Detection of Missing Baryons in Galaxy Groups with Kinetic Sunyaev–Zel’dovich Effect

S. H. Lim1,2, H. J. Mo1, Huiyuan Wang3,4, and Xiaohu Yang5,6

1 Department of Astronomy, University of Massachusetts, Amherst, MA 01003-9305, USA
2 Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada; slim@phas.ubc.ca
3 Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, Hefei, Anhui 230026, People’s Republic of China
4 School of Astronomy and Space Science, University of Science and Technology of China, Hefei 230026, People’s Republic of China
5 Department of Astronomy, Shanghai Key Laboratory for Particle Physics and Cosmology, Shanghai Jiao Tong University, Shanghai 200240, People’s Republic of China
6 IFSA Collaborative Innovation Center and Tsung-Dao Lee Institute, Shanghai Jiao Tong University, Shanghai 200240, People’s Republic of China

Received 2019 September 24; revised 2019 December 3; accepted 2019 December 17; published 2020 January 24

Abstract

We present the detection of the kinetic Sunyaev–Zel’dovich effect (kSZE) signals from groups of galaxies as a function of halo mass down to $\log(M_{500}/M_\odot) \sim 12.3$ using the Planck cosmic microwave background maps and stacking about 40,000 galaxy systems with known positions, halo masses, and peculiar velocities. The signals from groups of different mass are constrained simultaneously to take care of the projection effects of nearby halos. The total kSZE flux within halos estimated implies that the gas fraction in halos is about the universal baryon fraction, even in low-mass halos, indicating that the “missing baryons” are found. Various tests performed show that our results are robust against systematic effects, such as contamination by infrared/radio sources and background variations, beam size effects, and contributions from halo exteriors. Combined with the thermal Sunyaev–Zel’dovich effect, our results indicate that the “missing baryons” associated with galaxy groups are contained in warm-hot media with temperatures between $10^5$ and $10^6$ K.

Unified Astronomy Thesaurus concepts: Galaxy formation (595); Galaxy evolution (594); Galaxy dark matter halos (1880); Circumgalactic medium (1879); Extragalactic astronomy (506); Galaxy environments (2029); Galaxy groups (597); Galaxy clusters (584)

1. Introduction

According to the current scenario of galaxy formation, galaxies form and evolve in dark matter halos (see Mo et al. 2010, for a review). As a dark matter halo forms in the cosmic density field, the cold gas associated with it falls down into its potential well and gets shock-heated, eventually forming a hot gaseous halo with a temperature roughly equal to the virial temperature of the halo. However, various processes, such as radiative cooling, star formation, and feedback from supernovae and active galactic nuclei (AGNs), can affect the evolution of galaxies and the properties of the gaseous halos, so the distribution of baryons may be very different from that of the dark matter. Indeed, observations have shown that both the hot gas fraction and the total baryon fraction in present-day galaxy systems are much lower than the universal baryon fraction, especially in low-mass systems (e.g., David et al. 2006; Gastaldello et al. 2007; Pratt et al. 2009; Sun et al. 2009). Even for massive clusters of galaxies, the distribution of the gas is found to be different from that of the dark matter, although the total amount of hot gas is found to be close to the universal fraction (e.g., Arnaud et al. 2010; Battaglia et al. 2012). It has been suggested that a significant portion of the “missing baryons” may be in the form of diffuse warm-hot intergalactic media (WHIMs), with temperatures in the range of $10^5$–$10^7$ K, within and/or around dark matter halos (e.g., Cen & Ostriker 1999; Davé et al. 1999, 2001; Smith et al. 2011). Attempts have been made to use X-ray observations to detect such WHIMs associated with filaments and walls in the cosmic web (e.g., Werner et al. 2008; Fang et al. 2010; Eckert et al. 2015).

The Sunyaev–Zel’dovich effect (SZE; Sunyaev & Zeldovich 1972) offers another promising way to probe the WHIMs (Hill et al. 2018; Lim et al. 2018b; de Graaff et al. 2019; Tanimura et al. 2019). As the cosmic microwave background (CMB) photons pass through galaxy systems, such as clusters and groups of galaxies, they are scattered by the free electrons in these systems. The effect produced on the CMB by the thermal motion of electrons is referred to as the thermal SZE (tSZE), while that produced by the bulk motion of electrons is called the kinetic SZE (kSZE). Thus, cross-correlating galaxy systems (clusters and groups of galaxies, collectively referred to hereafter as galaxy groups) and their SZE in the CMB provides an avenue to probe the WHIMs associated with dark matter halos.

Great efforts have been made to measure the tSZE from observations and use it to constrain the gas associated with galaxy systems. Planck Collaboration XI (2013) used the Planck multifrequency CMB temperature maps and dark matter halos identified based on isolation criteria to study the tSZE down to a halo mass of $\sim 4 \times 10^{14} M_\odot$. Remarkably, they found that the hot gas fraction in halos is independent of halo mass, as expected from the simple self-similar model. Similarly, Greco et al. (2015) used the locally brightest galaxies to represent dark matter halos to extract the tSZE and found that their results are consistent with the self-similar model. In a recent paper, Lim et al. (2018a) used a large sample of galaxy groups (Lim et al. 2017b) to extract the tSZE associated with galaxy systems from the Planck Compton parameter map (Planck Collaboration XXII 2016). By stacking about half a million galaxy systems, they were able to obtain the tSZE as a function of halo mass down to $\log(M_{500}/M_\odot) \sim 12$, where $M_{500}$ is the halo mass enclosed by a radius in which the mean mass density is 500 times the critical density. They found that the thermal contents of the gas...
in low-mass halos are much lower than that expected from the cosmic mean baryon fraction and the virial temperature of halos, in contrast to the results obtained by Planck Collaboration XI (2013) and Greco et al. (2015). Hill et al. (2018) and Tests by Lim et al. (2018a) suggest that the discrepancy may originate from different treatments of projection effects in the measurements.

Detecting the kSZE signals from CMB observations is not a trivial task. First, the signals are weak: the kSZE amplitude is about 2 orders of magnitude smaller than the primary CMB fluctuation and typically 1 order of magnitude lower than the tSZE. As such, stacking a large number of similar galaxy systems is needed to detect the effect. Second, since the kSZE is directly proportional to the radial peculiar velocity of the galaxy system, and since the peculiar velocities of different systems have a symmetric distribution around zero, stacking individual systems without using the peculiar velocity information leads to cancellation rather than enhancement of signals. Third, the large beam sizes of current CMB experiments require assumptions of the locations and gas profiles of the galaxy systems to be stacked in order to extract the kSZE they produce. Finally, since the observed effects are projected on the sky, signals from low-mass systems may be contaminated by projections of the more massive systems along the same line of sight (LOS).

The detection of the kSZE has so far been made only for cluster-sized individual systems (e.g. Sayers et al. 2013; Adam et al. 2017) or from statistical measurements based on, e.g., the pairwise and cross-correlation methods (Hand et al. 2012; Hernández-Monteagudo et al. 2015; Hill et al. 2016; Planck Collaboration XXXVII 2016; Schaan et al. 2016; Soergel et al. 2016; De Bernardis et al. 2017). Using peculiar velocity fields reconstructed from galaxy distributions and the aperture photometry, Hernández-Monteagudo et al. (2015), Planck Collaboration XXXVII (2016), and Schaan et al. (2016) found that a significant fraction of baryons may be associated with the large-scale structure traced by galaxies. A similar conclusion was reached by Hill et al. (2016) by cross-correlating galaxies with CMB maps. The signals measured in these investigations are the averages over individual galaxies in the galaxy samples used, including the effects of the gas both confined to galaxy halos and unbound over large scales. These results, therefore, constrain the total amount of free electrons associated with the large-scale structure traced by galaxies but cannot be interpreted directly in terms of baryon fractions in halos of different masses. Thus, the missing baryon problem on halo scales, which has important implications for galaxy formation in dark matter halos, remains unresolved.

In this paper, we investigate the kSZE from halos of different mass, using galaxy groups and the Planck temperature maps, by extending methods similar to those used in Lim et al. (2018a) to the kSZE. As described below, our analysis differs from earlier studies in that our halo-based group catalog with reliable halo mass information allows us to probe the kSZE and baryon fractions in halos of different masses, and that we simultaneously constrain signals from different groups so that the LOS contamination by projection effects is taken into account. In addition, the combination of the kSZE measurements here with the tSZE measurements obtained in Lim et al. (2018a) allows us to obtain not only the total mass but also the effective temperature of the WHIMs associated with galaxy systems.

The structure of the paper is as follows. We describe the observational data for our analysis in Section 2 and our method to extract the kSZE in Section 3. We present our main results and inferences from combining the kSZE with the tSZE in Section 4. Finally, we summarize and conclude in Section 5.

2. Observational Data

2.1. The Planck CMB Map

The Planck observation (Tauber et al. 2010; Planck Collaboration I 2011) measures the all-sky CMB anisotropy in nine frequency bands from 30 to 857 GHz with angular resolutions ranging from $3\'\$ to $5\'$. In our analysis for the kSZE, we use the 100, 143, and 217 GHz channel maps from the Planck 2015 data release. To minimize Galactic contamination, the brightest 40% of the sky is masked using the masks provided in the data release. We also eliminate known radio and infrared point sources using the corresponding masks. From the reduced maps, subtractions are made of the tSZE,

$$\left(\frac{\Delta T}{T_{\text{CMB}}}\right)_{\text{tSZE}} = g(x)y \equiv g(x) - \frac{\sigma_T}{m_e c^2} \int P dl,$$

where $T_{\text{CMB}} = 2.7255$ K, $y$ is the Compton parameter, $g(x) = x \coth(x/2) - 4$ is the conversion factor at a given $x \equiv h\nu/(k_B T_{\text{CMB}})$, $\sigma_T$ is the Thompson cross section, $c$ is the speed of light, $m_e$ is the electron rest mass, and $P_e = n_e k_B T_e$ is the electron pressure, with $n_e$ and $T_e$ being the free electron density and temperature, respectively. The electron pressure is integrated over the path length, $dl$, along the LOS. We adopt the Compton parameter $y$ from the Planck Needlet Independent Linear Combination (Remazeilles et al. 2011) all-sky $y$-map (Planck Collaboration XXI 2016), which is constructed from the full mission data set of Planck, using a combination of different frequency maps to minimize the primary CMB fluctuations and contamination from foreground sources. Integrating over the Planck bands gives the conversion factors $g(x)T_{\text{CMB}} = -4.031$, $-2.785$, and $0.187$ K for the 100, 143, and 217 GHz maps, respectively. Readers are referred to the original papers for more details about the construction of the $y$-maps. Finally, a constant background is subtracted from each of the resulting maps to zero the average background over the full sky. As a test, we also applied the same analysis to the Planck Modified Internal Linear Combination Algorithm (Hurier et al. 2013) $y$-map, which is known to have a different level of dust contamination, and found no significant changes in our results.

The power spectra of dust emission from galaxies in the Planck maps at the frequencies relevant to the SZE analysis are more than 1000 times smaller than the CMB at $l \sim 100$ and subdominant in comparison to the primordial CMB for all $l$ (Planck Collaboration XXX 2014). In addition, the dust emission was found to be only weakly correlated with galaxy systems at low-$z$ (e.g., Vikram et al. 2017), as is expected from the fact that most of the dust emission comes from star-forming galaxies at higher redshift. Additional leakages of the dust emission through the subtraction of $y$-maps may be present in the maps but are very poorly constrained (Vikram et al. 2017). As long as such dust contamination is not correlated with the

https://pla.esac.esa.int
targets used for our analysis (see Section 2.2), it will only add noise to the data but will not bias our results, and the level of the noise will be much lower than that produced by the primary CMB fluctuations.

Thus, the final maps we use in our analysis still contain components other than the kSZE, such as the primary CMB, dust emission, instrumental noise, and residual contaminants. The uncertainties introduced by these contaminating components will be taken into account in our analysis, as described in Section 3.

2.2. Galaxy Groups

To extract the SZ signals associated with galaxy groups requires a well-defined group catalog. In our analysis, we use the group catalog of Yang et al. (2007), which is constructed from the Sloan Digital Sky Survey Data Release 7 (SDSS DR7: Abazajian et al. 2009) with the use of the halo-based group finder developed in Yang et al. (2005). All of the groups in the original catalog have accurate estimates of halo masses, spatial positions, and peculiar velocities. The halo masses of the groups are estimated from abundance matching based on the ranking of their characteristic luminosities. Tests using realistic mock catalogs show that the halo masses given by the group finder match well the true halo masses, with a typical scatter of 0.2–0.3 dex. Following conventions in SZ studies, we define a halo by a radius, $R_{500}$, within which the mean density is 500 times the critical density at the redshift of the halo. The corresponding halo mass is $M_{500}$. The masses provided in the group catalog, $M_{200}$, are converted to $M_{500}$ by assuming the Navarro–Frenk–White profile (Navarro et al. 1997) and a model for the halo concentration parameter (Neto et al. 2007). These properties of individual groups are used in a filter to extract the kSZE from galaxy groups over a large range of masses. They are also used to interpret the kSZE from galaxy groups and to model the halo concentration parameter ($n_{e,0}$, $c$, $\beta$).

The core radius of a group with concentration $c$, and $\beta = 0.86$ is the best-fit value obtained from South Pole Telescope cluster profiles (Plagge et al. 2010). Note that $r_{\text{vir}}$ and $c$ are determined by the halo mass and redshift of the galaxy group in question. In principle, the integrated signals extracted can depend on the assumed gas profile. However, because of the beam size of the Planck instrument, which is comparable to halo radius for a significant fraction of the groups in our sample, the impact of the specific choice for the spatial profile shape is significantly mitigated, as is discussed in Sections 3.2 and 4.2.

To extract the kSZE of a group from the Planck data, we apply a filter that takes into account the effect of the beam, $\hat{F}_i(k) = \hat{\gamma}(k) \hat{B}_i(k)$, where $\hat{F}_i(k)$ is the Fourier transform of the filter for each of the three frequencies, $I$, $Q$, and $\hat{\gamma}(k)$ is the Fourier transform of $\tau(R)$; and $\hat{B}_i(k)$ is the Fourier transform of the Gaussian beam function that mimics the convolution of the Planck observation in the frequency band $I$. We use the filter, $[F_i]$, defined in Equation (5) with redshift, virial radius, and concentration appropriate for the group, $i$, in question at the frequency $I$ to estimate the signal within the filter,

$$\{A_{B_i}\} = \int \{F_i(\theta)\} [D_i(\theta)] d^2 \theta,$$

where $\theta$ is the projected position relative to the group center in the sky, and $[D_i(\theta)]$ is the data map around the group.

The values of $\{A_{B_i}\}$ obtained this way can be affected by projection effects by other groups along the same LOS. Because of this, we do not use these values to represent the signal produced by individual groups and obtain the average quantities for groups of a given mass. Instead, we construct model maps that take into account the projection effects to compare with data. As described above, the kSZE signal expected from a given group is determined by its peculiar velocity and gas density profile, which is modeled by Equation (4). In our analysis, we use the reconstructed peculiar velocities described in Section 2.2. The virial radius, $r_{\text{vir}}$, and concentration, $c$, of a group are given by its halo mass, as described above. To model the total ionized mass associated with a halo, we assume that the amplitude of the profile, $n_{e,0}$, depends only on halo mass. We specify the halo mass dependence with the values of $n_{e,0}$ at $\log(M_{500}/M_\odot) = 12.3$, 12.7, 13.1, 13.5, 13.9, and 14.3, together with linear interpolations in $\log(n_{e,0}) - \log(M_{500})$ space to predict the
profile amplitude for any given $M_{200}$. Thus, the model is completely specified by a set of six model parameters that give the profile amplitudes at the six values of $M_{200}$. The numbers of groups in the six mass bins are listed in Table 1, together with the median values of $M_{200}$, defined in a way similar to $M_{200}$. For a given set of model parameters, denoted collectively by $\Theta$, we generate a theoretical map by convolving the center of each group with the 2D profile appropriate for its halo mass, redshift, and peculiar velocity. We convolve the resulting map with the beam function to mimic the beam effect of the observation. Note that the model maps include the projection effects generated by the superposition of the profiles of halos along an LOS, because they are constructed with the positions of all groups. Then, exactly the same method is used to extract signals from the model maps. We put the filter at the center of each of the groups in the model map, again according to the halo mass and redshift of the group, to extract the corresponding model signal,

$$\{A_{M,i}\}_l = \int \{F_l(\Theta)\}_l \{M_i(\Theta)\}_l d^2\theta, \quad (7)$$

where $\{M_i(\Theta)\}_l$ stands for the model map at frequency $l$ around group $i$.

With these measurements, we compute a $\chi^2$-based likelihood function,

$$\mathcal{L}(\Theta|A_D) \propto \exp(-\chi^2/2), \quad (8)$$

where

$$\chi^2 = [A_D - A_M(\Theta)]^T \text{Cov}^{-1}[A_D - A_M(\Theta)], \quad (9)$$

where $A$ is the one-dimensional array with the elements being $\{A_{M,i}\}_l$ for all groups and in the three frequency bands, thus each containing $3N_{\text{grp,tot}}$ elements, with $N_{\text{grp,tot}}$ being the total number of groups. The covariance of the estimate between group $i$ in band $l$ and group $j$ in band $J$ is calculated from 1000 sets of random shifts and rotations of the filters relative to the CMB sky but with their relative positions fixed by applying the filters to the maps as in Equation (6). This way, the values returned by the filters measure the total noise fluctuation, while the spatial correlation of the noise among groups is retained. We use the average correlation of the values from these 1000 random sets to estimate the covariance matrix.

Note that the covariance computed this way takes into account all noise components that are not correlated with the targets and are homogeneous over the sky, including the primary CMB fluctuation, instrument noise, dust emission, point sources that are not masked, and residual tSZE and kSZE from sources that are not correlated with our targets. The primary CMB fluctuation is treated as noise because the maps we use are not cleaned of it. It is very difficult to separate the primary CMB from the kSZE because they have the same frequency dependence. Our noise estimate may also contain parts of the kSZE from the targets because the targets cover a significant fraction of the sky, so some filters from the 1000 sets of random shifts and rotations may end up in regions covered by the targets. The impact of this, however, must be negligible, given that the kSZE is much weaker than other components of the noise, in particular the primordial CMB. The error in the velocity reconstruction and the intrinsic variance of the kSZE at a given mass can, in principle, also be included in the covariance matrix. However, the uncertainty expected from those errors combined is only a few percent of the primary CMB fluctuation and can safely be ignored without affecting our results. Uncertainties produced by the contamination that is correlated with the targets are not included in our covariance matrix. We will test the effect of such contamination in Section 4.2.

To efficiently explore the parameter space, we make use of the MULTINEST method developed in Feroz et al. (2009), which implements the nested sampling algorithm developed in Skilling (2006). The posterior distribution of the model

| Bin No. | 1     | 2     | 3     | 4     | 5     | 6     |
|---------|-------|-------|-------|-------|-------|-------|
| $\log M_{200} (M_\odot)$ | 12.4  | 12.8  | 13.2  | 13.6  | 14.0  | 14.4  |
| Number of systems | 23,997 | 9795  | 3780  | 1287  | 346   | 58    |
| $\mathcal{K}_{200}(10^{-2}\text{arcmin}^2)^a$ | 0.19  | 0.29  | 0.95  | 3.5   | 12    | 25    |
| $\sigma(\mathcal{K}_{200})(10^{-2}\text{arcmin}^2)^b$ | 0.059 | 0.096 | 0.51  | 1.1   | 2.2   | 8.4   |
| $\delta(\mathcal{K}_{200})(10^{-2}\text{arcmin}^2)^c$ | 0.063 | 0.064 | 0.49  | 0.95  | 4.7   | 6.0   |
| $z^d$ | 0.029 | 0.041 | 0.058 | 0.081 | 0.11  | >0.12 |
| $\text{cov}(i, 1)^e$ | (0.058)$^2$ | -0.00080 | (0.091)$^2$ | -0.0027 | (0.35)$^2$ | -0.037 | (1.5)$^2$ | (5.1)$^2$ | (4.5)$^2$ |
| $\text{cov}(i, 2)$ | -0.00027 | -0.016 | -0.043 | -0.14 | -0.82 | -3.3 |
| $\text{cov}(i, 3)$ | -0.071 | -0.078 | -0.23 | -1.4 |

Notes.

$^a$ The average of the median values of the posterior distribution from the seven samples used for our test of systematic effects.

$^b$ The variance of the median values of the posterior distribution among the seven samples.

$^c$ The variance divided by $\sqrt{6}$ among the six independent subsamples, each containing groups in an $\sim 1/6$ portion of the sky, to represent the statistical uncertainty in the estimates.

$^d$ The redshift where the median $\theta_{200}$ of halos equals the Planck beam size of 10$'$.

$^e$ Here cov(i,j) is the covariance of $\mathcal{K}_{200}$ between the $i$th and $j$th bins, calculated for the sample of all groups, according to the bin number assignment in the first column of the table. The values are in units of 10$^{-4}$ arcmin$^4$. 

$\mathcal{K}_{200}$—$M_{200}$ Relation

$\{A_{M,i}\}_l = \int \{F_l(\Theta)\}_l \{M_i(\Theta)\}_l d^2\theta, \quad (7)$

$\chi^2 = [A_D - A_M(\Theta)]^T \text{Cov}^{-1}[A_D - A_M(\Theta)], \quad (9)$

$\mathcal{L}(\Theta|A_D) \propto \exp(-\chi^2/2), \quad (8)$

$\text{cov}(i, 1)^e$ (0.058)$^2$ -0.00080 (0.091)$^2$ -0.0027 (0.35)$^2$ -0.037 (1.5)$^2$ (5.1)$^2$ (4.5)$^2$
parameters, i.e., the values of $n_e$, at the six halo masses, is used to make inferences from the data.

The amplitude of the density profile obtained from the filters, together with the assumed spatial profile shape, can be used to estimate the total number of electrons within $R_{500}$ for a group of a given mass,

$$d_A(z)^2 K_{500} \equiv \sigma_T \int_{R_{500}} n_e dV, \tag{11}$$

where $d_A(z)$ is the angular diameter distance of the group in question. We also define the following quantity, analogous to the convention used in the tSZE studies,

$$\tilde{K}_{500} \equiv K_{500}(d_A(z)/500 \text{ Mpc})^2. \tag{12}$$

We emphasize, however, that we are not measuring $\tilde{K}_{500}$ directly from the data but rather using it to represent the amplitude of the density profile, $n_e$. Once $\tilde{K}_{500}$ is obtained, we can also obtain corresponding quantities within other choices of radius, such as $\tilde{K}_{200}$ within $R_{200}$.

### 3.2. Testing with a Mock Observation

In this section, we test our method described above by applying it to a mock observation and examining how our method recovers the true $K_{500}$. The mock Planck map is constructed from the Magneticum (Dolag et al. 2016), a set of cosmological gas simulations of various volumes and resolutions run with an improved version of GADGET-3 (Springel 2005) and WMAP7 cosmology (Komatsu et al. 2011). The specific run chosen for our analysis samples a box of $L = 352 \ h^{-1} \text{ Mpc}$ with $2 \times (1584)^3$ particles, which results in a mass resolution of $m_{DM} \sim 10^8 M_{\odot}$ for the dark matter particles. Based on the simulation, we construct the kSZE light-cone maps using the SMAC code (Dolag et al. 2005), which covers the redshift range of $0.0173 < z < 0.194$ and 900 deg$^2$ of the sky. We generate the kSZE based on the dark matter particles only, assuming that gas particles follow the dark matter distribution with an assumed ratio between the gas and total mass densities. We then generate the CMB temperature maps at the same three frequencies as used in our analysis and add the primordial CMB anisotropy using the CMB power spectrum. Finally, we degrade the map with the Planck beam function and add the instrument noise. To mimic the impact of the primordial CMB anisotropy and the instrument noise on our results based on a sample that is nine times larger in sky coverage than the mock sample considered here, we reduce their amplitudes by a factor of $\sqrt{9}$. We neglect other contaminating sources, such as dust emission, residual tSZE, and kSZE, because they are negligible compared to the primary CMB fluctuations at the frequencies considered here.

The three panels of Figure 1 show the results of $\tilde{K}_{500}$ obtained by applying our method to the three stages of the constructed temperature maps, in comparison to the values obtained directly from the simulation. The left panel shows the results for maps that do not contain beam smearing and noise, the middle panel contains only beam smearing, and the right panel contains both. The black solid lines show the result for all halos recovered by assuming the $\beta$ profile of Equation (4), which we use as our fiducial case. The shaded bands show the 68% range of the posterior distribution. The median $\tilde{K}_{500}$ recovered by our method is about 20% higher for halos with $M_{500} \lesssim 10^{13.5} M_{\odot}$ relative to the true value. The bias does not seem to depend strongly on whether or not beam effects and noise are included and is within the statistical uncertainty expected from the observational data (see the right panel).

Since our method uses projected profiles (see Equation (3)) to extract kSZE signals, it has the ability to suppress the kSZE signals contributed by the gas that appears to be associated with the halos in question but is located outside the halos. However, it is likely that a fraction of the signals detected by our method actually come from the gas that is not associated with the halos in our sample, given that the typical correlation length of the peculiar velocity field is large. This may explain why our method overestimates the kSZE for low-mass halos. To check the effect of such contamination further, we artificially reduce the gas fraction within halos so as to increase the importance of background contamination. To do this, we assume that the gas fraction, adopted in constructing the mock maps, depends on halo mass through a simple power-law relation, so that the gas fraction changes from the universal fraction for the most massive bin to one-third of the universal fraction for the least massive bin. The gas fraction associated with dark matter...
outside these halos is still assumed to be at the universal fraction. The results obtained by applying our method to the mock maps are shown by the yellow dashed line. As one can see, the bias in the recovered $\bar{K}$ is similar to that in the fiducial case, indicating that the use of filtering profiles can effectively suppress contributions from gas outside halos.

We have carried out another test in which the shape of the density profiles used to model $n_e(r)$ is directly measured from the simulation, instead of that given by the $\beta$-model. Since the gas density in the mock maps is directly proportional to the dark matter density field, the profile shape used here is similar to that of dark matter halos. The results, shown by the blue dotted-dashed lines, indicate that using the true profile shape can reduce the bias when the beam effect and noise are negligible. This happens because the dark matter halo profiles are more concentrated than the $\beta$-model, so the contribution from large projected distances is suppressed more effectively. However, as shown in the middle and right panels, the beam size and, particularly, the noise significantly reduce the sensitivity of the results to the choice of the profile shape.

Based on the tests presented above, we conclude that our method may be biased toward a higher KsZE flux by up to ∼20% for low-mass halos due to the contamination by the background gas, but the bias is not significant given the statistical uncertainty in the present data. Note that the tests here are aimed at assessing the contamination by the gas that is not associated with the groups/halos in our catalog. In a forthcoming paper (S. H. Lim et al. 2020, in preparation), we use realistic mock CMB maps and group samples to check the impact of a number of other uncertainties. We find that the net bias produced by the uncertainties in group identification and halo mass assignment is not significant in the results obtained from current observational data. We also find that an error of ∼90 km s$^{-1}$ in halo peculiar velocities, a value similar to that in the reconstruction used here, does not lead to any significant impact on our results.

4. Results

4.1. The $\bar{K}$–$M$ Relation

The results obtained from the entire sample are shown as yellow triangles in Figure 2. The data points represent the median values obtained from the posterior distribution, while the error bars are the 68th percentile range. The dashed line in Figure 2 shows the “self-similar” model prediction, in which the total number of electrons within $R_{500}$ is

$$N_{e,500} = [(1 + f_H)/2m_p] \cdot f_B M_{500},$$

where $f_H = 0.76$ is the hydrogen mass fraction, $m_p$ is the proton mass, and $f_B = \Omega_B/\Omega_m = 0.16$ is the universal baryon fraction. Our data points follow well the self-similar model, indicating that the total ionized gas fractions in halos of different masses are comparable to the universal baryon fraction (see Section 4.3 for details).

Figure 2 shows that the $\bar{K}$–$M$ relation is approximately a power law. This motivates another way to extract the $\bar{K}$ associated with groups. Here we assume that

$$\bar{K} = A \times (M_{500}/10^{13.5} M_\odot)^{\alpha}$$

to generate the model maps and use the same method of filtering and minimization of the $\chi^2$ to constrain $\alpha$ and A. To fit this relation, we neglect potential intrinsic scatter of the relation and uncertainties in halo identifications based on the test results described in Section 3.2. The result for the entire sample is shown in Figure 3 with a yellow triangle. As a comparison, each of the small dots shows the result of a random sample. Each random sample is constructed by shifting and rotating the group sample by some random amount relative to the Planck maps before applying the filter. In this case, the relative positions of individual groups and their spatial clustering are preserved in the random samples, but the cross-correlation between the groups and the KsZE signals is destroyed. The distribution of the 200 random realizations is around (A, $\alpha$) = (0, 0), as expected from a zeroed mean background. The symmetry relative to (0, 0) is because the peculiar velocities have a roughly symmetric distribution around zero. The dipolar
pattern of the random samples in (\(A, \alpha\)) space, that positive (negative) values of \(A\) tend to correspond to positive (negative) values of \(\alpha\), indicates the presence of residual background fluctuation on scales larger than individual groups. In this case, the flux associated with a group in a random sample due to background fluctuation is proportional to the angular size of the group, with approximately the same probability of being positive or negative. Thus, the distribution of random groups in the \(\tilde{K}_{500}\)--halo mass plane is expected to have a wedge-like pattern symmetric with respect to the halo mass axis. Since we have more lower-mass systems, the averages of the \(\tilde{K}_{500}\) in the lower-mass bins are closer to zero and have lower random fluctuations. This has the effect of increasing the opening angle of the wedge within which the average \(\tilde{K}_{500}\)--versus--halo mass relations from different random samples are confined. The assembly of straight lines, each covering the whole halo mass range in such a wedge, thus tends to show the dipolar pattern seen.

We have also used random samples in which each group in the observational sample is assigned a random position in the Planck sky and found very similar results. The results indicate that the detection of the \(\tilde{K}_{500}\) is very significant relative to the random samples, as we will quantify below after the uncertainties in the measurements are tested.

### 4.2. Tests of Uncertainties

A number of contaminating effects can affect our measurements. Here we present the analyses we have carried out to test the reliability of our results against the contamination. One source of uncertainty comes from the errors in the peculiar velocities of groups adopted in our analysis. As shown in detail by Wang et al. (2012) using realistic mock samples, the errors in the reconstructed peculiar velocities have a symmetric distribution around zero. Since the kSZE of a group is directly proportional to its peculiar velocity, the uncertainty in the peculiar velocities is not expected to produce any bias in our results but will contribute to noise in the measurements.

Another source of contamination is from fluctuations in the background and foreground, such as the primordial CMB, signals produced by background sources that are not included in our sample, and residual Galactic foreground. If the background/foreground fluctuations are not correlated with the groups in our sample, then the contamination is not expected to lead to any bias in our results but can increase the noise in our measurements. To test this, we divide the total sample into three subsamples, each containing groups in an \(\sim 1/3\) portion of the sky coverage according to their Galactic latitudes, and repeat the procedures to obtain the \(\tilde{K}_{500}\) for groups in each of the three subsamples. The results of the three subsamples, shown as the blue crosses in Figure 2, are similar for all of the mass bins. In addition, as shown in our test using random samples (Figure 3), any residual background/foreground fluctuations are well below the signals we detect. These tests show that this type of contamination does not have a significant impact on our results.

Yet another source of contamination comes from the emissions of the groups in the observational wave bands, such as radio and infrared emissions and the tSZE. Although we have attempted to subtract the tSZE from the observational data, some residuals may still exist. One unique property of the kSZE is that two similar groups with opposite peculiar velocities produce temperature fluctuations with opposite signs, in contrast to the contaminating emissions mentioned above, which should be independent of the sign of the peculiar velocity. To check that the signals we detect are indeed produced by the kSZE, we divide the entire sample into two subsamples, one with \(v_r > 0\) and the other with \(v_r < 0\), and tune the model parameters independently for the two subsamples to achieve the best match to the data. The results obtained for these two subsamples are shown in Figure 2 as the violet and green triangles, respectively. The two give consistent \(\tilde{K}_{500}\) in all of the mass bins. The fact that the subsamples of opposite peculiar velocities give similar \(\tilde{K}_{500}\)--mass relations (i.e., opposite signals in \(k\) defined in Equation (2)) suggests that the contamination by emission sources does not change our results significantly, and that the signals we detect are indeed the kSZE.

As detailed in Section 3.2, the large beam size of Planck may contaminate the results with the kSZE signals from the gas outside halos. To test the contamination, we divide all halos into “resolved” ones and “unresolved” ones. We define a redshift, \(z_r\), at which the median \(\theta_{200}\) of the halos in a given mass bin is equal to the Planck beam size, 10'. The values of \(z_r\) for individual mass bins are listed in Table 1. For the \(z < z_r\) (resolved) and \(z > z_r\) (unresolved) subsamples, we assume two independent sets of parameters and jointly constrain them. The results so obtained for the resolved subsamples are shown as red plus signs in Figure 2. Clearly, the results obtained from the resolved sample are in good agreement with those obtained from the total sample within the uncertainty.

Finally, our filter uses Equation (4) to model the gas profile, while the true gas profiles may be different. To examine how our results are affected by the assumed profile, we have done tests by increasing the values of \(r_c\) by a constant factor, \(\mu\). We found that our results do not change significantly as long as \(\mu < 1.5\). When \(\mu\) becomes larger than 1.5, the values of \(\tilde{K}_{500}\)
obtained start to change noticeably for low-mass groups. The test shows that our results are not very sensitive to the gas density profile, as may be expected from the low angular resolution of the Planck data. However, this also indicates that the current data are not able to provide significant constraints on the details of the gas profiles around halos. A similar conclusion is reached in our forthcoming paper (S. H. Lim et al. 2020, in preparation), where we test a range of profiles predicted from hydrodynamic simulations using detailed mock CMB maps constructed to mimic the Planck data.

To summarize, all seven samples we have analyzed above, namely all groups, groups with \( \nu_r > 0 \), groups with \( \nu_r < 0 \), groups in three parts of the sky, and groups with \( z < z_r \), give consistent results (as summarized in Figure 2), demonstrating that our detection of the kSZ is reliable. For reference, in Table 1, we list the averages and variances of \( \tilde{K}_{200} \) for individual halo mass bins obtained from these seven samples. Since these samples are not independent, the variances listed can only provide a measure of the systematic effects we have tested, but they cannot be used to represent the statistical uncertainties in the estimates of the averages of \( \tilde{K}_{200} \). To get some insight into the statistical uncertainties in our estimates, we divide the groups into six independent subsamples, each containing groups in an \( \sim 1/6 \) portion of the sky coverage, and obtain \( \tilde{K}_{200} \) versus halo mass for each of the six subsamples. The uncertainties in the estimates of \( \tilde{K}_{200} \) obtained from these six independent subsamples are also listed in Table 1. Table 1 also lists the covariance matrix of \( \tilde{K}_{200} \) among the six mass bins for the sample of all groups, obtained by drawing from the posterior distribution and then by calculating

\[
\text{cov} = \langle (\tilde{K}_{200} - \langle \tilde{K}_{200} \rangle)(\tilde{K}_{200} - \langle \tilde{K}_{200} \rangle)^T \rangle,
\]

where \( \langle \cdots \rangle \) denotes the mean. As expected from the nature of our simultaneous constraining of the model parameters, the off-diagonal elements are all negative. To estimate the joint signal-to-noise from the six mass bins, we use the covariance matrix to define a multivariate normal distribution function and calculate the cumulative probability of a null detection. The probability of our measurements thus calculated corresponds to a significance level of about 6\( \sigma \).

The results with the power-law model (see Equation (14)) for all seven samples we have analyzed above are shown as the colored points in Figure 3. The data points are clustered in a region \( (A, \alpha) \in (0.018 \pm 0.0039, 1.1 \pm 0.4) \). The dispersion among the seven samples is comparable to that of \( A \) at a given \( \alpha \) obtained from the 200 random samples, indicating that the errors in the estimates are dominated by fluctuations in the background and foreground. With this dispersion to represent the uncertainty in the results, our detection is at a level of more than 5\( \sigma \), consistent with the estimate from the binned data given above.

4.3. The Gas Fraction and Temperature

Figure 4 shows the gas fraction within \( R_{200} \) obtained from the best-fit parameters and assumed gas profile. The data points are the averages of and dispersion among the seven samples, while the shaded band contains the predictions of the power laws enclosed by the ellipse in Figure 3. We also show the results corrected for the LOS contamination with the black open circles according to the black curve in the left panel of Figure 1. The inferred gas fraction is consistent with the universal baryon fraction, given the uncertainty of the data, and is much higher than the baryon fraction in stars and cold gas (shown as the dotted line).

Recently, Lim et al. (2018a) measured the tSZ produced by galaxy groups using the nearly all-sky group catalog of Lim et al. (2017b), which was constructed by applying the halo-based group finder to four large redshift surveys. The gas fraction inferred from the tSZ assuming the virial temperature, \( T_{\text{vir}} = \mu m_p G M_{200}/2 k_{\text{s}} R_{200} \), where \( \mu = 0.59 \) is the mean molecular weight, is shown in Figure 4. This fraction is much lower than that given by our kSZ data except for the most massive groups, indicating that the average temperature of the gas responsible for the kSZ in lower-mass groups is much lower than the virial temperature. The effective temperature, estimated by combining the gas mass obtained from the kSZ and the thermal energy content given by the tSZ, is shown as a function of halo mass in Figure 5. The derived effective temperature is about \( 10^5 - 10^6 \) K for halos with \( M_{200} \lesssim 10^{13.5} \) M\(_{\odot}\) and much lower than the corresponding virial temperatures. This relatively low temperature has its origin in the relatively low thermal pressure measured in Lim et al. (2018a), and we refer the reader to that paper for a detailed discussion about the comparison with other tSZ measurements.

The temperature obtained here can be compared with that obtained from X-ray observations. As shown in Pratt et al. (2007), the gas temperature in clusters of galaxies appears to decline in the outer parts. Since the effective temperature inferred from the SZE is sensitive to the low-density gas in the outer parts, this may partly explain the lower temperature found here than that inferred from X-ray observations. In addition, as
shown in Wang et al. (2014), for a given halo mass, the scatter in the X-ray luminosity is very large; at $M_{200} \sim 10^{13.5-14} M_\odot$, the scatter is more than 1 order of magnitude (see their Figure 7). As our SZE measurement is the average over all systems of a given halo mass, the thermal content inferred from it is expected to be lower than that obtained from X-ray-selected samples (e.g., Planck Collaboration X 2011).

5. Summary and Conclusion

We have examined the kSZE from gas in dark matter halos associated with galaxy groups as a function of halo mass down to log($M_{300}/M_\odot$) $\sim$ 12.3. Our analysis uses the stacking of about 40,000 galaxy groups to extract the kSZE from the Planck temperature maps in three different frequency bands and employs a filter to take into account the beam effect. The filters are applied simultaneously for individual groups so as to minimize the projection effects of halos along the same LOS. Accurate reconstructed peculiar velocities of the groups are used so that we can reliably convert the observed kSZE to the amounts of ionized gas associated with galaxy groups. A number of tests are done to examine the uncertainties in our results from residual background/foreground fluctuations, contamination by the tSZE and emissions from galaxy groups, the large beam size of Planck observations, contamination by large-scale coherent motion, and the gas density profile adopted. We found that our results are robust against these potential sources of uncertainties for the current data. In a forthcoming paper, we will check in detail the significance of these uncertainties by applying the methods in our analysis and the literature to the SZE light-cone maps generated from hydrodynamic simulations.

The strength of the kSZE as a function of halo mass is found to be consistent with the “self-similar” model, in which the baryon fraction is independent of halo mass, suggesting that the “missing baryons” on halo scales are found by the kSZE. Combined with the tSZE measured for galaxy groups by Lim et al. (2018a), our results indicate that the gas temperatures in low-mass halos are much lower than the corresponding halo virial temperatures. This suggests that it is the low temperature of the gas, not the total number of baryons, that is responsible for the low thermal energy contents in low-mass halos found in tSZE and X-ray observations. Our results, therefore, provide direct support for the hypothesis that the missing baryons in galaxy groups are contained in the WHIMs with temperatures of $\sim 10^3-10^8$ K.

Our results also demonstrate the potential of using the SZE to study both circumbulge media and the galaxy formation processes that produce them. Such studies have advantages over absorption line studies, in that they are not limited to a small number of LOSs, and gas metallicity and ionization states are not needed to obtain the total gas mass. In the future, when high-resolution SZE data are available, the same analysis as carried out here can be used not only to constrain the total amount of ionized gas associated with galaxy groups (dark matter halos) but also to investigate the density and temperature profiles of the gas around them. One may also use galaxy groups with different star formation and/or AGN activities to study how the ionized gas distribution is affected by these activities. Clearly, the synergy between the SZE and observations of galaxy systems in other wave bands should be exploited in the future to provide detailed information about both the WHIMs and the galaxy formation processes that produce them.

We thank the Planck Collaboration for making the full-sky maps public and Klaus Dolag for providing the SZE light-cone maps. We thank Eiichiro Komatsu, Colin Hill, Neil Katz, Volker Springel, Mark Halpern, Gary Hinshaw, Ludo Van Waerbeke, Douglas Scott, and the referee for helpful discussions and comments, which greatly improved the paper. This work is supported by the China 973 Program (No. 2015CB857002) and the National Science Foundation of China (grant Nos. 11233005, 11621303, 11522324, 11421303, 11503065, 11673015, 11733004). S.L. acknowledges partial support from a CITa National Fellowship.

ORCID iDs

S. H. Lim @ https://orcid.org/0000-0001-6860-9064
H. J. Mo @ https://orcid.org/0000-0002-9665-5380
Huiyuan Wang @ https://orcid.org/0000-0002-4911-6990

References

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543
Adam, R., Bartalucci, I., Pratt, G. W., et al. 2017, A&A, 598, 115
Arnaud, M., Pratt, G. W., Piffaretti, R., et al. 2010, A&A, 517, 92
Battaglia, N., Bond, J. R., Pfrommer, C., & Sievers, J. L. 2012, ApJ, 758, 74
Cen, R., & Ostriker, J. P. 1999, ApJ, 514, 1
Davé, R., Hernquist, L., Katz, N., & Weinberg, D. H. 1999, ApJ, 511, 521
Davé, R., Cen, R., Ostriker, J. P., et al. 2001, ApJ, 552, 473
David, L. P., Jones, C., Forman, W., Vargas, I. M., & Nulsen, P. 2006, ApJ, 653, 207
De Bernardis, F., Alois, S., Vavagiakis, E., et al. 2017, JCAP, 03, 008
de Graaff, A., Cai, Y.-C., Heymans, C., & Peacock, J. A. 2019, A&A, 624, 48
Dolag, K., Hansen, F. K., Roncarelli, M., & Moscardini, L. 2005, MNRAS, 363, 29
Dolag, K., Komatsu, E., & Sunyaev, R. 2016, MNRAS, 463, 1797
Eckert, D., Iazzuc, M., Shin, H., et al. 2015, Natur, 528, 105

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Effective gas temperature obtained by dividing the electron thermal energy obtained from the tSZE measurement by the total number of electrons obtained from the kSZE. Data points and error bars show the averages of and dispersion among all seven samples, respectively. The black open circles show the results corrected for the bias according to the black curve in the left panel of Figure 1. The shaded band is based on the ellipse shown in Figure 3. The dashed line shows the virial temperature as a function of halo mass.}
\end{figure}
