Astronomy & Astrophysics manuscript no. kszcnn
January 6, 2022

Convolutional Neural Network-reconstructed velocity for kinetic SZ detection

Hideki Tanimura[1], Nabila Aghanim[1], Victor Bonjean[2, 3], and Saleem Zaroubi[4, 5]

[1] Université Paris-Saclay, CNRS, Institut d'Astrophysique Spatiale, Bâtiment 121, 91405 Orsay, France
[2] Instituto de Astrofísica de Canarias, E-38205 Tenerife, Spain
[3] University of La Laguna, E-38206 Tenerife, Spain
[4] ARCO (Astrophysics Research Center), Department of Natural Sciences, The Open University of Israel, 1 University Road, PO Box 808, Ra'anana 4353701, Israel
[5] Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, the Netherlands
e-mail: `hideki.tanimura@ias.u-psud.fr`

2022/04/25 Journal club in Kohno lab
Presenter: Shuhei Inoue (M1)

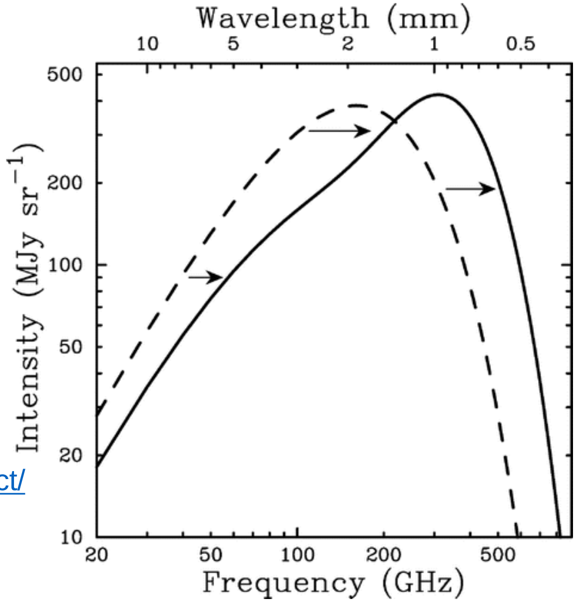
Outline

- **Section 1:** Introduction
- **Section 2:** Dataset
- **Section 3:** Estimation of LOS
- **Section 4:** ‘Stacking analysis’
- **Section 5:** Detection of kSZ signal
- **Section 6:** Fitting to β model and calculation of gas mass ratio
- **Section 7:** Discussions and conclusions

Section 1: Introduction

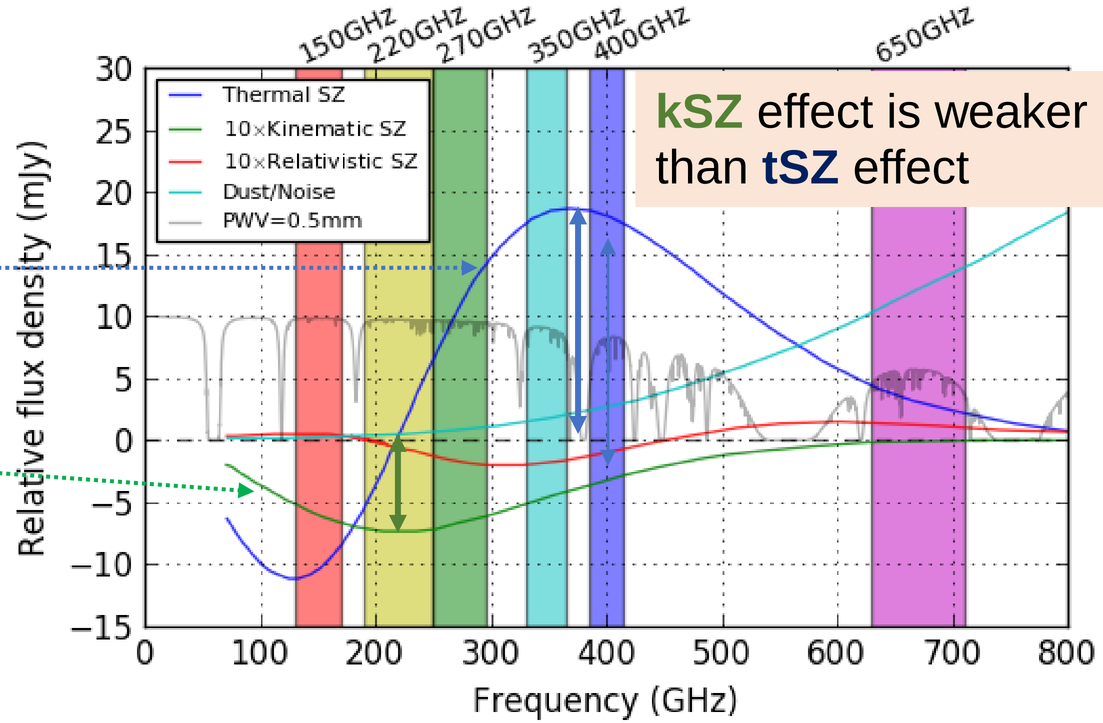
Sunyaev Zel'dovich (SZ) effect

- Galaxy clusters: largest gravitationally bound structures
- Sunyaev Zel'dovich (SZ) effect:
  - Caused by the scattering of the Cosmic Microwave Background (CMB) photons by hot and ionized plasma in the IntraCluster Medium (ICM)
- Classification of SZ effect:
  - Thermal SZ (tSZ) effect:
    - Due to the thermal motion of the cluster
  - **Kinetic SZ (kSZ) effect:**
    - Due to the peculiar motion of the cluster



Carlstorm et al. (2002)

https://lambda.gsfc.nasa.gov/product/suborbit/APEX/bolo.berkeley.edu/apexsz/science.html



kSZ effect is weaker than **tSZ** effect

kSZ effect

- Potential: constrain cosmological and astrophysical model
- Some studies use kSZ signal to measure
  - Optical depth
  - Gas mass fraction
- kSZ effect is sensitive to
  - Virialized gas
  - The gas surrounding halos
  - ➔ suited for studying the gas distribution around galaxy clusters
  - ➔ solve a debate “What amount of diffuse gas is present around the halos ?”

Detection of kSZ effect (1)

- Classical Method
  1. The pairwise method (Hand et al. 2012; Soergel et al. 2016 etc..)
  2. Cross-correlation method (Hill et al. 2016)
     - Redshift information is not needed.
- Most of the measurements: Matched filter or aperture photometry is needed
  - ➢ Matched filter: (optimizes estimated S/N; with known spatial template and SZ spectrum) Melin et al.2006
    - May be biased if the assumed profile is incorrect
      - Due to uncertainty of the gas density profile
  - ➢ Aperture photometry:
    - S/N ratio is depending on aperture

➔ Limitation: S/N = 2~4σ

Detection of kSZ effect (2)

- **Optimized stacking analysis (Tanimura et al. 2021 [T21])**
  - Method
    - Use linearized continuity equation to estimate Line of Sight (LOS) velocity
      - $v(\boldsymbol{x}) = \frac{f(\Omega)aH(a)}{4\pi} \int d^3\boldsymbol{y}\delta(\boldsymbol{y}) \frac{\boldsymbol{y}-\boldsymbol{x}}{|\boldsymbol{y}-\boldsymbol{x}|^3}$
        - $f(\Omega) = \Omega_m^{0.55}$
        - $\delta(\boldsymbol{y})$: overdensity of matter at $\boldsymbol{y}$ position
        - a: scale factor
        - $H(a)$: Habble parameter
    - The peculiar velocities to align the signs of the kSZ signals and stack them. (Section. 4)
      - To avoid cancelation and detect only the kSZ signals.
  - Advantage: without
    1. assumptions of spatial distortion of gas around galaxy clusters
    2. aperture
  - ➔ S/N = 3.5σ (< 4σ ..)

- New approach: Convolutional Neural Network (CNN)
  - Reconstruct velosity field (Wu et al. 2021)
    - From dark matter density field in numerical simulation.
  - They show that
    - CNN approach: $|v_{\mathrm{residual}}| \lesssim 150\ \mathrm{km/s}$
    - Linear perturbation theory: $|v_{\mathrm{residual}}| \sim 300 - 400\ \mathrm{km/s}$
      - v_{residual}: difference between the predict and true velocity
    - CNN approach is more effective in the large density places.
- This paper
  - CNN approach was used extend to real data: Wen-Han-Liu (WHL) galaxy cluster
  - Training using Magneticum (hydrodynamical) simulation (Sec 2.5)
    - ➔ **estimate <u>LOS velosity</u>** (then kSZ radial profile)

Flow: This paper and T21 approach

1. Estimation of **LOS velocity**

- This paper: Convolutional Neural Network (CNN) (Chap 3.2)
- T21: Linearized continuity equation
  - $v(\boldsymbol{x}) = \frac{f(\Omega)aH(a)}{4\pi} \int d^3\boldsymbol{y}\delta(\boldsymbol{y}) \frac{\boldsymbol{y}-\boldsymbol{x}}{|\boldsymbol{y}-\boldsymbol{x}|^3}$

  using Magneticum simulation

2. Estimation of **kSZ**

- Stacking method (Sec. 4)
- Null test (Sec. 4)

3. Estimation of **gas-mass fraction**

- β-model fitting (Sec. 6)

Section2: Datasets

2.1. Galaxy cluster catalog

- Galaxy groups and clusters
  - Total: 158,103 from Sloan Digital Sky Survey (SDSS)
  - Selection: 30,431 clusters
    - ➢ $0.25 < z < 0.55$
    - ➢ $M_{500} > 10^{13.5}\, h^{-1}\, M_{\odot}$
      - ✓ estimated from total luminosity
      - ✓ calibrated by 1191 galaxy clusters (X-ray&tSZ)
    - ➢ No large mask by point sources and galaxies

R_{500}: the enclosed density which is 500 times the *critical* density

M_{500}: enclosed mass within a sphere of radius R_{500}

$h \equiv H_0/(100\mathrm{km/s/Mpc}) = 0.704$

2.2. Galaxy catalog

- Purpose: To estimate LOS velocity of WHL cluster

- Galaxy (after selection ?)
  - North hemisphere … 953,193
  - South hemisphere … 372,542

  From
  1. Baryon Oscillation Spectroscopic Survey (BOSS) LOWZ Galaxy
  2. Constant-MASS (CMASS) Galaxy

- Selection
  - $0.25 < z < 0.55$
    - Number of density is fairly flat

2.3. Planck maps from PR3

- Data: All-sky map from the Planck
- Data release: 2018
- Frequency: 217 GHz (tSZ null)
- Mask:

1. Galactic plane and point source detected by all frequencies
2. Radio and infrared point source

- ➔ exclude ~50% of the sky

2.4. Planck maps from PR4

- Improvements from PR3:
  - ➢ Including foreground polarization
    - ➔ Breaking scanning-induced degeneracies
  - ➢ Correction mismatch at all frequencies
  - ➢ Collection of 8% more data
- Half-ring data for null tests
  - Splitting into two data
  - Produced two maps

2.5. Magneticum simulation

- One of the largest cosmological hydrodynamical simulations
  - ➢ Based on Λ-CDM model (parameter: from Komatsu et al. 2011)
    - ✓ $\Omega m = 0.272$
    - ✓ $\Omega b = 0.046$
    - ✓ H0 = 70.4 $\mathrm{km\ s^{-1}\ Mpc^{-1}}$
  - ➢ Produces simulation boxes with different sizes and resolutions
    - ✓ Largest boxsize = $2688\ h^{-1}$Mpc
    - ✓ Number of particles $= 2 \times 4536^3$
    - ✓ Uses simulation data at z = 0.47 (median redshift of WHL galaxy)

- Purpose: to train and test the machine-learning approach

3. Machine-learning reconstructed LOS velocities

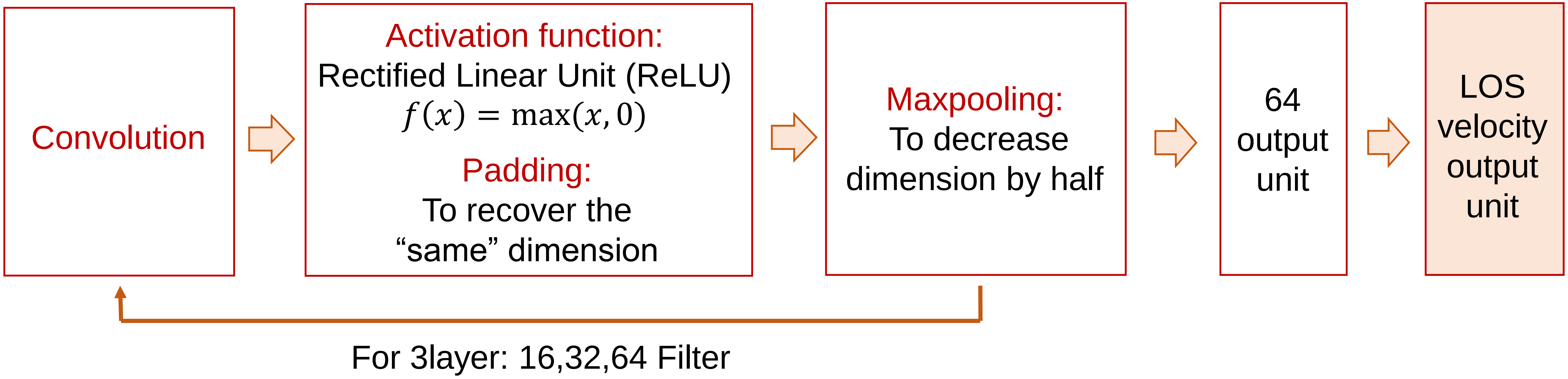
3.1. Training with the Magneticum simulation

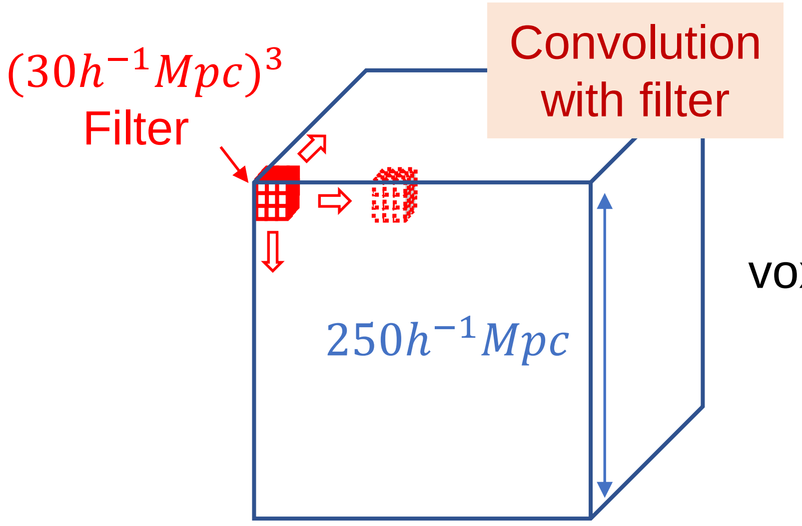
- Magneticum simulation:
  - To learn the correlation between the LOS velocities of the galaxy clusters and their surrounding galaxies
  - Only v_z is learned (z: LOS direction)
- Method:
  1. Selection: to match with WHL galaxy
     - Galaxy cluster: $M_{500} \geq 10^{13.5} h^{-1}\ M_{\odot}$ M_{*}: galacy mass
     - Galaxy: $M_{*} \geq\ 1.0 \times\ 10^{10}\ h^{-1}\ M_{\odot}$
  2. Boundary condition:
     - Clusters: in the center of $(250\ h^{-1}\ \mathrm{Mpc})^3$ cubic box
     - Grid size (= voxel size):
       - T21: $(5h^{-1}\ \mathrm{Mpc})^3\ >\ (3h^{-1}\ \mathrm{Mpc})^3$: expected Redshift-Space Distortion (RSD) length
       - $\text{This Paper:}\ (10h^{-1}\ \mathrm{Mpc})^3$
     - Smoothed by a Gaussian kernel of $2h^{-1}$ Mpc
     - ➔ e.x.) in T21, galaxy overdensities $\delta(\boldsymbol{y})$ are calculated

3.1. Training with the Magneticum simulation

- **Method: (continue)**
  3. Sprit simulation box:
     - $(25^3\ h^{-1}\ \mathrm{Mpc})^3$ into 8 independent regions
       - 7 for the training and valudation ($N_{\mathrm{galaxy}} = 418{,}374$)
       - 1 for the test ($N_{\mathrm{galaxy}} = 59{,}767$)
  4. Training the LOS velosity:
     - Input a series of 25^3 voxels (overdensity fields) into CNN

3.2 Neural network architecture





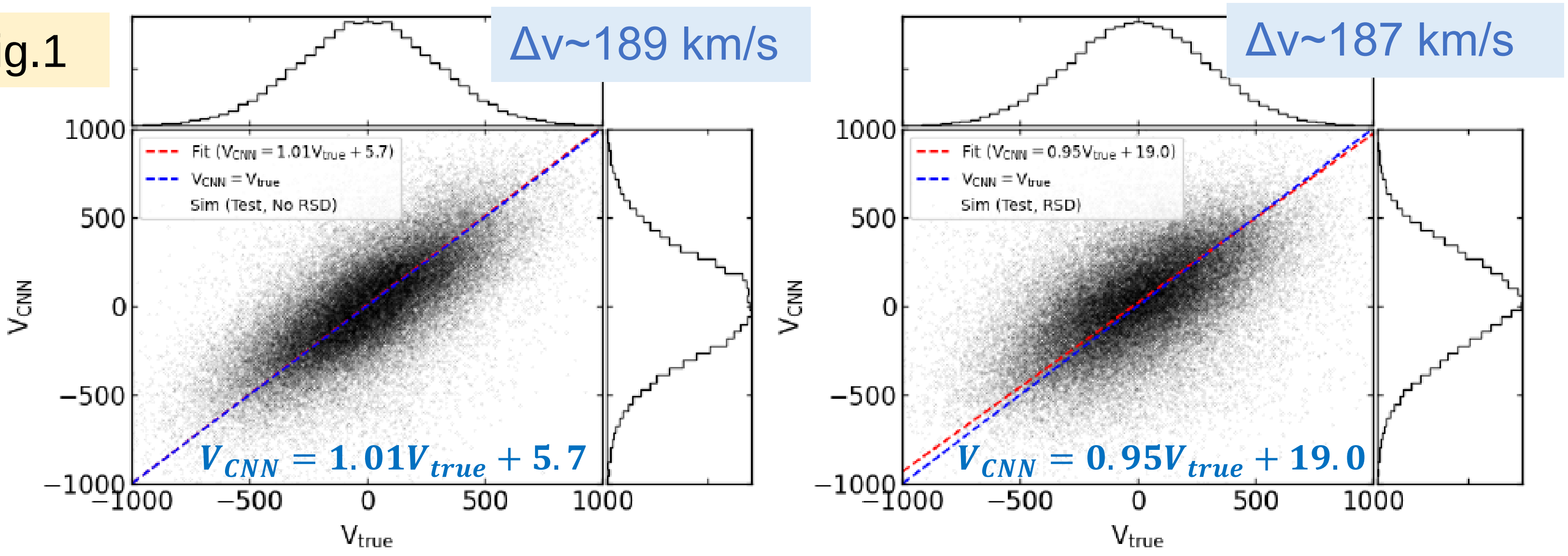
voxel size = $10h^{-1}$ Mpc

3.3. Test with the Magneticum simulation

without RSD | with RSD

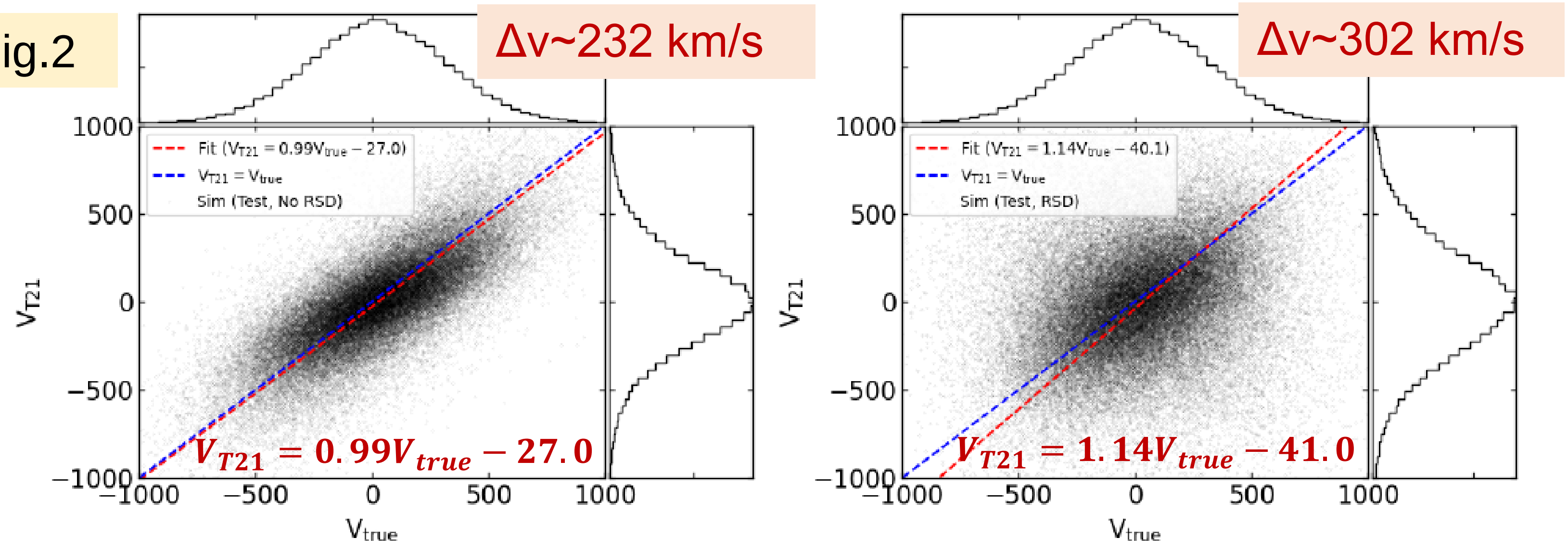
CNN — less sensitive to RSD

Fig.1



T21 — sensitive to RSD

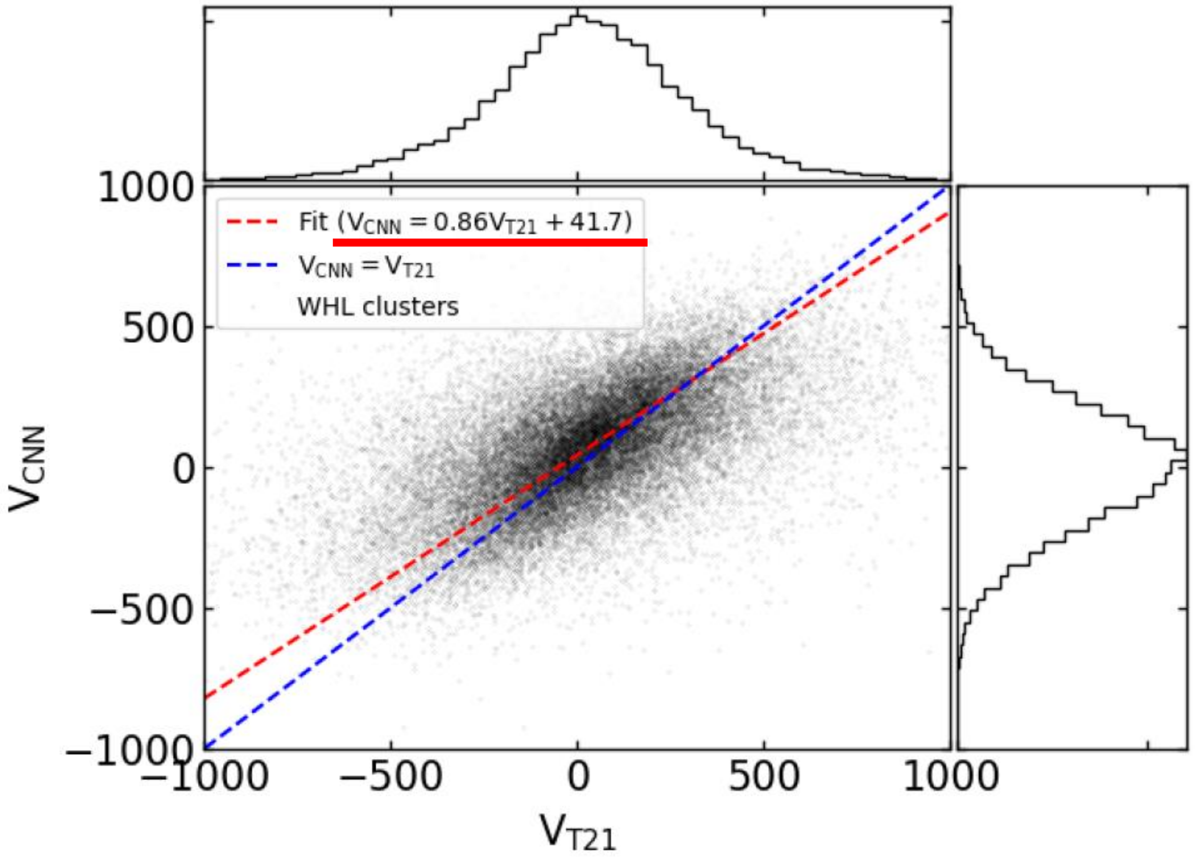
Fig.2



⇨ CNN is less sensitive to RSD than T21

3.4. Apply for real (WHL) galaxy clusters

Fig.3



Method	LOS velocity
CNN&T21 (WHL clusters)	$V_{CNN} = 0.86\,V_{T21} + 41.7$
CNN (Mag. sim. with RSD)	$V_{CNN} = 0.95\,V_{true} + 5.7$
T21 (Mag. sim. with RSD)	$V_{T21} = 1.14\,V_{true} - 41.0$

$$0.86 \sim 0.83 = \frac{0.95}{1.14}$$

from WHL clusters

CNN (Mag. sim. with RSD)

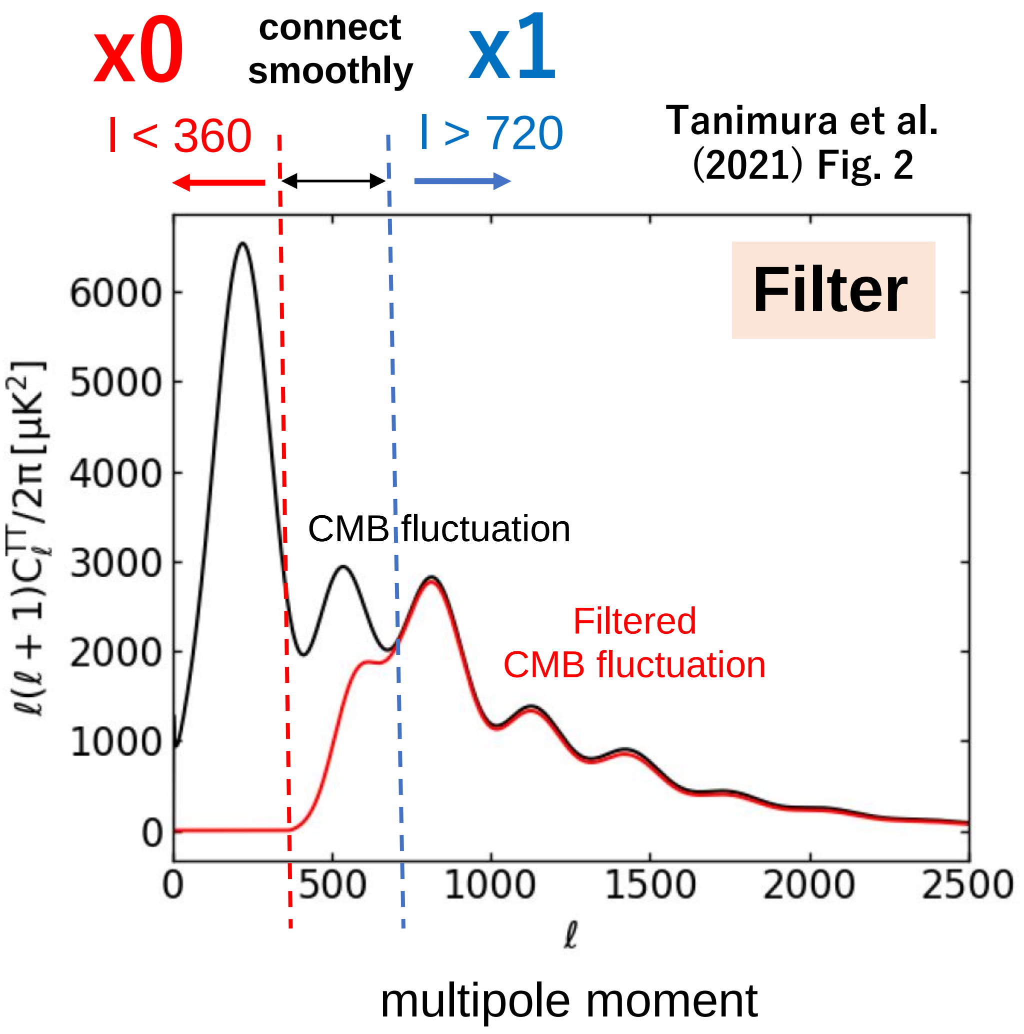
T21 (Mag. sim. with RSD)

The results of Magneticum simulation and WHL galaxy clusters are similar.

4. Stacking analysis

4. Stacking analysis

1. **Filter the Planck PR3 & PR4 map**
   - primordial CMB fluctuation: ~100 μK
   - kSZ signal: ~1 μK
   - ➔ Filter is essential to reduce contamination
   - Filter function
     - l < 360 (> 30 arcmin): 0
     - l > 720 (< 15 arcmin): 1
     - 360 < l < 720: connect smoothly
   - ➔ CMB fluctuation reduced to ~40 μK



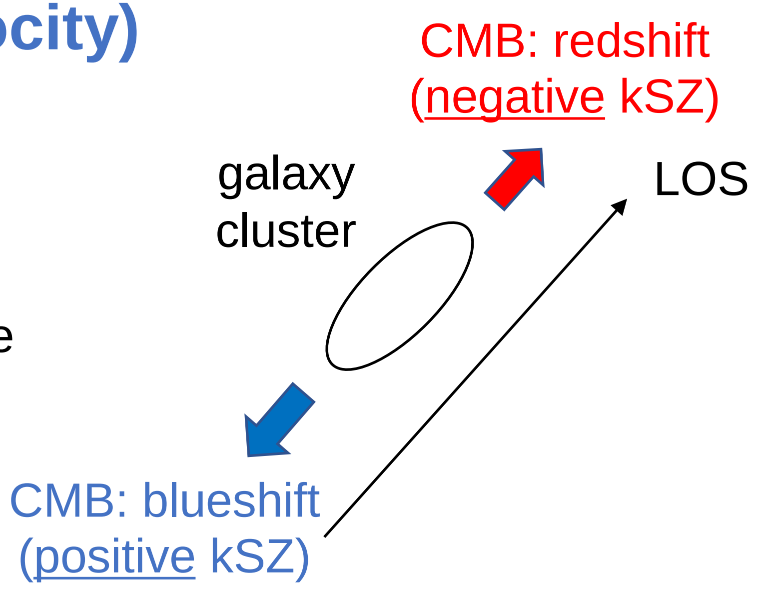
l < 360 ➔ CMB dominant
l > 720 ➔ kSZ signal dominant

2. Placed each cluster in the center of the 2D grid

- $-10 < \frac{\theta}{\theta_{500}} < 10$, (10x10 bins)
  - R_{500}: Radius of galaxy cluster up to 500 times the critical density at the redshift z
  - θ_{500}: angular distance up to 500 times the critical density at the redshift z

3. Stacking of the grid maps (weighted by the LOS velocity)

- $T(R) = \frac{\Sigma_i T_i(R) \times v_{i,\mathrm{LOS}}/\sigma_i^2}{\Sigma_i |v_{i,\mathrm{LOS}}|/\sigma_i^2}$
  - $T_i(R)$: The temperature value of the i-th cluster at the radial distance
  - $v_{i,\mathrm{LOS}}$: LOS velosity of the i-th cluster
  - σ_i: Variance of the temperature values of the i-th cluster



⇨ When $v_{i,\mathrm{LOS}} < 0$ ➔ positive kSZ
When $v_{i,\mathrm{LOS}} > 0$ ➔ negative kSZ

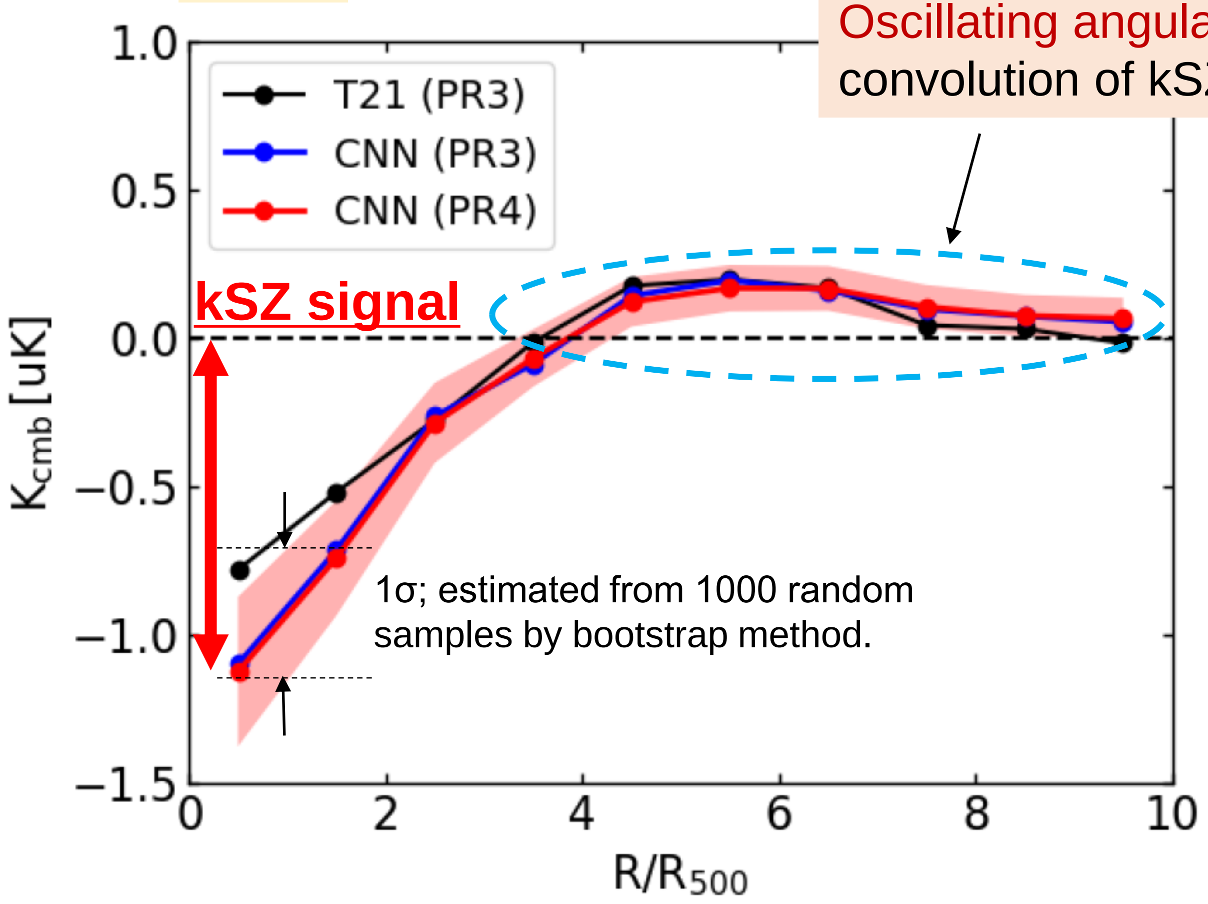
} Both are negative effects for the $T(R)$

Except for the kSZ,
signals are canceled.

5. kSZ detection from WHL galaxy clusters

5.1 kSZ detection with Planck PR3&PR4 maps

Fig.4



Oscillating angular pattern is due to the convolution of kSZ and CMB with filter

Velocity-weighted kSZ radial profile around the 30,431 WHL galaxy clusters

Measuring Scale: $4 \times \theta_{500}$

Method (data)	S/N
T21 (PR3)	3.5σ
CNN (PR3)	4.7σ
CNN (PR4)	**4.9σ**

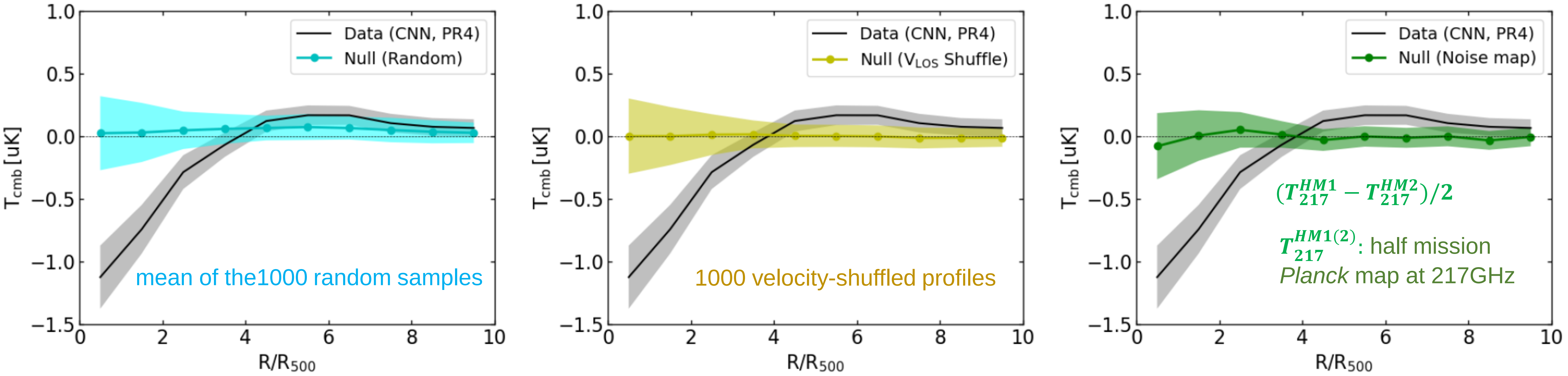
Results

- S/N ratio of CNN is higher than T21
  - ➔ CNN is suitable for kSZ detection. (?)
- S/N ratio of CNN(PR4) is slightly higher than that of CNN(PR3)
  - ➢ Probably because PR4 have less statistical and systematic error than PR3

• Null tests: 5.2 kSZ detection with Planck PR4 maps

1. Place the center of the galaxy clusters at random positions on the sky
   - repeated 1000 times
   - assess the *rms* fluctuations of foreground and background signals
2. Randomly shuffled LOS velocity of the galaxy clusters and they stacked
   - repeated 1000 times
   - evaluate mean and standard deviation
3. $(T_{217}^{HM1} - T_{217}^{HM2})/2$ T_{217}^{HM1}: half mission 1(2) planck map at 217GHz

Fig.5



Average of the three tests consistent with 0 ➔ unbiased measurements

6. Gas mass fraction in WHL galaxy clusters

6. Gas mass fraction in WHL galaxy clusters

- **Goal:**
  - Modeling of kSZ measurement
  - Estimation of gas-mass fraction

- Relative variation of CMB temperature due to the kSZ

From velocity-weighted kSZ radial profile →

$$\frac{\Delta T_{kSZ}}{T_{CMB}} = -\sigma_T \int n_e \left(\frac{\boldsymbol{v}\cdot\boldsymbol{n}}{c}\right) dl \simeq -\tau \left(\frac{\boldsymbol{v}\cdot\boldsymbol{n}}{c}\right)$$

← From average LOS measurement of WHL galaxy clusters (Sec 3.4)

※ uncertainty of the sign was corrected

τ: β-model fitting

$$\tau = \sigma_T \int n_e dl$$

σ_T: Thomson scattering cross section

n_e: Electron number density

$\boldsymbol{v}\cdot\boldsymbol{n}$: LOS velocity

- $\boldsymbol{v}\cdot\boldsymbol{n}$ correlation length ($\sim 80\ h^{-1}$ Mpc)
- n_e correlation length ($\sim 5\ h^{-1}$Mpc)

➔ $\boldsymbol{v}\cdot\boldsymbol{n}$ can be regarded as constant

6. Gas mass fraction in WHL galaxy clusters

β-model:

$$\boldsymbol{n_e(r) = n_{e,0}\left[1 + \left(\frac{r}{r_c}\right)^2\right]^{-\frac{3\beta}{2}}}$$

- $n_{e,0}$: the center electron number density
- $\beta = 0.86$
- $r_c = 0.2 \times R_{500}$

  : core radius of the electron distribution

} From South Pole Telescopes clusters (Plagge et al. 2010)

6. Gas mass fraction in WHL galaxy clusters

Optical depth:

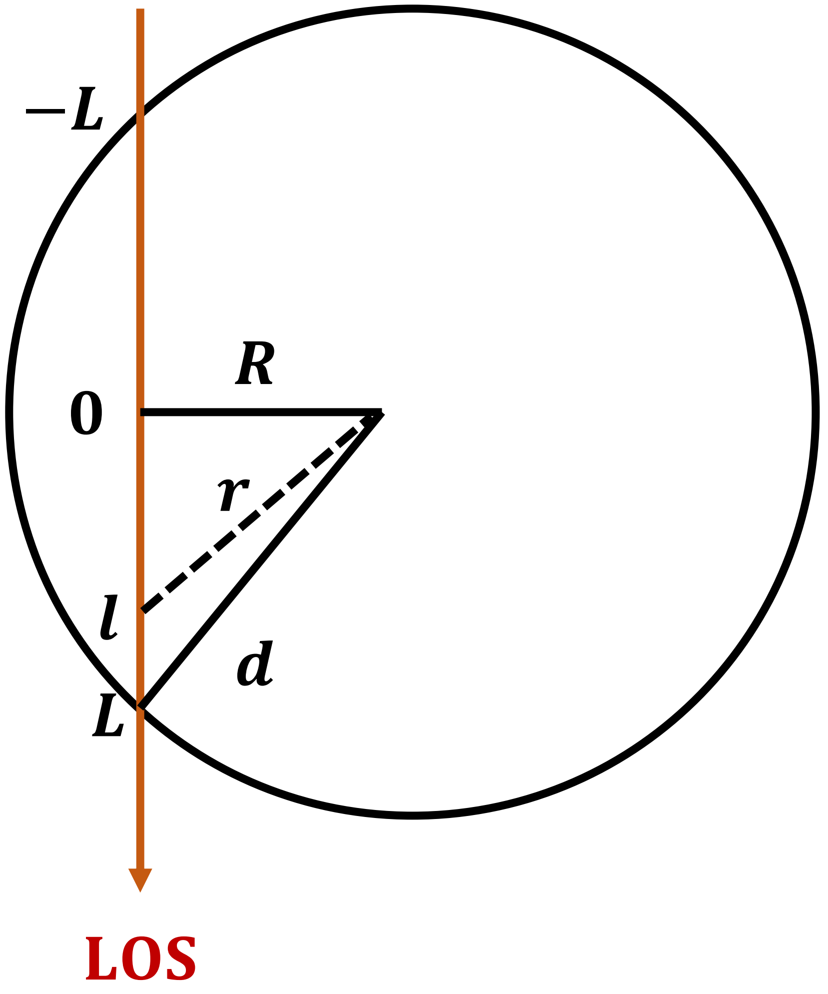
$$\tau(R) = \sigma_T \int \frac{2rn_e(r)}{\sqrt{r^2 - R^2}} dr$$

R: tangential distance from a galaxy cluster

$$l = \sqrt{r^2 - R^2}$$

$$dl = \frac{rdr}{\sqrt{r^2 - R^2}}$$

$$\tau(R) = \sigma_T \int_{-L}^{L} n_e(r) dl = \sigma_T \int_{R}^{d} \frac{2rn_e(r)}{\sqrt{r^2 - R^2}} dr \quad (?)$$



6. Gas mass fraction in WHL galaxy clusters

Optical depth and gas mass:

$$\tau_{e,500} = \int_0^{R_{500}} \sigma_T n_e(T) dV$$

$$\bar{\tau}_{e,500} = (2.0 \pm 0.4) \times 10^{-3}$$

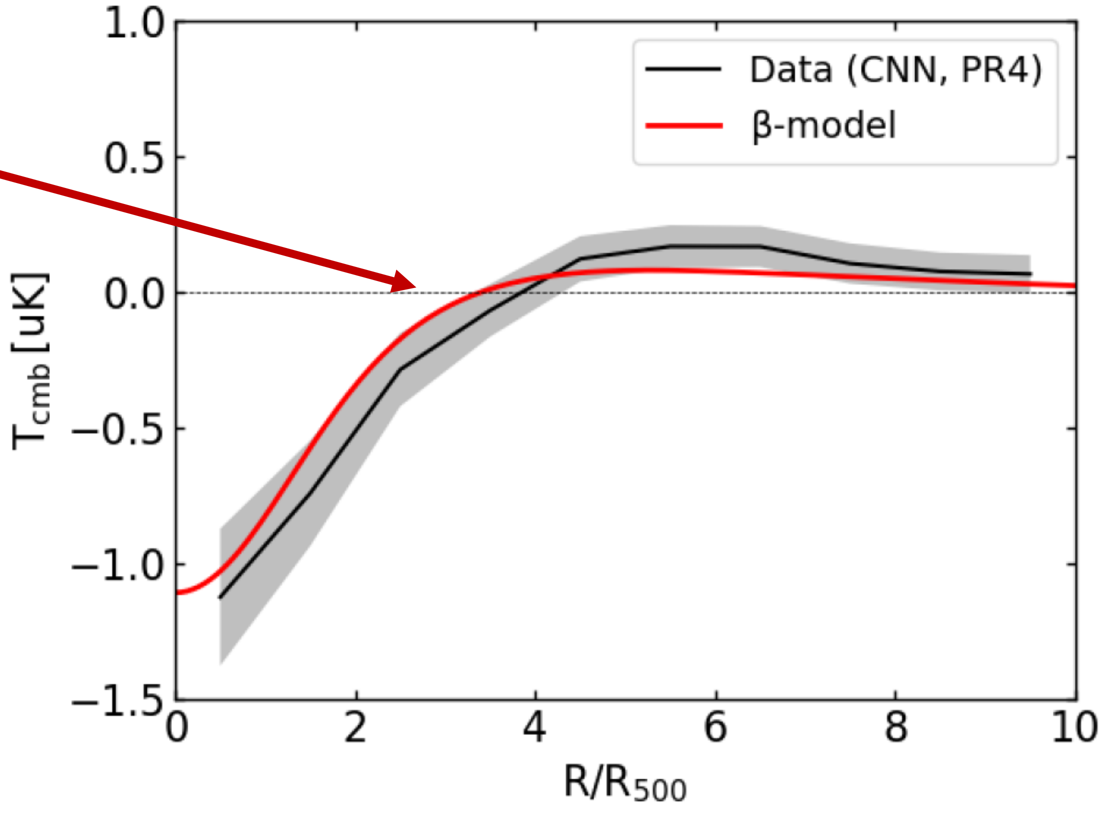
$$M_{gas,500} = \int_0^{R_{500}} n_e(r) \mu_e m_p \, dV$$

$$\bar{M}_{gas,500} \sim 0.9 \times 10^{13} \, h^{-1} M_\odot$$

$$\bar{M}_{500} \sim 1.0 \times 10^{14} \, h^{-1} M_\odot$$

$$f_{gas,500} = \frac{\bar{M}_{gas,500}}{\bar{M}_{500}} = 0.09 \pm 0.02$$

Fig.6 Velocity-weighted kSZ radial profile



7. Discussion and Conclusion

Gas-mass fraction: comparison with other methods

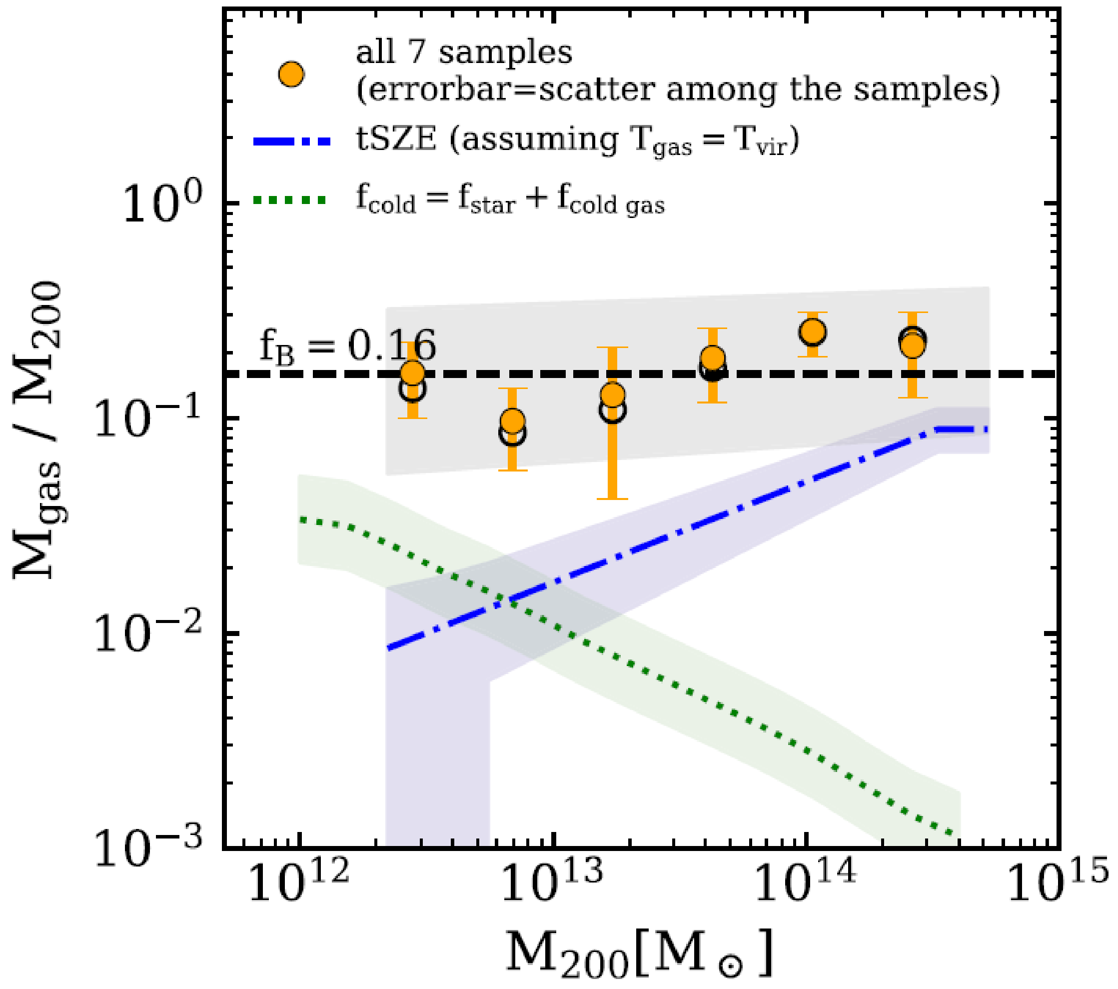
Paper	Object	Galaxy clusters	Method	z	$f_{gas,500}$
This paper	kSZ	WHL clusters	CNN&stacking	0.25-0.55	0.09±0.02
Tanimura et al. 2021	kSZ	WHL clusters	LCE&stacking	0.25-0.55	0.12±0.04
Soergel et al. 2016	kSZ	DES redMaPPer cluster catalog	Pairwise method	< 0.65	0.08±0.02
This paper Tanimura et al. 2021	-	(WHL clusters)	Magneticum Simulation	~0.47	~ 0.13
Nelson et al. 2019	-	?	IllustrisTNG (TNG300-1) simulation	?	~0.13
Gonzalez et al. 2013	X-rays	Abell catalog	Spectral Fitting Surface Brightness Profile Fitting	~0.1	~0.1

- $f_{gas,500}$ of this paper and Soergel et al. (2016) are lower than from the hydrodynamical simulations
  - but, consistent with ~2σ
- kSZ and X-ray measurements are consistent

Unknown: Discrepancy fraction of gas mass

- Lim et al. (2020) claims that $f_{gas} \simeq f_{baryon} = 0.16\ (?)$
  - ➔ not consistent with this paper, Soergel et al. (2016), and Gonzalez et al. (2013)
  - ✓ kSZ measurement
  - ✓ z ~ 0.12
- Possible Reason:
  - Difference between z ?
    - Lim et al (2020): z ~ 0.12
    - This paper and Soergel (2016): z ~ 0.5
    - But, Gonzalez et al. (2013): $z \sim 0.1 \simeq 0.12$
- Schaan et al. (2021): 6.5σ
  - ✓ Atacama Cosmology Telescope (ACT) DR5
  - ✓ z~0.55 (CMASS), z~0.31 (LOWZ)
  - ✓ Stacking method

Lim et al. (2020) Figure. 4



Future works: The CNN approach will be applied to ACT data.

Summary:

- The CNN new approach to estimate LOS velocity (then kSZE) was used
- The CNN approach was less sensitive to RSD than the T21 approach
- Average of the three Null tests using Planck PR4 is consistent with 0.
  - ➔ It was found to be unbiased measurement.
- S/N ratio of CNN (4.7σ; PR3) is higher than that of T21 (3.5σ; PR3)
- S/N ratio of CNN (using PR4) is 4.9σ
- $\tau_{e,500} = (2.0 \pm 0.4) \times 10^{-3}, f_{gas} = 0.09 \pm 0.02$ were estimated
  - lower than from the hydrodynamical simulations (f_{gas} ~0.13), but ~2σ
  - Consistent with X-ray measurement (f_{gas} ~0.1) by Gonzalez et al. (2013)
- The reason of the Discrepancy fraction of gas mass claimed by Lim et al. (2020) was still unknown.

My Questions

- How accurate is the β-model fitting ?
  - No matter how much we improve the accuracy of kSZ signal detection, we cannot get accurate parameters if the β-model fitting is uncertain.
- How about cross-correlate method instead of stacking method ?
  - Stacking method (T21) S/N=3.5σ may be not so good compared to Hill et al. 2016 S/N= 4σ
  - But of course, simple comparisons cannot be made because of the difference in the observations.
- What about the possibility of the CNN approach Improvement ?
  - e.g.) Fine tuning of the boundary conditions