distribution is know, the average y(r) signal is:

$$\overline{y}_{r_0,z,\lambda_1,\lambda_2} = \int y_{r_0} P(M|z, \lambda_1, \lambda_2) dy_{r_0} \tag{9}$$

where r_0 is a certain point from the center of galaxy clusters. With probability chain rule, and assuming that richness and y-Compton parameter are not correlated, we can write:

$$\begin{aligned} \overline{y}_{r_0,z,\lambda_1,\lambda_2} &= \int P(M|z, \lambda_1, \lambda_2) dM \\ &\int y_{r_0} P(y_{r_0}|M, z) dy_{r_0} \end{aligned} \tag{10}$$

The second integral is Battaglia profile (Eq.3), up to now we indicate it by $y_{r_0}(M, z)$; the first integral instead could be written as:

$$P(M|z, \lambda_1, \lambda_2) = \int_{\lambda_1}^{\lambda_2} P(M|\lambda, z)P(\lambda, z)d\lambda = \int_{\lambda_1}^{\lambda_2} P(\lambda|M, z)P(M, z)d\lambda \quad (11)$$

where we use again the probability chain rule. Here $P(M, z)$ is the halo mass function at redshift z and it is independent by the integration variable, and $P(\lambda|M, z)$ describes the richness distribution for halos of mass M and redshift z. Now, the mean point of our analysis is to assume that the richness follows a log-Normal distribution with mean and standard deviation as follow:

$$\ln \lambda \sim N(\mu; \sigma^2_{\ln(\lambda|M)}) = N(A_\lambda + B_\lambda \ln M; \sigma^2_{\ln(\lambda|M)}) \quad (12)$$

where we can use the $\lambda - M$ relation [5] to link certain value of richness with the corresponding mass value:

$$M(\lambda, z) \equiv < M|\lambda, z >= M_0 \left(\frac{\lambda}{\lambda_0}\right)^{F_\lambda} \left(\frac{1+z}{1+z_0}\right)^{G_z} \quad (13)$$

where $M_0 = 2.35 M_\odot$, $F_\lambda = 1.12$ and $G_z = 0.18$ are parameter of the model with pivot values $\lambda_0 = 30$ and $z_0 = 0.5$. In order to generalized the scatter function, theoretical study suggest to use:

$$\sigma^2_{\ln(\lambda|M)} = \frac{1}{e^{\mu+\frac{1}{2}\sigma_0^2}} + \sigma_0^2 \quad (14)$$

Once again, σ_0 is a parameter of the theory. With some mathematical passages, we can rewrite the Eq.11:

$$P(M|z, \lambda_1, \lambda_2) = \int_{\lambda_1}^{\lambda_2} P(M|\lambda, z)P(\lambda, z)d\lambda \quad (15)$$

$$= \frac{1}{2}[erfc(\chi_1(M) - erfc(\chi_2(M))]P(M|z) \quad (16)$$

where we use the following notation:

$$\chi_{1/2}(M) = \frac{\log M_{1/2} - \log M}{\sqrt{2\sigma^2_{\log(\lambda_{1/2}|M)}/B^2_\lambda}} \quad (17)$$

Finally, The expected y(r) signal in a certain richness bin, namely the total cross correlation function between SZ maps and galaxy cluster groups, is equal:

$$\bar{y}_{r_0,z,\lambda_1,\lambda_2} = \int \frac{1}{2}[erfc(\chi_1(M) - erfc(\chi_2(M))] \times P(M|z) y_{r_0}(M, z)dM \quad (18)$$

5 Data analysis and result

The analysis was divided using two sets of data: the first set contains theoretical data, called mock data, and the second set contains the real data, namely observational data. In both of this analysis I evaluated the contribution of one halo term against one plus two halo terms, considering the standard deviation of log-Normal distribution of λ constant for each richness bins, and using the new standard deviation function (Eq. 14). I tried to recover some reasonable constraint in two parameters: the amplitude of pressure profile (Eq. 3, 4) labeled with A_p, and σ_0 which governs the behavior of the new standard deviation function.
The first analysis is necessary for understanding if the software works well and if the new standard deviation function (Eq. 14) is correct. In addition, as we will see from the using of mock data it is possible to recover a lot interesting informations. Furthermore, using the real data is always important and allows to test our theoretical model and to obtain physical constraints in the parameters.

5.1 Mock Data

Instead of using the experimental data for *MCMC* analysis, I used the output of the model with certain value of parameters in order to test the fitting code. I produced mock data sets with different configurations, which are listed in the table below. This data have been fitted by the corresponding model in a Monte Carlo Markov Chain. On the following plots, we present the constraints on A_p obtain with the first columns of Tab. 2. We see that we recover, for both one halo and one plus two halo term, approximately the same parameter value that I used for generating the mock data:

	Data sets	
1 halo	A_p = 18.1 ; no σ-relation	A_p = 18.1 ; σ_0 = 0.225
1 and 2 halos	A_p = 18.1 ; no σ-relation	A_p = 18.1 ; σ_0 = 0.225

Figure 2: mock data used in the first *MCMC* analysis.

one halo term: $A_p = 18.023^{+0.208}_{-0.091}$;

one and two halo terms: $A_p = 18.085^{+0.073}_{-0.328}$.

This tells us that the software works well. The

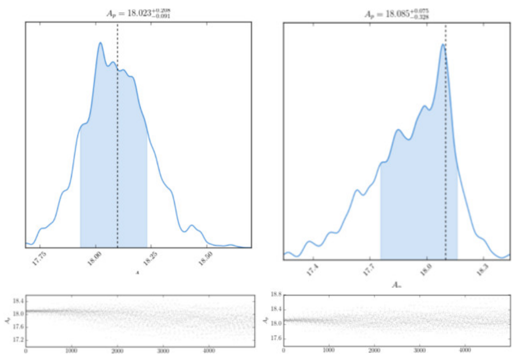

Figure 3: Results from *MCMC* analysis using theoretical data coming from the models without the new scatter function.

same *MCMC* analysis was done for the second column, this time in order to see if, with the adding of the scatter relation, the software still recovers reasonable results:

one halo term: $A_p = 19.49^{+0.44}_{-0.32}$;

one and two halo terms: $A_p = 18.085^{+0.073}_{-0.328}$.

We can see the results in the next plots. Since the constraint value in A_p is similar to 18.1, the scatter relation is correct and this is another proof that the software works well. Other information we can understand from this analysis: the particular shape of the the random walk in the parameters space suggest that the *MCMC* diverge than, as of now, we can say that A_p and $sigma_0$ can't binded at the same time.
The last study that I did with mock data was an *MCMC* analysis with only one halo term, using the theoretical data generated with one plus two halo model. Since from the observation we never

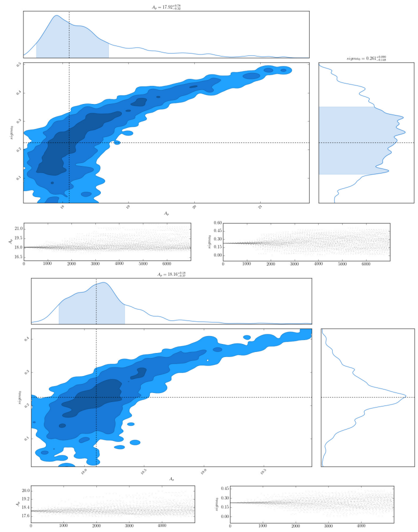

Figure 4: Results from *MCMC* analysis using theoretical data coming from the models with the new scatter function.

can separate the two contribution, this kind of analysis allows us to determine how much the difference is between the constraint value of certain parameter and its real value when we try to fit the one plus two halo data with only one halo model. The A_p values constrained from this analysis are:

without scatter relation: $A_p = 22.11^{+0.26}_{-0.33}$ with positive bias of 22.25%;

with scatter relation: $A_p = 19.49^{+0.44}_{-0.32}$ with positive bias of 8.76%.

5.2 Observational Data

In the end, the same approach was applied to the real data. Then, with the last *MCMC* analysis I tryed to extrapolate some reasonable constraints from the real data using only the one halo model, one time without the scatter relation and another time with the scatter relation.

In the analysis without the scatter relation, we recover a value of A_p smaller with respect to the simulations study. More precisely, $A_{p,observational\ data} = 16.07^{+0.18}_{-0.26}$ against $A_p = 18.1$ [Battaglia et al.]. In addition, the real value of A_p should be approximately 22% less

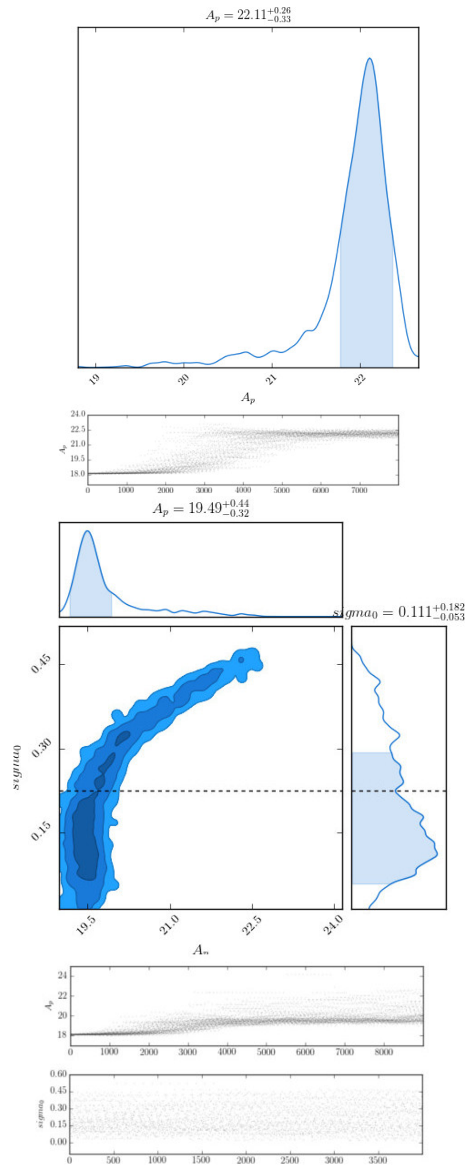

Figure 5: Constraint plots for bias analysis.

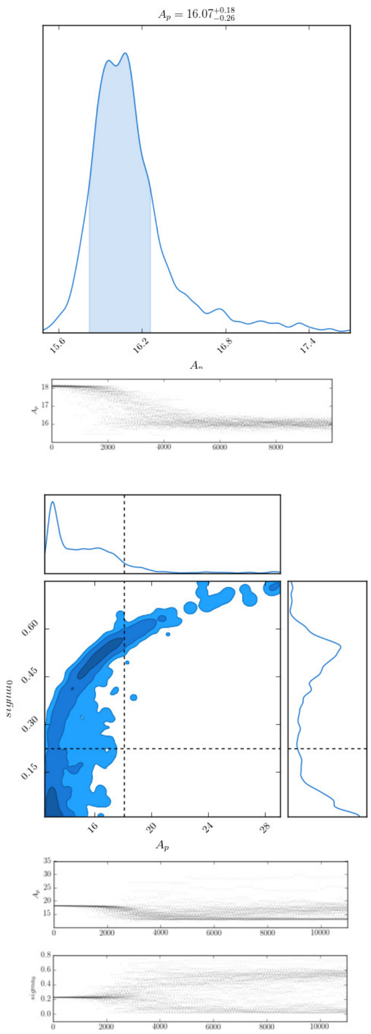

Figure 6: *MCMC* analysis of one halo term using observational data. In the first plot A_p is free parameter, and in the second case A_p and σ_0 are free parameters.

with respect to the constraint value obtained here.
Instead, form the analysis with the scatter relation, namely the second plot, we can't observe some limit in the parameters. Moreover, the shape of random walk in the parameters space is equal to the shape in mock data analysis.

5.3 Results

The main goal of my work is to demonstrate if the model and the software were able to perform parameter estimation. In view of the fact that we recovered the parameters input values in the mock analysis tell us that the software works well even with the new scatter relation. In addition, since the analysis with only one halo term gave reasonable results, in a future work we could concentrate most of all on this term so to make the analysis software faster.

Moreover, from the *MCMC* analysis with the observational data we obtained a smaller constraint value on A_p, considering the bias, with respect to the value which comes from simulations study. Then, we can claim that simulation study is not so accurate.
Eventually, from both sets of data we understand that A_p and σ_0 are strongly correlated and highly degenerated. Then, they can't be constrained at the same time.
One way to proceed in the future could be study more, from theoretical point of view, the behavior of A_p and σ_0 in order to find some physical relation which helps to put some strong limit. Also, one upgrade could be using new observational data like South Pole Telescope data, which are of excellent quality. Improving the observational data set will allows us to extend the parameter space.

References

[1] J. Zuntz, M. Paterno, E. Jennings, D. Rudd, A. Manzotti, S. Dodelson, S. Bridle, S. Sehrish, and J. Kowalkowski, "CosmoSIS: Modular cosmological parameter estimation," *Astronomy and Computing*, vol. 12, pp. 45–59, Sept. 2015.

[2] A. Cooray and R. Sheth, "Halo models of large scale structure," , vol. 372, pp. 1–129, Dec. 2002.

[3] V. Vikram, A. Lidz, and B. Jain, "A Measurement of the Galaxy Group-Thermal Sunyaev-Zel'dovich Effect Cross-Correlation Function," , vol. 467, pp. 2315–2330, May 2017.

[4] N. Battaglia, J. R. Bond, C. Pfrommer, and J. L. Sievers, "On the Cluster Physics of Sunyaev-Zel'dovich and X-Ray Surveys. II. Deconstructing the Thermal SZ Power Spectrum," , vol. 758, p. 75, Oct. 2012.

[5] P. Melchior, D. Gruen, T. McClintock, T. N. Varga, E. Sheldon, E. Rozo, A. Amara, M. R. Becker, B. A. Benson, A. Bermeo, S. L. Bridle, J. Clampitt, J. P. Dietrich, W. G. Hartley, D. Hollowood, B. Jain, M. Jarvis, T. Jeltema, T. Kacprzak, N. MacCrann, E. S. Rykoff, A. Saro, E. Suchyta, M. A. Troxel, J. Zuntz, C. Bonnett, A. A. Plazas, T. M. C. Abbott, F. B. Abdalla, J. Annis, A. Benoit-Lévy, G. M. Bernstein, E. Bertin, D. Brooks, E. Buckley-Geer, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, C. E. Cunha, C. B. D'Andrea, L. N. da Costa, S. Desai, T. F. Eifler, B. Flaugher, P. Fosalba, J. García-Bellido, E. Gaztanaga, D. W. Gerdes, R. A. Gruendl, J. Gschwend, G. Gutierrez, K. Honscheid, D. J. James, D. Kirk, E. Krause, K. Kuehn, N. Kuropatkin, O. Lahav, M. Lima, M. A. G. Maia, M. March, P. Martini, F. Menanteau, C. J. Miller, R. Miquel, J. J. Mohr, R. C. Nichol, R. Ogando, A. K. Romer, E. Sanchez, V. Scarpine, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, F. Sobreira, M. E. C. Swanson, G. Tarle, D. Thomas, A. R. Walker, J. Weller, and Y. Zhang, "Weak-lensing mass calibration of redMaPPer galaxy clusters in Dark Energy Survey Science Verification data," , vol. 469, pp. 4899–4920, Aug. 2017.