Synchrotron emission from virial shocks around stacked OVRO-LWA galaxy clusters

Kuan-Chou Hou,1,* Gregg Hallinan,2 and Uri Keshet1†

1 Physics Department, Ben-Gurion University of the Negev, POB 653, Be’er-Sheva 84105, Israel
2 Department of Astronomy, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125

15 March 2023

ABSTRACT

Galaxy clusters accrete mass through large scale, strong, structure-formation shocks. Such a virial shock is thought to deposit fractions $\xi_e$ and $\xi_B$ of the thermal energy in cosmic-ray electrons (CREs) and magnetic fields, respectively, thus generating a leptonic virial ring. However, the expected synchrotron signal was not convincingly established until now. We stack low-frequency radio data from the OVRO-LWA around the 44 most massive, high latitude, extended MCXC clusters, enhancing the ring sensitivity by rescaling clusters to their characteristic, $R_{500}$ radii. Both high (73 MHz) and co-added low (36–68 MHz) frequency channels separately indicate a significant (4–5$\sigma$) excess peaked at $(2.4–2.6)R_{500}$, coincident with a previously stacked Fermi $\gamma$-ray signal interpreted as inverse-Compton emission from virial-shock CREs. The stacked radio signal is well fit (TS-test: 4–6$\sigma$ at high frequency, 4–8$\sigma$ at low frequencies, and 8–10$\sigma$ joint) by virial-shock synchrotron emission from the more massive clusters, with $\dot{m}\xi_e\xi_B \approx (1–4) \times 10^{-4}$, where $\dot{m} \equiv M/(MH)$ is the dimensionless accretion rate for a cluster of mass $M$ and a Hubble constant $H$. The inferred CRE spectral index is flat, $p \approx 2.0 \pm 0.2$, consistent with acceleration in a strong shock. Assuming equipartition or using $\dot{m}\xi_e \approx 0.6\%$ inferred from the Fermi signal yields $\xi_B \approx (2–9)\%$, corresponding to $B \approx (0.1–0.3) \mu G$ magnetic fields downstream of typical virial shocks. Preliminary evidence suggests non-spherical shocks, with factor 2–3 elongations.

Key words: galaxies: clusters: general - galaxies: clusters: intracluster medium - intergalactic medium - magnetic fields - radio continuum: general

1 INTRODUCTION

Galaxy clusters and filaments are the largest gravitational bound objects in the Universe, with clusters located at the nodes of the cosmic web of large-scale structure (LSS). A cluster of mass $M$ is thought to grow by accreting surrounding matter at a rate $\dot{M} = \dot{m}MH$, where $H$ is the Hubble parameter and $\dot{m}$ is a dimensionless parameter of order unity. As the accreted gas is violently decelerated, strong, collisionless, so-called virial or structure formation shocks form near the virial radius of the cluster. These virial shocks define the edges of clusters, so observing them can provide useful information about the formation of LSS, in particular the local accretion rate. As the accreted gas is pristine and weakly magnetized, virial shocks also provide a useful laboratory for studying collisionless shock physics. A virial shock is distinct from the weak shocks found in its heated downstream, known as the intracluster medium (ICM); such weak shocks, usually observed at $r \lesssim 1$ Mpc radii following a merger, are more difficult to model as they propagate into a pre-heated, magnetised plasma, already enriched by metals and cosmic-rays.

Virial shocks are thought to accelerate charged particles to highly relativistic, $\gtrsim 10$ TeV energies, in resemblance of strong supernova remnant shocks that share similar, $\sim 10^8$ km s$^{-1}$ velocities. The cosmic ray (CR) electrons (CREs) and ions (CRIs) accelerated by such strong shocks are thought to develop a nearly flat, $dN/dE \propto E^{-p}$, energy spectrum with $p \approx 2$ (equal energy per logarithmic CR energy bin), as expected in diffusive shock acceleration theory (for a review, see Blandford & Eichler 1987) when scattering is not too anisotropic (Keshet et al. 2020a). A fraction $\xi_e$ of the downstream thermal energy, deposited in CREs, is subsequently radiated away, forming a distinctive non-thermal signature, which should surface at the extreme ends of the electromagnetic spectrum (Loeb & Waxman 2000; Totani & Kitayama 2000; Waxman & Loeb 2000; Keshet et al. 2003; Miniati 2002; Keshet et al. 2004).

The high energy CREs accelerated by the virial shock cool rapidly, primarily by inverse-Compton scattering off cosmic microwave background (CMB) photons, resulting in a thin shell of energetic CREs. These CREs produce a nonthermal, detectable leptonic ring around the cluster, as predicted analytically (Loeb & Waxman 2000; Waxman & Loeb 2000; Totani & Kitayama 2000) and calibrated using cosmological...
Simulations (Keshet et al. 2003; Miniati 2002; Keshet et al. 2004). The rings are expected to be fairly spherical, but somewhat elongated toward the main filament feeding the cluster (Keshet et al. 2003).

Evidence for the virial inverse-Compton signal first surfaced in the Coma cluster: an ideal virial-shock target, thanks to its high mass, proximity (redshift $z \approx 0.023$), and location in a low-foreground region near the north Galactic pole. Consistent signals were reported in $\sim 220$ GeV data from VERITAS (Keshet et al. 2017), $\sim$ GeV data from Fermi-LAT, and $\sim 0.1$ keV data from ROSAT (Keshet & Reiss 2018). The inferred virial ring is elongated toward the large-scale filament between Coma and the nearby cluster A1367, with a semi-minor axis $\sim 2.1R_{500}$ coinciding with the virial radius, and is thin; searches for a thick or spherical LAT signal did not recover the signal. Here, $R_500$ is a characteristic cluster radius, where subscript $\delta$ designates an enclosed mass density $\delta$ times above the critical mass density of the universe. The three aforementioned signals are all consistent with a normalisation $\dot{m}\xi \approx 0.3\%$ (up to an uncertainty factor $\sim 3$) and with a $2.0 \lesssim p \lesssim 2.2$ spectrum (Keshet & Reiss 2018).

Galaxy cluster signals can often be amplified by stacking data around multiple clusters, utilising the self-similarity of the latter. By co-adding Fermi-LAT data around 112 massive, high latitude, extended clusters, rescaled by their $R_{500}$ radii to a dimensionless, $\tau \equiv r/R_{500}$ scale, binning the outcome radially, and utilising the anticipated flat spectrum, a high significance ($> 5\sigma$) stacked $\gamma$-ray ring was identified (Reiss et al. 2017; Reiss & Keshet 2018, henceforth RK18). The stacked signal is best fit by a normalised shock (subscript $s$) radius $\tau_s \equiv r_s/R_{500} = 2.3 \pm 0.1$, a CRE injection rate $\dot{m}\xi \approx (0.6 \pm 0.1)\%$, and a spectral index $p = 2.1 \pm 0.2$, all consistent with predictions and with the Coma signal. While stacking increases the sensitivity, it carries an elevated systematic uncertainty (here admitting best-fitting peak radii in the range $2.2 \lesssim \tau \lesssim 2.5$, and an additional uncertainty factor $\sim 2$ in $\dot{m}\xi$) and washes away non-spherical components and variations among clusters, due to the co-addition and radial binning. Moreover, the projected shock signal of an individual cluster should not be precisely circular, so stacking effectively biases $\dot{m}$ towards lower values.

A direct tracer of virial shocks, independent of any particle acceleration, is their anticipated (Kocsis et al. 2005) imprint on the thermal Sunyaev-Zel’dovich (SZ; Sunyaev & Zeldovich 1972) signal. A spatial correlation was identified between the $\gamma$-ray ring in Coma and an outdropward in the y-parameter inferred from WMAP data, supporting the association of both signals with the virial shock (Keshet et al. 2017). A highly significant, $8.6\sigma$ drop in y-parameter was later identified in the cluster A2319, which has the highest signal-to-noise detection in the Planck SZ catalogue, indicating a strong shock around $\tau \approx 3$ with a Mach number $Y \gtrsim 3.25$ (at the 95% confidence level; Hurier et al. 2019; Keshet et al. 2020b). A joint search for SZ and Fermi $\gamma$-ray signals in the massive clusters Coma, A2319, and A2142 showed a significant drop in $y$ coincident with a weak $\gamma$-ray excess in all three clusters, consistent with $0.2\% \lesssim \dot{m}\xi \lesssim 0.7\%$ (Keshet et al. 2020b). Combining the SZ signal with galaxy counts gives a measure of the accretion rate (Hurier et al. 2019), allowing independent measurements of $\xi \sim 0.5\%$ and $\dot{m} \simeq 1.1$ in A2319 (Keshet et al. 2020b). More recently, stacking the Planck y-parameter around 10 galaxy groups (Pratt et al. 2021) indicated a significant drop around $2.0 \lesssim \tau \lesssim 2.6$, and stacking the $y$-parameter among 500 clusters using South Pole Telescope data (Anbajagane et al. 2022) indicated a projected drop starting at $\tau \approx 2$, both consistent with a virial shock at the stacked $\gamma$-ray radius of RK18. Interestingly, stacking the 500 clusters also indicated a second drop starting at $\tau \approx 6$; if verified, this would suggest elongated virial shocks with a $\sim 2.5$ ellipticity ratio, as indicated by the virial ring found in Coma by Keshet et al. (2017); Keshet & Reiss (2018).

In addition to the inverse-Compton signal, the CREs accelerated by the virial shock should also emit a synchrotron radio signal, as they gyrate in the post-shock magnetic fields (Waxman & Loeb 2000; Keshet et al. 2004). Assuming that the shock deposits a fraction $\xi_B \approx 1\%$ of the thermal energy in magnetic fields (as inferred from cluster halo observations; see Waxman & Loeb 2000; Keshet et al. 2004, and references therein), the cumulative signal from all clusters should constitute a considerable fraction ($\approx 30\%$ for $\xi_B = 10^{-4}$) of the diffuse extragalactic radio background below 500 MHz, where it dominates fluctuations on $1' \lesssim \theta \lesssim 1^2$ scales (Keshet et al. 2004). The signal, similar in morphology but somewhat more patchy than its inverse-Compton counterpart, has proven difficult to detect, although a correlation found between the WMAP synchrotron signal and the VERITAS $\gamma$-ray ring in Coma corresponds to $\xi \approx 1\%$ (Keshet et al. 2017).

The flat, $\alpha \approx p/2 \approx 1$ radio spectral index of the anticipated virial shock signal from cooled CREs is somewhat softer than that of the Galactic foreground, giving low-frequency interferometers an advantage in searching for this elusive signal (Keshet et al. 2004). The Owens Valley Radio Observatory Long Wavelength Array (OVRO-LWA; hereafter LWA) is particularly suited for such a search, providing clean sky maps of the northern sky in the (36-73) MHz frequency range. Searching the low-frequency sky for the virial shock signature is complicated by the weakness of the signal, an abundance of unrelated point and diffuse sources, including radio halos and relics associated with the galaxy clusters themselves (Brunetti et al. 2008; Keshet 2010; Bonafede et al. 2014, 2018; Cassano et al. 2019; Hoang et al. 2019; Ozinga et al. 2020), and instrumental difficulties associated in particular with beam sidelobes of sources in the centres of clusters. Such studies therefore require careful modelling and extensive sensitivity tests.

We select the 44 most massive, nearby, extended clusters in the northern sky from the Meta-catalogue of X-ray Clusters (MCXC; Piffaretti et al. 2011), after avoiding regions of strong Galactic contamination or a suboptimal point spread function (PSF). LWA data are excised around these clusters, cleaned from point sources and from smooth background and foreground signals, rescaled to the $R_{500}$ extent of each cluster, co-added over all clusters, and radially binned, to test if a residual virial-shock signal can be identified. A robust (in particular, independent of PSF residuals) and significantly-significant signal, found around the virial radius and coincident with the previously stacked $\gamma$-ray signal, is then modelled as the cumulative synchrotron emission from cluster virial shocks.

The paper is organised as follows. We outline the LWA data and the cluster sample in §2. The stacking analysis, including background modelling and Monte-Carlo simulations of control cluster samples, is described in §3. In §4, we outline the results of the stacking, in particular a strong central signal
and a significant virial excess. Simple models for the central emission are presented and fit to the data in §5. Models for virial-shock emission are fit to the peripheral signal, alone and jointly with the central signal, in §6. Finally, the results are summarised and discussed in §7.

The Supplementary Material provides sky maps (§A), technical details concerning different aspects of the analysis, including cluster co-addition (§B), background and foreground removal (§C), control sample analyses (§D), noise and correlation corrections (§E), modelling the central emission from clusters (§F), additional tests constraining residual sidelobe artefacts (§G), sensitivity and consistency tests (§H), the theory of synchrotron emission from virial shocks (§I), cross contamination by central and virial modelling residuals (§J), methods of inferring the peripheral magnetic fields (§K), the list of clusters in our sample (§L), and extended 2D significance images (§M).

We adopt a flat $\Lambda$CDM cosmological model with a Hubble constant $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$, a mass fraction $\Omega_m = 0.3$, and a cosmic baryon fraction $f_b = 0.17$. Assuming a 76% hydrogen mass fraction, we use a mean particle mass $\bar{m} \simeq 0.59 m_p$, where $m_p$ is the proton mass. Error bars designate 68% containment projected for a single parameter.

2 DATA AND DATA PREPARATION

2.1 LWA sky maps

The LWA is a 288-antenna interferometer with a maximum baseline of $\sim 1.5$ km (Anderson et al. 2019; Eastwood et al. 2019). We use data produced from 28 hours of observation during 2017–2018. The data, in the form of eight low-frequency maps of the sky north of declination $\delta = -30^\circ$, were constructed with Tikhonov-regularised $m$-mode analysis imaging (Eastwood et al. 2018). The maps were taken at frequencies 36.53, 41.76, 46.99, 52.46, 57.46, 62.69, 67.9 2 GHz, with a 24 kHz bandwidth, a 15′–20′ FWHM angular resolution, and $\sim 800$ mJy/beam thermal noise. More precise parameters, and the thermal noise level $T_0(\nu)$ in each frequency $\nu$, are discussed in Eastwood et al. (2018) and summarised in Table 1.

For the high-frequency, $\nu_h = 73.152$ MHz channel, the sky map was prepared with a substantially better, $\sim 11$′ FWHM resolution. Therefore, we separate the study of data in this channel (henceforth referred to as the high-frequency channel) from the analysis of the other, lower-frequency data sets (referred to as the low-frequency channels).

Point sources were removed from the maps using a version of the CLEAN algorithm, adapted to deconvolve the PSF for $m$-mode analysis imaging by drift-scanning telescopes like the OVRO-LWA (see algorithm details in Eastwood et al. 2018). To better understand the point-source contribution and the performance of CLEAN, we utilise two sets of maps in each channel, denoted ‘cleared’ and ‘restored’, provided by the OVRO-LWA collaboration. The cleared maps are the usual output of CLEAN, after the flux associated with each source was removed by convolving a point source with the PSF. In the restored maps, the flux of each source thus removed was subsequently restored using a simpler, Gaussian beam. Figure 1 shows the resulting sky map of the cleared data in the high-frequency channel.

Our nominal analyses are based on the cleared data. To further avoid contamination from bright source residuals, we mask all HEALPix pixels within a radius 0.3 of $F_{1.4\,\text{GHz}} > 1$ Jy sources in the NRAO VLA Sky Survey (NVSS; Condon et al. 1998). Our results are not sensitive to small variations in the masking radius or in the flux cut, as illustrated in §H.

The PSF of the LWA varies with declination. The PSF sidelobes become increasingly less circularly-symmetric toward lower declination, in particular for $\delta < 20^\circ$ (Eastwood et al. 2018). This gradual distortion of the PSF renders sidelobe artefacts increasingly difficult to test and correct. Therefore, we impose a cut on the cluster sample, and include only clusters with $\delta > 20^\circ$ in our sample, thus selecting for a well-behaved, approximately circular PSF. The basic information and properties of the maps are summarised in Table 1.

The central regions of galaxy clusters are often radio-bright, due to both point sources and diffuse emission. Searching for an annular signal around such a cluster is challenged by the possible presence of sidelobes that were not fully accounted for. This difficulty is addressed and tested throughout the text, in particular in §4.1.4, §4.2.3, §F1, and §G, as summarised in §7. Another difficulty is that if the virial ring is sufficiently clumpy, it may in part be misidentified by CLEAN as a collection of point sources and thus removed; we show in §4.1.2 that this is not the case.

The sky maps are discretised using a HEALPix scheme (Górski et al. 2005) of order 11 (nside = 2048), involving $3 \times 2^{24} \simeq 5 \times 10^7$ pixels on the sky, so each pixel spans a solid angle

$$\delta \Omega \simeq 2.5 \times 10^{-7} \text{sr}.$$  

(1)

The mean, $\sim 0.3$ separation between pixels is sufficiently small with respect to the angular resolution of the LWA to avoid discretisation errors. Our modelling utilises tabulated PSF maps, provided for the full declination range by the OVRO-LWA collaboration.

Figure 1. Brightness temperature map of the cleared LWA data in the high-frequency channel using a Mollweide projection, in the full (declination $\delta > -30^\circ$) field of view: White circles of radius 5′500 are shown around the 44 clusters in our sample. The cuts discussed in §2.2 are shown as curves: around the Galactic plane (dotted magenta lines), Loop-I (dashed blue circle), Fermi bubbles (dash-dotted yellow curve), and $\delta = 20^\circ$ (long-dashed red curve). A larger version and a map based on restored data may be found in §A.
2.2 Cluster sample

We select groups and clusters of galaxies (henceforth, clusters) for radio stacking from the MCXC catalogue. This catalogue, based on the ROSAT all sky-survey, was homogenised to an overdensity of 500, thus providing the characteristic length and mass scales, required for our analysis, for 1743 clusters. In addition to the position of each cluster on the sky and the cluster mass \( M_{500} \) enclosed within \( R_{500} \), the catalogue specifies the redshift \( z \) and radius \( R_{500} \) of each cluster, so the corresponding angular radius \( \theta_{500} \) can be computed.

To construct an optimal sample of clusters for the virial shock search, we select only clusters that satisfy all of the following criteria:

(i) Massive clusters: the anticipated synchrotron signal arises from massive clusters, so we select clusters with a mass \( M_{500} > 10^{14} M_\odot \). Note that the mass dependence of synchrotron emission is stronger than its inverse-Compton counterpart, due to expected scaling of the magnetic field.

(ii) Resolvable \( R_{500} \) core: an angular radius \( \theta_{500} > 0.2 \), chosen such that \( R_{500} \) is marginally resolved at the \( \sim 12' \) PSF scale of our nominal, high frequency channel.

(iii) Avoiding excessively extended, bright clusters of \( \theta_{500} > 0.4 \), which are too bright and extended for our analysis; in particular, such clusters complicate the background removal procedure.

(iv) Avoiding contamination from the Galactic plane by excluding clusters located at latitudes \( |b| < 20^\circ \).

(v) Avoiding possible contamination from Loop-I, by excluding a circle on the sky centred on Galactic coordinates \( \{ l, b \} = (337.2, 24.6) \) with a \( 5^\circ \) radius. This region is based on LWA data, and is largely consistent with previous findings (e.g. Haslam et al. 1971; Dickinson 2018).

(vi) Avoiding possible contamination from the Fermi bubbles, by excluding the regions defined as \( |l| < 55^\circ \) and either \( l < 45^\circ \) or \( l > 315^\circ \), (approximating the bubbles structure found in Keshet & Gurwich 2017).

(vii) Optimising for PSF: as in the cut on the data, we select only clusters with \( \delta > 20^\circ \) to avoid the PSF deterioration at lower declination.

Out of the 1743 clusters in the MCXC catalogue, the above cuts leave us with a sample consisting of \( N_c = 44 \) high-quality clusters. These are nearby clusters, found at redshifts \( z < 0.07 \). The clusters in our sample and their basic properties are listed in the supplementary material §L. Their median values are \( z \simeq 0.03, \theta_{500} \simeq 0.24, \) and \( M_{500} \simeq 0.56 \times 10^{14} M_\odot \). Variations in the above cuts are discussed in §H; our results are shown to be robust to reasonable changes in all cluster selection criteria.

3 STACKING METHOD

Along with the stacking of multiple rescaled maps around clusters, detecting the virial shock signal requires careful removal of backgrounds and foregrounds (henceforth background, for brevity), angular data-folding or radial binning, and Monte-Carlo simulations to calibrate the noise and correlation parameters and validate the relevant sky statistics.

The procedures of scaling, co-adding, and binning the data are outlined in §3.1. The stacked quantities and their computation are detailed in §3.2. Different background removal methods are discussed in §3.3. Finally, we describe the Monte-Carlo simulations of control cluster samples in §3.4. These control samples, consisting of random position in the sky analysed as if they contain clusters analogous to the real sample, are used for calibration purposes throughout §3.1–§3.3, where we refer to them as 'control clusters'; in later sections, where we inject mock signals into these random sky locations for modelling purposes, we refer to them as 'mock clusters'.

### Table 1. LWA stacked-data properties.

| \( \nu (\Delta \nu) \) [MHz] | \( T_0 \) [K] | FWHM (\( \delta = 45^\circ \)) | Restored FWHM | Cleared \( \eta_c \) | \( \eta \) | \( \psi \) | Restored \( \eta_c \) | \( \eta \) | \( \psi \) |
|-----------------------------|--------------|-------------------------------|---------------|-----------------|------|-----|----------------|------|-----|
| (1)                         | (2)          | (3)                          | (4)           | (5)             | (6)  | (7) | (8)            | (9)  | (10) |
| 36.528 (0.024)              | 595          | 20' 2.2 \times 16'.9         | 16'.1         | 1.24 (1.20)     | 1.17 (1.14) | 0.70 (0.71) | 1.91 (1.86) | 1.90 (1.83) | 0.85 (0.85) |
| 41.760 (0.024)              | 541          | 18'.5 \times 16'.0           | 14'.5         | 0.99 (0.95)     | 0.93 (0.90) | —              | 1.54 (1.49) | 1.51 (1.45) | — |
| 46.992 (0.024)              | 417          | 17'.4 \times 15'.2           | 14'.0         | 1.06 (1.03)     | 1.02 (0.99) | —              | 1.63 (1.57) | 1.61 (1.55) | — |
| 52.224 (0.024)              | 418          | 16'.2 \times 15'.9           | 13'.1         | 0.89 (0.86)     | 0.87 (0.84) | —              | 1.33 (1.28) | 1.30 (1.25) | — |
| 57.456 (0.024)              | 354          | 15'.9 \times 15'.0           | 12'.9         | 0.84 (0.80)     | 0.81 (0.78) | —              | 1.18 (1.13) | 1.17 (1.13) | — |
| 62.688 (0.024)              | 309          | 15'.8 \times 14'.9           | 12'.7         | 0.77 (0.74)     | 0.75 (0.72) | —              | 1.05 (1.02) | 1.05 (1.01) | — |
| 67.920 (0.024)              | 281          | 15'.9 \times 14'.7           | 12'.6         | 0.71 (0.68)     | 0.69 (0.67) | —              | 0.94 (0.91) | 0.93 (0.90) | — |
| 73.152 (0.024)              | 154†         | 12'.2 \times 10'.9†          | 9'.3          | 0.97 (0.94)     | 0.94 (0.91) | —              | 1.26 (1.22) | 1.24 (1.20) | — |

Columns: (1) Central frequency (and bandwidth) (MHz) of each channel; (2) Noise temperature (in K, from Eastwood et al. 2018); (3) FWHM major axis \( \times \) minor axis of the synthesized beam at a typical declination \( \delta \sim 45^\circ \); (4) Effective FWHM of the restored data in the \( 20^\circ < \delta < 60^\circ \) range (see §5.1); (5) Correction factor \( \eta_c \equiv \eta (2.25 \leq \tau < 2.5) \) for the product \( T_0 \cdot \text{FWHM} \) in the virial shock region, inferred from control samples for the beam- (cluster-)stacked cleared data; (6) Correction factor \( \eta \equiv \eta (0 < \tau < 10) \) uniformly averaged over the full control range for the cleared data; (7) Correlation parameter of low-frequency channels uniformly averaged over the range \( 0 < \tau < 10 \), for the cleared data; (8) \( \eta_c \) for the restored data; (9) \( \eta \) for the restored data; (10) \( \psi \) for the restored data.

† — In the high-frequency channel we use a dedicated high-resolution data set, provided by the OVRO-LWA collaboration, rather than the nominal-resolution \( 16'.8 \times 14'.6 \) data set of Eastwood et al. (2018). The \( T_0 = 154 \) K of the nominal-resolution map is shown to approximately hold at high-resolution. The \( \delta = 45^\circ \) FWHM provided here is extracted from the tabulated high-resolution PSF.
of course does not bias the measurements of the spectrum, carried out in §5 and §6.

We may now stack the flux excess (3) around the clusters in our sample to obtain an image of the mean flux excess per cluster at a given channel $\nu$, 

$$\Delta F(\nu, \tau \equiv \{\tau_x, \tau_y\}) = \frac{1}{N_c} \sum_{c=1}^{N_c} \sum_{h=1}^{N_p(\tau, c)} \Delta F(\nu, c, h) \ ,$$

where the second sum runs over the $N_p(\tau, c)$ HEALPix pixels mapped from cluster $c$ onto the desired $(\tau_x, \tau_y)$ pixel. Radial binning, used in the remainder of this subsection, is obtained by formally identical summations,

$$\Delta F(\nu, \tau) = \frac{1}{N_c} \sum_{c=1}^{N_c} \sum_{h=1}^{N_p(\tau, c)} \Delta F(\nu, c, h) \ ,$$

where the second sum here runs over the $N_p(\tau, c)$ HEALPix pixels that fall in the radial $\tau$ bin of cluster $c$. We henceforth focus mostly on the radially binned Eq. (5).

The significance $S$ of the excess intensity over the background in some bin $\tau$ of the cluster $c$ may be estimated as

$$S(\nu, \tau, c) = \frac{\sum_{h=1}^{N_p(\tau, c)} \Delta T(\nu, c, h)}{\eta(\nu)T_0(\nu)\sqrt{N_p(\tau, c)n}} \ ,$$

where $n$ is the number of (correlated, HEALPix) pixels in the beam,

$$n \simeq \frac{\pi (\text{FWHM}/2)^2}{\delta\Omega} \simeq 60 \left(\frac{\text{FWHM}}{15'}\right)^2 \ .$$

Here, we assumed that the noise is predominantly thermal, at the $T_0(\nu)$ antenna noise temperature, and given by Poisson statistics among the $\sim N_p/n$ independent beams mapped onto the $\tau$ bin. The product $T_0 \cdot \text{FWHM}$ varies with the PSF across the sky, so its effective, co-added value could deviate from its nominal, tabulated value. Therefore, in Eq. (6) we introduced order-unity correction factors $\eta$. These factors vary among channels, but in principle should be fairly independent across the sky, as shown below, and so roughly independent of $\tau$ and $c$.

The factors $\eta$ are calibrated by sampling the sky with control clusters, as discussed in §3.4 and in §E. In the cleared data, i.e. after CLEAN, $\eta$ varies between $\sim 0.7$ and $\sim 1.3$ as a function of frequency. In the restored data, $\eta$ tends to be larger and varies between $\sim 0.9$ and $\sim 1.9$ for different frequencies. In both cleared and restored data, $\eta$ shows small, $\lesssim 5\%$ spatial variations across the stacked control maps. Hence, we define mean factors $\bar{\eta}$, by averaging $\eta$ in the domain $0 < \tau < 10$ uniformly in $\tau$. The resulting averaged factors are used throughout the following analysis. Table 1 provides the averaged factors, as well as the local factors $\eta_\ell$ estimated near the anticipated shock position. The factors $\eta$ and $\eta_\ell$ are consistent within $\lesssim 6\%$.

There are different ways to stack the signal over clusters. We consider two opposed methods for estimating the significance of an excess signal stacked over multiple clusters, for given $\nu$ and $\tau$. We refer to the first, more standard method, as beam co-addition (equivalent to photon co-addition in RK18). Here, we compute the excess brightness summed over all clusters, and estimate its significance over random fluctu-
Near the virial radius, we find co-addition, consider for example the high-frequency channel. In the second method, referred to as cluster co-addition (also used in RK18), we co-add the significance of the coincident excess \( S(\nu, c, h) \) in individual clusters as unit-normal random variables of equal weights, leading to

\[
S^{(cl)}(\nu, \tau) = \frac{\sum_{c=1}^{N_c} \sum_{h=1}^{N_h} \Delta T(\nu, c, h)}{\eta(\nu)T_0(\nu)/\sqrt{n \sum_{c=1}^{N_c} N_h(\tau, c)}},
\]

(8)

directly generalizing Eq. (6) for multiple clusters. As this method weighs the stacked excess against the stacked solid angle, it has the advantage of being less sensitive to individual clusters with a small \( \theta_{500} \) that leaves fewer independent beams for determining \( S(\nu, c, h) \).

As each cluster contributes equally, regardless of its angular extent, this method avoids the large weights that beam co-addition effectively assigns to both the excess and the estimated fluctuations of the most extended clusters.

The radially-binned Eqs. (8) and (9) are readily generalised for a \( S(\nu, \tau_x, \tau_y) \) map, by replacing the \( \tau \) bins by \( (\tau_x, \tau_y) \) pixels, as in Eq. (4). The beam and cluster co-addition methods, analysed separately below, are generally found to yield results in good agreement with each other, provided that bright NVSS sources are masked. The correction factors \( \eta \) are calibrated separately for beam and cluster co-addition after masking NVSS sources, showing small, \( \leq 5\% \) differences between the two stacking methods for both cleared and restored data.

To demonstrate the agreement between beam and cluster co-addition, consider for example the high-frequency channel. Near the virial radius, we find \( \lesssim 0.03 \) (\( \lesssim 0.07 \)) changes in \( S \), which correspond to \( \lesssim 3\% \) (\( \lesssim 6\% \)) fractional differences, for the cleared (restored) data (see §4.1.2). The largest deviations in \( S \) between the two methods, \( \lesssim 0.4 \) (\( \lesssim 0.3 \)), are found in the central, \( \tau \lesssim 1 \) regions of the clusters. Substantial differences between beam and cluster co-addition can be seen in the restored data when bright NVSS point sources are not masked (see §F1).

We use control samples to verify (see §3.4 and §E) that the distribution of the excess emission above the background, stacked around mock clusters randomly placed on the sky, is approximately normal in each channel when using the cleared data. In the restored data, this distribution is slightly skew-normal but, within a \( \sim 0.1 \) accuracy in \( S \), can be approximated as normal. The control samples are used to calibrate the noise parameters \( \eta \), as summarised in Table 1. The distribution of the calibrated significance \( S \) attributed to the excess is then found to approximately follow a unit-normal (i.e. zero mean and unit variance) distribution. Namely, the distribution of control samples is found to be consistent with the calibrated Eqs. (8) and (9), out to at least the \( \pm 3\sigma \) confidence level of a normal distribution.

Our analysis is largely based on the above, single-channel expressions, applied to the high-frequency channel, where the angular resolution is superior. The low-frequency channels are mostly used to measure the spectrum and verify the signal. However, we also co-add the different channels, adopting the significance estimate

\[
S(\tau) = \frac{\sum_{\nu=1}^{N_{\nu}} S(\nu, \tau)}{N_{\nu}^{(1+\psi)/2}},
\]

where \( N_{\nu} = 7 \) is the number of co-added, low-frequency channels. Anticipating some degree of inherent correlation between the different channels, here we introduced a correlation parameter \( \psi \), which can in principle range from 1 (full correlations) to 0 (no correlations) or even become negative (for anti-correlations), and is determined empirically below based on control samples.

These control samples indicate strong positive inter-correlations among the seven low-frequency channels. The simple correction factor in Eq. (10) is motivated by the approximately unit-normal \( S \) distribution it produces once \( \psi \) is calibrated; we thus find that \( \psi \sim 0.7 \) (\( \psi \sim 0.85 \)) for the cleared (restored) data. The strong correlations appear to emerge primarily from the similar small-scale structures seen across the radio sky in the different channels, rather than from any instrumentation effects, because point source restoration significantly strengthens the correlations and because the correlations among different channels roughly scale with their PSF. As such, Eq. (10) provides a reasonable and simple alternative to the determination of the full covariance matrices. Notice that introducing a positive \( \psi \) provides a conservative estimate of \( S \): it effectively lowers the significance attributed to the appearance of the virial excess emission in multiple channels.

The channel co-addition Eq. (10) is applied to both beam and cluster stacking methods, with typically similar results. For example, the two methods show \( \lesssim 0.2 \) (\( \lesssim 0.3 \)) differences in \( S \) corresponding to \( \lesssim 7\% \) (\( \lesssim 10\% \)) fractional changes, for the cleared (restored) data around the virial radius (see §4.2.2). In terms of \( \psi \), we find \( \lesssim 1\% \) differences between beam and cluster co-addition.

It is instructive to compute the stacked intensity excess profile, \( \Delta I \equiv \nu \Delta I_{\nu}(\nu, \tau) \), and relate it to the significance \( S \) of the local excess. The procedure is similar to the above stacking of \( S \), as discussed in §B.

### 3.3 Background modelling

The virial signal is typically weaker than the Galactic foreground, the extragalactic background, and the emission from the centre of the cluster, so a careful removal of the background is essential. We consider several different models for the \( T_b(\nu, c, h) \) background field.

The simplest model is a uniform background around each cluster, which we define as the mean brightness temperature \( T_b(\nu, c) \) measured within an angular separation \( \theta < \theta_b \) of the centre of the cluster. We find that \( \theta_b \approx 5^\circ \) is sufficiently large to establish the virial ring, and not too large to be offset by the large-scale curvature of the Galactic foreground. Our results do not vary substantially with small changes in \( \theta_b \) around this value (see §C). This pertains both to the above uniform background and to the polynomial backgrounds discussed next. Therefore, we adopt \( \theta_b = 5^\circ \) as our nominal value.

More accurate models allow spatial variations, to better capture the local background. As our cluster selection avoids extended Galactic structures, namely the Galactic plane, Loop-I, and the Fermi bubbles (see selection criteria (iv) –
(vi) in §2.2, we apply identical masks on the data, with a negligible effect on the results. After bright point sources and these Galactic structures are masked, the remaining background varies mainly on scales larger than the anticipated, $\lesssim 1^\circ$ extent of the virial signal. This remaining background can be fairly well approximated using a polynomial fit on larger scales. For each cluster, we thus consider the region $\theta < \theta_b$, and fit the enclosed data in each band by an order $N_0$ polynomial in $(\tau_x, \tau_y)$.

In §C, we consider different choices of $N_0$ and $\theta_b$, quantifying their implications using the control samples presented in §3.4 below. Overall, we find that the statistical distribution of $S(\tau)$ is approximately symmetric, normal and normalised, as constructed above, provided that $N_0 \geq 2$. In contrast, a low, $N_0 \leq 1$ order polynomial is not sufficient to capture the curvature in the background, and leads to $S(\tau)$ distributions that are neither symmetric, nor normal, nor normalised. However, higher orders $N_0$ cause increasingly more of the virial excess itself to be fitted as part of the background, thus spuriously diminishing the estimated signal.

These considerations indicate that $N_0 = 2$ or $N_0 = 3$ best balance a fitting power sufficiently high to render $S(\tau)$ normal and normalised, but sufficiently low to avoid over-fitting much of the local excess due to the cluster as part of the background. Odd-order terms in a polynomial are integrated out by the radial binning, so for any odd $N_0$, the results are very similar, although not identical, to those of $N_0 - 1$. Therefore, and for better accuracy in the $(\tau_x, \tau_y)$ stacked maps, we adopt $N_0 = 3$ as the nominal polynomial order. Interestingly, similar background properties were found by RK18 when fitting the stacked brightness of Fermi clusters. For our nominal, $N_0 = 3$, the results are not sensitive to the fitting range for $\theta_b \geq 5^\circ$. For sensitivity tests and a discussion of alternative choices of $N_0$, see §C.

3.4 Control sample tests

In order to estimate the statistics of the stacked quantities, calibrate the parameters $\eta$ and $\psi$, test Eqs. (8)-(10) for $S$ (and their counterparts (B3)-(B4), and (B6) for $I$), examine different background models, and investigate possible systematic biases, we generate control samples of mock clusters and stack the real LWA data around them with the same pipelines applied to the MCXC cluster sample. For each test, we thus generate $N_{\text{mock}} = 5000$ control samples, each chosen to mimic the real cluster sample. This value of $N_{\text{mock}}$ samples the analysed sky many ($\gtrsim 40$) times, leading to sufficiently converged estimates ($< [2\%, 2\%, 3\%]$ in $\{1, 2, 3\} \sigma$ cleared-data containment around $\tau \sim 2.4; < 1\%$ in $\eta$ and $\xi$).

Each control sample consists of 44 control clusters with random positions on the sky, chosen under the same geometric constraints as the real sample, and in addition, constrained to lie at least $1.5^\circ$ away from any MCXC catalogue source and any bright (1.4 GHz flux exceeding 1 Jy) point source from the NVSS catalogue. We verify that the latter constraint, avoiding bright NVSS sources for more efficient control samples, has no significant effect on our nominal results, as pixels are already masked within $0.3^\circ$ from such bright NVSS point sources. The control clusters are assigned with the same relevant parameters (angular scale and, for modelling the virial signal, also mass and redshift) as the real clusters, and are identically analysed. The significance $S(\nu, \tau)$ of the mock stacked excess is then estimated for each cluster sample.

Figure 2 shows the resulting distribution of $S$ around the control cluster sample, for nominal analysis parameters, in particular $N_0 = 3$ and $\theta_b = 5$ for background removal, and a mask of radius $0.3^\circ$ around $F_1 \equiv 0.1$ Jy NVSS sources. The figure shows results for both the high-frequency (left panels) and the co-added low-frequency (right) channels, using either cleared (top row) or restored (bottom row) data. The results are stacked using both beam (thin lines) and cluster (thick lines) co-addition, found to be in good agreement with each other.

The results are shown after calibrating the single-channel noise correction factors $\eta$, and the multi-channel correlation factor $\psi$. These parameters are tuned such that $S$ is normalised to have unit variance, $\sigma(S) = 1$, when averaged over $\tau$, in each channel separately and when the low-frequency channels are co-added; the resulting $\eta$ and $\psi$ values are not far from unity (see Table 1). The figure then contrasts, for all relevant $\tau$ bins separately, the nominal values of $S(y$-axis labels) inferred from Eqs. (8)-(10), against the mock confidence intervals (dashed curves) extracted from the control samples up to the maximal, $\pm 3\sigma$ levels available with $N_{\text{mock}}$, and against the normal-distribution intervals $\mu \pm n\sigma$ (dotted curves) anticipated from the locally measured mean $\mu$ and standard deviation $\sigma$ of the control samples for $n = 1, 2, 3$. The respective curves should therefore coincide with each other and with integer values of $S$ only when the underlying distribution of $S$ is locally symmetric, normal and normalised; see §D and §E for details.

Overall, we find that the stacking equations for $S$ [Eqs. (8)-(10)] and $I$ (§B) can be applied to the cleared data, with minor errors with respect to the control sample distribution, even when the latter is extrapolated as a normal distribution to high-confidence levels. Our tests indicate that in the radial range $0.5 < \tau < 10$, we may then slightly underestimate the significance of an excess signal by $-0.06 \lesssim |\Delta S/S| < 0$ in the high-frequency channel, and overestimate it by $0 < |\Delta S/S| \lesssim 0.02$ in the co-added low-frequency channels. In particular, around the virial radius, the nominal confidence levels are conservative (slightly exaggerated) and consistent with the extrapolated control sample distributions within $\gtrsim -4\% \lesssim +3\%$, for the high (co-added low) frequency data. For clarity, when displaying the $S$ profiles measured from the real sample, we overplot the extrapolated significance levels of the corresponding control sample.

4 STACKING RESULTS

We begin in §4.1 by presenting our nominal results, based on the high-frequency channel. To test these results and constrain the spectrum, in §4.2, we present the co-added results for the seven low-frequency channels. We focus on the cleared data, showing restored data results mainly to clarify the role of point sources. The correction factors, namely $\eta$ for channel noise and $\psi$ for cross-channel correlations, are used throughout, as inferred above from the control cluster catalogues.
B1

4.1 Nominal analysis: high-frequency channel

In our nominal analysis, we focus on the high-frequency, 73.15 MHz channel. Here, the spatial resolution is considerably better than in other channels, complications due to PSF sidelobes are minor, and the control samples are well behaved even without any noise corrections. We inspect the brightness profile in §4.1.1, and the significance of the implied excess in §4.1.2; a pronounced excess near the virial is shown to be robust in §4.1.3, and argued to be genuine and not associated with beam sidelobes in §4.1.4.

Nominal parameters are used (henceforth, unless otherwise stated), including the standard cuts giving our nominal, 44 cluster sample, a radial binning with $\Delta\tau = 0.25$ resolution, polynomial background modelling of order $N_0 = 3$ within $\theta_b = 5^\circ$, masking data within 0.3 of NVSS point sources with a flux exceeding 1 Jy, and applying the (small) $\eta$ correction.

4.1.1 Brightness profile

Figure 3 shows, for both cleared and restored data sets, the radial, $\tau$-profile of the mean excess brightness, averaged over the 44 clusters in our sample, as defined in §B [namely, using Eqs. (B1) and (B2), with 1σ uncertainty intervals (error bars) computed using Eqs. (B3) and (B4)]. Nominal parameters are used.

Both beam and cluster co-addition methods are shown in the figure, and are found to be in good agreement with each other, here and quite generally, except where point sources make a strong contribution. This agreement indicates that the diffuse features are fairly evenly stacked among the clusters, rather than dominated by a handful of outlier clusters.
We verify this behaviour by manually checking the excess in individual clusters, estimating the mean and standard deviation of the excess, and removing the strongest contributing clusters, as demonstrated for the virial signal in §4.1.2. For simplicity, we focus on beam co-addition for the remainder of §4.1.1.

As expected, a strong signal is identified in the centres of clusters, and is partly removed by the modified CLEAN algorithm. The original signal, as inferred from the restored data, which more faithfully represents the signal including point sources, is peaked approximately at the centre of each cluster, with an excess intensity of $\Delta I \approx 1.2 \times 10^{-10}$ erg s$^{-1}$ cm$^{-2}$ sr$^{-1}$ when averaged over clusters. This component, likely attributed to a combination of point sources, such as radio galaxies, and diffuse emission, such as from radio halos and minihalos, is further discussed in §5.

Outside the clusters, at large, $\tau \gtrsim 3$ radii, well beyond the virial radius, the signal is in general consistent with noise, with only expectedly small, $\lesssim 1\sigma$ fluctuations. The cleared and the restored data are not in good agreement here, indicating that deviations of the restored data from the background are largely associated with point sources. The low-significance cleared-data excess emission around $\tau \approx 3.5$ ($\sim 2\sigma$) and around $\tau \approx 5.5$ ($\sim 2.5\sigma$) are discussed below in §4.1.2.

In the peripheries of clusters, around $2 \lesssim \tau \lesssim 3$, both data sets show a localised, significant excess peaked in the 2 $\tau$ bin, coincident with the virial shock signal found in RK18 by stacking Fermi data (shown as a vertical yellow band in the figures), as well as with other observations and previous theoretical predictions outlined in §1. This excess, when averaged over clusters, is $\Delta I \approx 5.0 \times 10^{-12}$ erg s$^{-1}$ cm$^{-2}$ sr$^{-1}$ in the cleared data, and ~10% smaller in the restored data. The similar radius, magnitude, and shape of the excess in both data sets indicate that this signal is not associated with, nor strongly contaminated by, point sources. We henceforth refer to this signal as the virial excess.

### 4.1.2 Significance profile

We quantify the significance of the local excess above the background using both the nominal $S$ derived in §3.2 for a normal distribution, and the control samples described in §3.4. Figure 4 shows the radial profile of the excess significance $S$, computed according to Eqs. (8) and (9), as a function of $\tau$, for both the cleared (top panel) and the restored (bottom) data, using both beam (abbreviated “bm”; symbols) and cluster (abbreviated “cl”; curves) co-addition. The sensitivity of the results to alternative parameter choices is explored in §H. Our analysis is based on the $0 \leq \tau < 10$ region, sufficiently large to capture three times the radius of the $\gamma$-ray excess, or equivalently, more than four times the FWHM even in the lowest-frequency channel; the significance figures below focus on the $0 \leq \tau < 7.5$ range only for visibility.

As discussed in §3.4, we overplot the confidence-level estimates based on the corresponding control samples, extrapolated to high orders in the normal-distribution approximation. Namely, the figure shows (as dotted lines) the mean of the beam co-added $S$, offset by integer multiples of its standard deviation, based on the respective control samples. The nominal confidence levels are consistent with these extrapolated confidence levels (e.g., compare dashed vs. dotted curves in Fig. 2), within the accuracy discussed in §3.4. Namely, outside the central $\tau < 0.5$, our nominal significance estimates are consistent with the extrapolated levels within ~4% (~10%) for the cleared (restored) data.

Consider first the restored data, shown in the bottom panel of Fig. 4. Here, the central signal within $\tau \lesssim 1$ is highly significant, reaching the $\sim 18\sigma$ level. Beyond a local minimum (the ‘dip’, henceforth) seen around $\tau \approx 1.5$, the nominal stacked signal (circles and thick dot-dashed curve) reaches the $\sim 3\sigma$ confidence level in the peak, $2.25 \leq \tau < 2.5$ bin, with only $\lesssim 1\sigma$ fluctuations at large radii beyond $\tau \approx 3$. Both the signal and the fluctuations appear to be somewhat extended in $\tau$. Using thicker, $\Delta \tau = 0.5$ bins (squares and thin double-dot-dashed curve) raises the significance of the virial excess to the $\sim 4\sigma$ level in the $2 \leq \tau < 2.5$ bin, while the fluctuations reach $\sim 2\sigma$. The significance is also increased if these $\Delta \tau = 0.5$ bins are first offset by $0.25$ in $\tau$ (triangles), but only to the $\sim 3.4\sigma$ level in the $2.25 \leq \tau < 2.75$ bin. The location of the peak is consistent with the stacked Fermi data analysis in either choice of angular binning, and in both stacking methods.

Next, consider the cleared data (top panel), after point sources were removed by CLEAN and not restored. Here, the central peak is diminished to the $\sim 6\sigma$ level, as expected...
due to its point source contribution, whereas the virial excess becomes more significant, with a $\sim 4.2\sigma$ peak in the same, $2.25 \leq \tau < 2.5$ radial bin. Thus, while point source removal does not appreciably change the magnitude or position of this virial signal (see Fig. 3), it raises its significance by lowering the background noise level. Such behaviour is consistent with this excess arising from diffuse, rather than point-source, emission. The peak remains broad; using thicker, $\Delta \tau = 0.5$ bins again raises the significance, here to the $\sim 5\sigma$ confidence level both in the $2.0 \leq \tau < 2.5$ bin and in the offset, $2.25 \leq \tau < 2.75$ bin. The nominal excess is comprised of the cumulative contribution of $\sim 70\%$ of the clusters in our sample, with a mean $S \simeq 1.1 \pm 0.7$ in the $2 < \tau < 3$ range. No sub-sample strongly governs the signal, so removing the few most significant, $\{3.3, 2.4, 2.2, \ldots\}\sigma$ contributions does not significantly alter the stacked results.

Unlike the virial excess, discussed in some detail below in §4.1.3, the central, $\tau \lesssim 1.5$ signal is highly sensitive to point source removal, which substantially lowers its brightness (Fig. 3) and corresponding significance (Fig. 4). While most of the central flux is thus attributed to point sources, the residual stacked signal in the cleared data suggests a $\sim 6\sigma$ diffuse central component, extending in out to $\sim 1R_{500}$. This diffuse signal shows some $|\Delta S/S| \lesssim 7\%$, or equivalently $|\Delta S| \lesssim 0.4$, difference between the two stacking methods (compare symbols and corresponding curves in Fig. 4). In contrast, the virial peak shows only a $\sim 0.4\%$ difference between the two stacking methods, suggesting that the virial excess is more evenly distributed among clusters than the central cleared signal.

Although the noise level is reduced in the cleared data, point source removal appears to amplify some of the $\lesssim 1\sigma$ fluctuations, found in restored data outside the clusters ($\tau > 3$), resulting in apparently extended, $\sim 2\sigma$ excess signals in cleared data near $\tau \simeq 3.5$ and near $\tau \simeq 5.5$. Consider first the $\tau \sim 3.5$ excess. This signal is seen only in the high-frequency channel, and not in other frequencies (see §4.2.2), so we cannot substantiate its viability. Furthermore, this signal is dominated by only three clusters, each showing $\tau > 2.5\sigma$ excess at $\tau \simeq 3.5$. Two of these clusters harbour a (partly masked) extended, $> 1$ Jy NVSS point sources at the same radius. Excluding these three clusters lowers this excess to $< 1\sigma$, with only a minute effect on the virial signal (see stars in Fig. 4). Hence, we disregard this $\tau \approx 3.5$ signal as a fluctuation. In contrast, the $\tau \sim 5.5$ signal — like the virial excess — reflects multiple small contributions from a large fraction ($\sim 70\%$) of our clusters, and is persistent, seen in all channels. However, this signal is less significant than the virial excess and more sensitive to CLEAN, and it becomes far less significant or even vanishes entirely when the declination selection criterion is varied (see §7.3). We further discuss this $\tau \approx 5.5$ signal in §7.3.

4.1.3 Virial excess

Figure 4 and the preceding discussion indicate that the virial excess in the high-frequency channel is robust and significant at the $4-5\sigma$ confidence level. The signal peaks around $\tau \approx 2.4$, coincident with its $\gamma$-ray counterpart, identified earlier by stacking Fermi data. The excess is somewhat extended in $\tau$, contributing to bins over the $2.0 \leq \tau \lesssim 2.8$ range. The location, magnitude, and shape of the signal inferred from beam vs. cluster co-addition are in good agreement with each other. This agreement supports the robustness of these statistics, and indicates that the stacked excess arises from many of the clusters in our sample, rather than being dominated by a handful of outlier clusters. The agreement between the excess parameters in cleared vs. restored data indicates that the excess is diffuse, and not dominated or distorted by point sources. These conclusions are confirmed by inspecting the individual significance profiles of all clusters in our sample.

The rescaled, stacked virial signal is sufficiently strong to be seen even without full radial binning. Figure 5 shows a partly-folded map of the significance $S(\tau_x, \tau_y)$ of the stacked excess, prepared as in the radial plots but without radial binning. Namely, an excess flux map is first generated for each of the 44 clusters in our sample, using the cleared data. Each such map is rescaled by the respective $R_{500}$, and rotated randomly about the centre of the cluster to wash out any co-added large-scale gradients. The resulting maps are then stacked using beam co-addition, according to Eq. (8). To better show the signal, the resulting image is then folded onto a quadrant or an octant (see labels). The nominal, high-frequency signal is folded onto quadrant 2 (upper left sector). To further emphasize the signal, we fold the same image again, onto octant 5 (just below quadrant 2). The bottom quadrant and octant show the low-frequency signal, obtained by co-adding the seven low-frequency channels as discussed in §4.2. Quadrant 1 (upper right) combines the high-frequency channel with the co-added, low-frequency channels. Both the central, $\tau \lesssim 1$ signal, and a virial ring peaked around $\tau \sim 2.4$ are evident. The $2.2 \lesssim \tau \lesssim 2.5$ range, identified (RK18) as a virial ring in stacked $\gamma$-rays, is also shown (dotted orange

![Figure 5](image-url)

**Figure 5.** Folded significance map of the cleared data excess, after clusters were rescaled, randomly rotated, and beam co-added (see labels and colourbar). The data are folded onto either quadrants or (in quadrant q3) octants, to better highlight the signal. Shown are the high-frequency channel quadrant q2 and octant o6), the co-added low-frequency channels (quadrant q4 and octant o6), and the two results combined (quadrant q1); see labels. The $2.2 \lesssim \tau \lesssim 2.5$ best-fit range of shock radii based on the RK18 stacking of Fermi data is also shown (between orange dotted circles). The same figure with a larger $\tau$ range is shown in §M.
circles). Figures with a larger $\tau$ extent and modified selection criteria are shown in §M.

In order to further test the robustness of the virial signal, we show (see §4.2.3) that it presents in all channels, and carry out a series of sensitivity and convergence tests in §C and §H. In particular, we examine the sensitivity of the signal to the cluster selection criteria, the background modelling order, the background fitting area, and different radial bin sizes. Overall, the results are found to be insensitive to reasonable variations in these parameters.

4.1.4 Ruling out a sidelobe origin

Given the strong central signal, especially before point sources are removed, sidelobes become a major concern: one must carefully test if the virial ring may arise simply from stacked interferometric ripples around the central signals. In principle, such a spurious effect should be ruled out by the CLEAN algorithm, as the identified point sources near the centre were convolved with the PSF before they were removed, along with their sidelobes, and the remaining central signal in the cleaned data is not sufficiently strong to generate the virial excess as a sidelobe ripple. Nevertheless, putative artefacts due to inaccuracies in the algorithm should be ruled out before the virial excess can be securely established.

In subsequent parts of the analysis, we rule out a putative virial excess stemming from the residuals of sidelobe ripples of the central signal that were not removed by CLEAN. This conclusion is derived in several independent methods, including: (i) a significant virial excess signal is found also in those clusters that do not show a central signal (as discussed below and shown in Fig. 6); (ii) Monte-Carlo simulations of PSF-convolved central emission models do not show the observed virial ring (§F1); (iii) multiple channels show the signal at similar ring radii, despite variations among their PSF patterns (§4.2); (iv) removing one or a few outlier clusters with bright NVSS sources in their centre strongly diminishes the raw (restored, unmasked) stacked central signal, with no significant effect on the virial excess (§G); (v) intentional attempts to generate an artificial virial excess by choosing control, centrally-bright clusters, fail to produce a significant signal (§G); (vi) the spectrum of the virial excess is found to be a robust power-law (§6.4), whereas the central signal shows a non-trivial spectrum varying spatially and among clusters (§F1); and (vii) the virial excess is well-fit by the expected model: spatially coincident with the previously stacked $\gamma$-ray signal and of a normalisation comparable to that inferred in Coma (Keshet et al. 2017), not far from equipartition (§7).

Although most clusters show a bright central signal, at least before point sources are removed, some of them do not. The first method we invoke to test for sidelobe artefacts is to focus on those clusters that do not show a central excess. In such clusters, an excess signal near $\tau \approx 2.4$ clearly cannot be dismissed as sidelobes of emission from the centre. We define such centre-faint clusters as those showing an intensity deficit, $\Delta I < 0$, in their central, $\tau < 0.5$ region, after the large-scale background was removed. For the selection of these clusters, we use the raw data, i.e. the restored data without additional NVSS point-source masking, in order to include the full imprint of all point sources that may have left some sidelobes contamination in the cleared data even after CLEAN. We adopt nominal parameters, focusing on the high-frequency channel.

After removing the background, 13 out of the 44 clusters in the sample are classified as centre-bright, while 31 clusters are classified as centre-faint. Figure 6 shows the stacked excess intensity profiles of these centre-faint vs. centre-bright clusters, for both cleared (top panel, with masking) and restored (bottom, without masking) data. As the cleared data show, the centre-faint clusters make a substantial contribution to the virial ring signal, comparable to and even exceeding that of the centre-bright clusters, when accounting for the different sizes of each sub-sample. Namely, the virial excess in the cluster-faint clusters is significant at the $\sim 2.5\sigma$ confidence level, slightly higher than anticipated from Poisson statistics when taking into account the smaller size of this sub-sample.

The mean intensity of the centre-faint clusters agrees with that of the centre-bright clusters in the virial radius bin. As expected, the two sub-samples show different, $\lesssim 2\sigma$ fluctuations outside the clusters, beyond $\tau \approx 3$. The centre-faint signal is somewhat broader than found for the full sample, extending out to $\tau \approx 3$, in resemblance of the low-frequency signal discussed in §4.2. We conclude that the agreement between the virial signal as inferred from centre-faint vs. centre-bright clusters supports the virial signal as a genuine effect, rather than an artefact due to a putative failure of CLEAN.
to remove PSF sidelobes of the emission from the centres of clusters.

4.2 Low-frequency channels

Next, consider the seven low-frequency channels, in order to test the results of the nominal, high-frequency analysis in §4.1 and to extract some spectral information. In these frequencies, the resolution is substantially lower than in the high-frequency channel, PSF sidelobes are more pronounced, and the control samples indicate noise characteristics that require somewhat larger correction factors \( \eta \), as provided in Table 1. In order to obtain robust results, we therefore co-add the seven channels, taking into account the correlation factor \( \psi \), unless otherwise stated.

4.2.1 Brightness profile

Figure 7 presents the mean excess brightness profiles of the co-added low-frequency channels, for both the cleared and restored data. The figure shows \( \Delta I(\tau) = \nu \Delta I_\nu \), which is optimised for an \( I_0 \propto \nu^{-1} \) spectrum, computed as described in §3B [using Eq. (B5)], with uncertainties computed according to Eq. (B6)]. Again, we use nominal parameters as in §4.1, and find that the two stacking methods yield consistent results. For simplicity, we focus here (throughout §4.2.1) on beam co-addition.

![Figure 7. Brightness excess profile of the seven co-added low-frequency channels, averaged over clusters. Notations are the same as in Fig. 3.](image)

The brightness profile in the low-frequency channels is similar to the nominal, high-frequency behaviour discussed in §4.1, with expected differences due to the poorer resolution available here. Strong central emission in the restored data is found to peak, when averaged over clusters and channels, at \( \Delta I \approx 1.1 \times 10^{-10} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1} \) (using beam co-addition), quite similar in shape and in magnitude to the high-frequency channel. Outside the clusters, at \( \tau \gtrsim 3 \), a very broad, \( \sim 2\sigma \) excess is seen throughout the \( 4.5 \leq \tau \leq 6 \) range, coincident with the \( \tau \approx 5.5 \) high-frequency excess discussed in §4.1.2 and in §7.3.

In the periphery of the clusters, a significant excess is found around \( \tau \approx 2.6 \), broadly consistent with the virial signal seen in the high-frequency channel in terms of location, width, and amplitude. The excess is again quite similar in cleared and restored data, suggesting that it is mostly diffuse in nature. However, some differences between the two data sets suggest that contamination by point sources is stronger and more difficult to correct at low frequencies, due to the poor resolution and a possible abundance of soft-spectrum sources.

This virial excess in the cleared data peaks, when averaged over clusters and channels, at \( \Delta I \approx 2.8 \times 10^{-12} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1} \) (for beam co-addition). This value is similar to its high-frequency counterpart, supporting the detection of the high-frequency signal and affirming that it is not a PSF artefact, as discussed above (in §4.1.4) and further below (in §7 and §G). The virial excess at low frequencies is slightly (\( \sim 8\% \)) less significant than in the high-frequency channel, possibly because the signal is convolved with more extended PSFs at low frequencies, rendering the signal broader (see §C); a more careful analysis of the signal in different channels, along with a derivation of the implied spectrum, is deferred to §4.2.4 (backward modelling) and §6.4 (forward modelling).

The virial signal is slightly broader in the low-frequency channels than it is in the high-frequency channel, as one might expect from the lower resolution. Here, the signal peaks in the \( 2.5 \leq \tau < 2.75 \) bin, adjacent to the \( 2.25 \leq \tau < 2.5 \) bin of the high-frequency peak. The broader signal at low frequencies shows a significant excess coincident with the \( 2.2 \leq \tau \leq 2.5 \) range of the stacked \textit{Fermi} signal and with the high-frequency LWA channel, but it extends slightly beyond \( \tau \approx 3 \). In addition to the lower resolution, this broadening could be in part attributed to the contribution of faint, soft point sources. This possibility is supported by the substantial differences between the profiles obtained from cleared vs. restored data, for example the pronounced minimum around \( \tau \approx 3.7 \) in the latter. We revisit the issue in more detail in §4.2.3.

4.2.2 Significance profile

Following the same procedure applied to the nominal, high-frequency channel in §4.1.2, next we compute the significance of the excess in the low-frequency channels. Figure 8 shows the resulting significance profiles \( S \) for the cleared (top panel) and restored (bottom) data, both for individual channels (thin curves without symbols) and for the excess co-added over the seven low-frequency channels (symbols and thick curves). Confidence levels extrapolated from the corresponding (channel co-added) control samples are also shown (dotted curves), using the same procedure as for Fig. 4.

Like in the high-frequency case, the central signal is very significant at low frequencies, with a \( \sim 25\sigma \) excess in the restored data. The dip here is seen at \( \tau \approx 1.75 \), just outside its high-frequency counterpart, presumably because of the lower spatial resolutions. When using the cleared data, the central signal is reduced to the \( \sim 12\sigma \) level, consistent with a dominant point-source contribution as found in the high-frequency case.

For \( \Delta \tau = 0.25 \) resolution, both cleared and restored data sets show a virial signal peaked around \( \tau = 2.6 \), seen separately in almost every individual channel, as well as in the co-added result. The exception is the 52.2 MHz channel, where the peak is found farther out, around \( \tau \approx 3 \). The virial signal is not strongly affected by the removal of point sources, although it does strengthen from \( \sim 2\sigma \) in the restored data to \( \sim 3.8\sigma \) in the cleared data, as expected from a diffuse signal contaminated by point sources, and as seen in the nom-
Focusing in more detail on the virial excess, we examine its high-frequency counterpart, and is comprised of small contributions between beams and clusters co-addition methods. In particular, we find ~7% differences between the two stacking methods in the cleared data. However, these discrepancies are not significant, and they appear to be at least in part associated with residual point sources. As expected, the cleared data signals show better agreement between the two stacking methods than found in the restored data.

In low frequencies, the virial signal peaks consistently at radii somewhat larger than the τ ~ 2.4 peak found in the high-frequency channel, an effect attributed in part to the better resolution of the latter. Six of the seven low-frequency peaks in the 2.5 ≤ τ < 2.75 bin when using the cleared data, just outside the adjacent, 2.25 ≤ τ < 2.5 bin of the high-frequency peak. With a finer, Δτ ≃ 0.167 (6 bins per τ; see §H) resolution, we again find the low-frequency peak adjacent to the high-frequency peak with the same high resolution, indicating that the effect is small. We conclude that the peak shifts from τ ~ 2.6 in the co-added low-frequencies, to τ ~ 2.4 in the better resolved, high-frequency.

Interestingly, when using the restored data, some of the peaks shift to slightly larger radii. This suggests that the discrepancy with respect to the high-frequency channel may in part be associated with low-frequency contamination from soft point sources, not fully removed by CLEAN. Another possibility is that the virial emission from individual clusters has a broad distribution, extending out to τ ~ 3 or even beyond that, with some preference around τ ~ 2.4 that is picked up when sufficient high-resolution data are stacked. Such an interpretation could explain why the virial signals both for the low-frequency results and for the centre-faint, high-frequency results (see §4.1.4) have a similarly broad τ distribution, with a peak similarly shifted to τ ~ 2.6.

The comparable peak radii found in the different low-frequency channels provide additional evidence that the virial excess is unlikely to be an artefact due to uncleared sidelobe contamination from the central signal. Indeed, such putative ripples would show in different radii, corresponding to the frequency-dependent PSF, as we verify directly in §F1.

4.2.4 Spectra

Next, consider a crude estimate of the spectra of the central and virial signals, before deprojecting and modelling them in §5 and §6. The specific brightness of the two signals in all eight channels is shown in Fig. 9, as a function of frequency. For this purpose, we define the central signal as the τ < τ_{max} excess, exploring different choices of τ_{max}; the figure focuses on the most central, τ_{max} = 0.17 region accessible.
with present resolution. As the virial signal is extended radially at low frequencies, we define it for spectral purposes as the mean excess in the broad, $2.25 \leq \tau < 3.0$ range. When considering the central signal, we examine both the raw excess, in the restored data without NVSS masking, and the cleaned excess, in the cleared data with masking. The former includes the contribution of point sources, and is useful as it more accurately reflects the spectrum of any putative sidelobe effects due to the central signal. The latter is a closer representation of the underlying, diffuse central signal, although some point source contamination is likely to persist.

**Figure 9.** Spectrum (error bars with dotted lines to guide the eye) of the raw (restored data, non-masked; top red) and the cleaned (cleared data, NVSS-masked; middle blue), central ($0 \leq \tau < \tau_{\text{max}} = 0.17$) signal, as well as the virial signal (in a broad, $2.25 \leq \tau < 3.0$ region; cleared data; bottom green). At $\nu > 50$ MHz frequencies, where the PSF depends weakly on frequency, the central signal can be reasonably well fit as a power-law spectrum (solid curves with 1σ confidence levels), and extrapolated (dashed curves) to lower frequencies. For the virial signal, systematic (PSF and background) and statistical uncertainties are prohibitively large for such a backward determination of the spectrum.

In principle, the spectrum should be estimated by forward modelling, where different model signals are injected into control samples, which are then analysed with the real pipeline in order to determine the best-fitting model. We use this method to analyse the spectrum of the central signal in §5, and of the virial signal in §6. Here, we use a crude backward modelling to directly fit the binned spectrum in Fig. 9. The results of such modelling may be useful in regimes where projection effects are favourable and PSF corrections can be approximately accounted for. As in §3.4, let us define the spectral index $\alpha$ through $I_\nu \propto \nu^{-\alpha}$.

The raw central excess is the strongest among the three signals in the figure, and its statistical uncertainties (red error bars on the top curve) are negligible. At frequencies above 50 MHz, where the PSF no longer varies strongly with frequency (see Table 1) and the resolution is sufficient for $\tau_{\text{max}} = 0.17$, the data are found to be well-fit by a pure power-law, $\alpha = 0.8 \pm 0.1$, with $\chi^2_\nu \simeq 0.4$ chi-squared per degree of freedom. At lower frequencies, the spectrum appears to be harder, but this is likely due to the strong frequency dependence of the PSF. For larger $\tau_{\text{max}}$, the spectrum remains well-fit by a pure power law above 50 MHz, but the integrated spectrum gradually softens. Such softening suggests a considerable contribution from underlying diffuse emission, possibly in the form of radio halos and minihalos which often exhibit such a behaviour. At larger scales, the steepening appears to saturate at some $1.5 \lesssim \alpha \lesssim 2$; for $\tau_{\text{max}} = 1$, we find $\alpha = 1.5 \pm 0.1$ ($\chi^2_\nu \simeq 1.5$). At small scales, extrapolation to $\tau_{\text{max}} = 0$ gives $\alpha = 0.7 \pm 0.1$, consistent with a population of point sources.

It is interesting to compare the spectra of possible sidelobe ripples from the central signal, to the $\alpha \sim 1$ spectrum of the virial shock, as expected based on theory and on previous high-energy studies, and favoured in §6 based on forward modelling. The raw, steep, central signal can be approximated as multiple point sources stacked close to the centre of the cluster, so a strong, sharp ripple is expected to show $\alpha \approx 0.7$, substantially harder than the virial signal. In contrast, the cumulative emission from an extended central region can give rise to a more diffuse ripple with a much softer, $\alpha \sim 1.5$ spectrum. Therefore, mistaking a ripple for an $\alpha \gtrsim 1$ virial signal is unlikely. Such contamination cannot, however, be ruled out entirely based on spectral arguments alone, as the inferred spectrum is sensitive to the precise handling of point sources. For instance, removing one cluster with a bright NVSS source in its centre changes the raw spectrum in Fig. 9 from 0.8±0.1 to 1.0±0.2, which is similar to the expected virial signal. Naively, the ripples should remain softer than $\alpha = 1$, but making such a determination requires a more careful analysis.

The cleaned central signal is much weaker than the raw signal, its statistical errors are larger, and some spectral curvature emerges in backward modelling. Consequently, while the signal can still be fit by a power-law spectrum, the uncertainty in the resulting spectral index, ranging from $\alpha = 0.9 \pm 1.0$ at $\tau_{\text{max}} = 0.17$ to $\alpha = 1.4 \pm 0.7$ at $\tau_{\text{max}} = 1$, is substantial. Note that this mostly diffuse signal is quite similar to the expectation from stacked radio halos, in terms of its spatial and spectral distributions. The virial signal is even weaker, strongly dominated by statistical uncertainties, and the location of its peak radius is frequency-dependent, so a backward determination of the spectrum is not useful. The spectra of the two signals in the cleared data are addressed more carefully below, using forward modelling (see §5, §6.4 and §6).

### 5 CENTRAL EXCESS

Before studying the virial excess, it is important to separate out the central, $\tau \lesssim 1.5$ excess, to minimize its risk of biasing the measurement of the virial signal. We examine the central signal in both restored and cleared data. As in §4.2.4, the raw excess is studied using the restored data without masking NVSS sources, while the cleaned excess after point source removal is studied using the cleared data after masking NVSS sources.

The virial signal could in principle contribute, in projection and especially where the PSF is extended, to the central excess. As we show in §1, this contribution is only at the $\lesssim 5\%$ level for the raw data, and so can be neglected. For the cleared signal, this contribution can reach $\sim 20\%$, so the central and virial signals are simultaneously modelled in §6.1. For simplicity, and to $\sim 20\%$ accuracy, in this section, we neglect the virial contribution even for the cleared data.

We utilise the different appearances of the central signal
in restored and cleared data to better model the emission and its effect on the virial signal. Importantly, we examine if sidelobes of the central signal may be able to mimic the virial excess. In §5.1, we outline models for the central component in each data set. The fitting procedure, through $\chi^2$ minimisation, is described in §5.2. We outline the results in §5.3 and provide the technical details in §6.

5.1 Central excess models

We model the central emission as a combination of a point source (subscript ‘p’) and an extended (subscript ‘e’) component, with respective power-law spectral indices $\alpha_p$ and $\alpha_e$. The range of redshifts spanned by the clusters in our sample is small, so the specific flux projected along the line of sight in normalised coordinates $\tau = \{\tau_x, \tau_y\}$ can be approximately described by a simple one-dimensional model,

$$F_{cen} \equiv F_p \nu^{-\alpha_p} \delta(\tau) + A_e \nu^{-\alpha_e} \frac{\Theta(\tau_{cut} - \tau)}{(\tau^2 + \tau_{cut}^2)^{1/2}} , \quad (11)$$

where $\nu \equiv \nu/\nu_0$ is the normalised (to $\nu_0 = 73.152$~MHz) frequency, $\tau$ is the core radius, and $\Theta$ is the Heaviside step function. The normalizations $F_p$ and $A_e$, and the normalised core radius $\tau_c$ and cutoff radius $\tau_{cut}$, are for simplicity assumed to be identical for all clusters. As the typical core radius, $0.1 R_{200}$, is smaller than the FWHM of the LWA PSF even in the high-frequency channel, model results are not sensitive to the exact choice of $\tau_c$ around this value; hence, we henceforth use $\tau_c = 0.1$. For illustrative purposes, we define the total flux in the high-frequency channel, $F_{\nu, e} \equiv F_{\nu, p} + F_{\nu, e}$, where $F_{\nu, p}$ and $F_{\nu, e} \equiv 2 \pi \tau_c^2 \zeta^2 A_e/(2 - \zeta)$ are the specific fluxes of the point-like and extended components, respectively, integrated over $\tau$ at $\nu = \nu_0$.

Recall that bright sources in the LWA maps are deconvolved with the PSF, and the associated emission is either removed (for cleared data; Eastwood et al. 2018) or restored with a Gaussian restoring beam (for restored data; see typical FWHM in Table 1). When modelling the cleared data, Gaussian restoration is irrelevant, and we simply convolve the model with the PSF.

In contrast, in the restored data, which shows a much brighter central excess than in the cleared data, most ($\sim 75\%$; see §6) of this excess has been identified by CLEAN and restored. Indeed, the central signal is stronger and more sharply peaked in the restored data than it is in the cleared data, as seen from the stacked intensity (by a factor of $\sim 4$; Figs. 3 and 7) and significance (Figs. 4 and 8) profiles. Therefore, our model for the restored central excess (11) is convolved with the Gaussian restoring beam and then fit to the data. This model can then be convolved with the PSF, to constrain the maximal putative sidelobe artefacts.

5.2 Fitting procedure

For both restored and cleared data sets, we run Monte-Carlo simulations and minimize $\chi^2$ to determine the best-fitting parameters. The procedure involves sampling a grid of possible model parameter values, and testing each parameter set against $N_{\text{mock}}$ mock samples. Each mock sample contains 44 mock clusters, like the real sample, distributed randomly across the allowed regions of the sky as described in §3.4. For each mock cluster, we inject the model emission Eq. (11), convolve the mock sky image with either the PSF or a restoring Gaussian beam, as explained above, and pass the resulting mock image through our nominal stacking and analysis pipelines. We use $N_{\text{mock}} = 30$ such mock samples for each choice of parameter set, sufficient for convergence within $\lesssim 0.1\%$ in $\chi^2$.

Each control sample $j$ yields a significance distribution $M_j(\nu, \tau, c) = S(\nu, \tau, c)$, as described in §3.2. We average among the $N_{\text{mock}}$ samples of a given parameter set in order to estimate its model outcome, $M(\nu, \tau, c) \equiv N_{\text{mock}}^{-1} \sum_{j=1}^{N_{\text{mock}}} M_j(\nu, \tau, c)$. The $\chi^2$ value can then be computed, as usual, by summing over all relevant channels, radial bins, and clusters,

$$\chi^2 \equiv \sum_{\nu} \sum_{\tau} \sum_{c=1}^{N_c} [S(\nu, \tau, c) - M(\nu, \tau, c)]^2 , \quad (12)$$

as $S(\nu, \tau, c)$ is approximately a unit normal random variable. As in §4, we separately analyse the high-frequency channel and the combined low-frequency channels. The seven low-frequency channels are jointly fitted with the same parameters of the model. For the central excess modelling, we compute $\chi^2$ in the radial range $0 \leq \tau \leq 1.5$ (and verify that the results are not sensitive to exact choice of outer radius).

A simple alternative to Eq. (12) is obtained by choosing the normalised stacked signal $S(\nu, \tau, c)$ as the unit normal random variable, instead of its individual cluster counterpart $S(\nu, \tau, c)$. This choice leads to

$$\chi^2 = \sum_{\nu} \sum_{\tau} [S(\nu, \tau) - M(\nu, \tau)]^2 , \quad (13)$$

reducing the number $N$ of degrees of freedom (DOF) by a factor $\sim N_c$. Here, we again averaged over mock samples, $M(\nu, \tau, c) \equiv N_{\text{mock}}^{-1} \sum_{j=1}^{N_{\text{mock}}} M_j(\nu, \tau, c)$, where $M_j(\nu, \tau) = S(\nu, \tau, c)$ is obtained by stacking over clusters, as defined in §3.2, in either beam-weighted or cluster-weighted methods. This prescription aims to better fit the stacked signals (shown in Figs. 4 and 8, typically giving a lower $\chi^2$ per DOF), at the expense of the sensitivity to the mass-dependence of individual cluster models. The method is more conservative than Eq. (12), typically providing similar parameter estimates at lower confidence levels, so we primarily use Eq. (13). Quantitatively, Eqs. (12) and (13) typically yield $\lesssim 20\%$ differences between their best fit parameter estimates. The values of $\zeta$ and $\tau_{cut}$ change slightly more, by $\sim 30\%$, but these variations typically lie within the $1\sigma$ uncertainty of these parameters. A larger and more significant difference is obtained for the normalisation $m \xi_S \xi_B^{(2+\nu)/4}$ of the virial signal. Here, using Eq. (13) lowers the best fit by $\sim (20–55)\%$; see discussion in §6.2.

The frequency summations in Eqs. (12) and (13) do not take into account correlations between channels, and so tend to overestimate the $\chi^2$ expected for co-adding independent low-frequency channels. Indeed, these equations reasonably identify the best-fitting parameters, but the correlations can lead to unreasonably small parameter uncertainties. We may crudely correct for correlations by choosing the corrected $S(\tau)$, defined in Eq. (10), as the random variable. Along with the corresponding, corrected model $M(\tau) = N_{\nu}^{-1+(\nu)/2} \sum_{\nu=1}^{N_{\nu}} M(\nu, \tau)$, we then obtain

$$\chi^2 = \sum_{\tau} [S(\tau) - M(\tau)]^2 , \quad (14)$$
further reducing $N$ by a factor of $\sim N_v$. Equation (14) cannot be used to model the spectrum, and so is mainly invoked below only to gauge the correlations.

For model parameters other than the spectrum, we can compare the best fits obtained from Eqs. (13) and (14). While the best-fit values agree, the former equation leads to consistently and predictably underestimated uncertainties, due to channel correlations. The results become fully consistent with each other if one multiplies the $\chi^2$ of Eq. (13) by a constant correction factor $C$. When co-adding cleared data in all low-frequency channels, this correlation yields consistent results for all model parameters, provided that $C \approx 0.25$ (see §6).

An alternative way to correct for correlations is by generalising Eq. (10) to all model parameters. Namely, assuming that the uncertainty in any model parameter $a$ should satisfy $\sigma_{a,N} = N^{(\psi-1)/2} \sigma_a$. Here, $\sigma_a (\sigma_{a,N})$ denotes the uncertainty obtained from one channel ($N$ channels), and $\psi$ was calibrated in §3.2. We find that Eq. (13) yields a similar relation, $\sigma_{a,N} = N^{(\psi-1)/2} \sigma_a$, but with a different correlation parameter $\Psi$ affected by the numerical procedure, found to be $\Psi \approx 0.45$ for the restored data and $\Psi \approx 0.01$ for the cleared data. Hence, when deriving the uncertainty $\sigma_{a,N}$ in $a$ from Eq. (13) with summation over $N$ channels, we correct the result as

$$\sigma_{a,N} = N^{(\psi-\Psi)/2} \sigma_{a,N}^* ,$$

and again recover a good agreement with the outcome of Eq. (14) (see demonstrations in §E). Anticipating parabolic $\chi^2$ minima near the best-fit values, we find that $C \approx N^{(\Psi-\psi)/2} \approx 0.26$ for the cleared data, consistent with the independent finding above, and $C \approx 0.46$ for the restored data.

Unless otherwise stated, we henceforth compute $\chi^2$ using Eq. (13) and correct the co-added results for the low-frequency channels using Eq. (15). Other methods are occasionally demonstrated to yield similar results. We verify that the results are not sensitive to the choice of $N_{mock}$. For brevity, in the following text, we sometimes quote $\chi^2$ (or, when relevant, $C\chi^2$) values only for beam co-addition, but provide both beam (denoted ‘bm’) and cluster (‘cl’) values in the tables and most figures.

### 5.3 Central modelling results

For the raw data, the Gaussian-convolved central model, including both point-like and extended components, provides a good fit in both high-frequency ($\chi^2$ per DOF value of $\chi^2 = 0.8$) and co-added low-frequency ($C\chi^2 = 0.6$) channels, but the large parameter uncertainties suggest that the model is oversimplified. Here, the spectral indices of the two components typically fall in the range $0.7 \lesssim \alpha \lesssim 1.3$, but the point-like component softens considerably in the higher-frequency channels. The spectral indices of the two components are then presented in §6.5. Finally, the magnetic field downstream of the shock is estimated in §6.6.

### 6 VIRIAL EXCESS SIGNAL

With the central signal quantified in §5, we now turn to the virial excess, focusing on the cleared data, where much of the point-source contamination has been removed. We test if the observed signal agrees with simple models for the synchrotron emission from the virial shock, which are briefly provided in §6.1 based on a derivation in §I. To this end, we combine linear regression with TS tests, as described in §6.2. The results are then presented in §6.3 for the special case of a flat injection spectrum, $p = 2$, giving rise to a cooled photon spectrum $\alpha_{vir} \approx 1$ from the virial shock. The results are then generalised for an arbitrary spectrum in §6.4. We examine the dependence of the virial excess signal upon the cluster mass in §6.5. Finally, the magnetic field downstream of the shock is estimated in §6.6.

#### 6.1 Virial shock synchrotron model

Our virial shock modelling assumes that the shock deposits a fraction $\xi_\text{E}$ of the downstream thermal energy in CREs, and a fraction $\xi_B$ in magnetic fields. While CRE cooling is dominated by Compton losses off the CMB for the relevant, $B < 1 \mu G$ fields, a fraction $\sim B^2/(8\pi u_{cmb})$ of the energy is radiated as synchrotron emission, where $u_{cmb}$ is the CMB energy density. The CREs cool rapidly, so their radiative signature is expected to be well-localised near the virial shock, reflecting the temporal and spatial changes in energy injection rate across the shock. Therefore, while we model the emission from a virial shock, for simplicity, by invoking a stationary flow in spherical symmetry, the actual shock morphology can be highly irregular, and the energy injection rate through the shock can show strong variations with position and time.

The co-addition procedures we apply to the LWA data effectively average out these spatial and temporal variations, preferentially picking up the component of the virial shock signal that is most conducive to our stacking and circular binning. The stacking and analysis procedures can also pick up foreground and background contaminations, such as the changes in background due to the quenching of star-forming galaxies that crossed inside the virial shock. We consider two
different spatial models for the virial-shock component of the stacked radio signal: emission from a thin spherical shell at a normalised radius $\tau_s$ (henceforth the shell model), and emission from a thin ring of the same radius $\tau_s$ on the plane of the sky (henceforth the planar model).

For concreteness, and following RK18, we adopt an isothermal $\beta$-model to describe the gas distribution in a cluster, and assume hydrostatic equilibrium. The anticipated signal, derived for both shell and planar models in §I, then follows the scaling

$$
\nu I_0 \propto \tilde{m}_\xi \xi_B \frac{2+p}{2} \tau_s \left( \frac{2}{p+4} \right) \times \frac{M_{500}}{(1+z)^2} \times \nu^{\frac{p-2}{2}} f(\tau/\tau_s),
$$

where we defined $M_{14} \equiv M_{500}/10^{14} M_\odot$ and the normalised angular dependence $f$ of the relevant (shell or planar) model. Given the cluster mass and redshift, the free parameters of the model are the normalised shock radius $\tau_s$, the CRE injection spectral index $p$, and the product $\tilde{m}_\xi \xi_B \left( \frac{2+p}{4} \right)$, which weighs the gas accretion rate, the CRE acceleration efficiency, and the magnetisation efficiency. The proportionality constant in Eq. (16) depends on $p$, the cosmology, and the range $(\gamma_1, \gamma_2)$ of CRE Lorentz factors [see Eq. (120)]. In the anticipated limit of a strong shock, $p = 2$, so $\nu I_0$ becomes frequency independent, the free parameters reduce to $\tau_s$ and $\tilde{m}_\xi \xi_B$, and the proportionality constant depends on the CRE energy range only through the combination $\ln(\gamma_2/\gamma_1)$ [see Eq. (121)].

### 6.2 Fitting procedure

To determine the best-fitting parameters and their uncertainties, we apply the same procedures here as in §5.2. Namely, for each model and set of parameters, we average over multiple Monte-Carlo simulations of control samples incorporating the injected model, compare the outcome to the data, and minimize the resulting $\chi^2$. Again, the analysis is performed separately for the high-frequency channel and for the co-added seven low-frequency channels.

The $\chi^2$ minimisation is equivalent to maximising the likelihood $L$, given by

$$
\ln L = -\frac{1}{2} \chi^2.
$$

For the co-added low-frequency channels, we use the corrected value $\tilde{C}\chi^2$ instead of $\chi^2$, as defined and tested in §5.2, to crudely correct for inter-channel correlations. Again, this has no effect on the estimated parameters, and only corrects their estimated uncertainty intervals.

To quantify the significance of the virial excess, we use the test statistic (e.g., Mattox et al. 1996)

$$
\text{TS} \equiv -2 \ln \frac{L}{L_+} = \chi^2 - \chi^+_2,
$$

where subscript ‘-’ (‘+’) refers to the model without (with) the virial shock. After maximising the likelihood of each model over its free parameters, the resulting TS is assumed to approximately have a $\chi^2$ distribution with $N = N_+ - N_- \approx N_-$ degrees of freedom, corresponding to the extra parameters added by the virial shock (Wilks 1938). The implied significance of the excess is denoted $\sigma_{\text{TS}}$.

Beam and cluster co-additions give similar results for the virial excess; for brevity, the following text focuses mainly on beam co-addition. In particular, quoted values of TS, $\chi^2$, and, when relevant, also $\tilde{C}\chi^2$, refer to beam co-addition unless otherwise stated. Both beam-weighted and cluster-weighted results are provided in most figures and in the supplementary material.

Figures 4 and 8 show that the virial excess is localised in the $2 \lesssim \tau \lesssim 3$ bins in the high-frequency data, and in the $2 \lesssim \tau \lesssim 3.5$ bins in the low-frequency data. Therefore, our nominal $\chi^2$ analysis focuses on the radial range $0 \lesssim \tau \lesssim \tau_{\text{fit}}$ with $\tau_{\text{fit}} = 5$, including both central and virial signals. We verify that the results are not sensitive to reasonable variations in $\tau_{\text{fit}}$ (ranging from 4 to at least 7.5).

The parameters inferred for the central and virial models are not entirely independent of each other. In §5, we modelled the central source parameters while focusing on small radii and neglecting the virial contribution. In §6.3 below, we simultaneously model the virial excess and the main parameters of the central signal. Comparing the results indicates that the virial excess has a small but noticeable, $\lesssim 15\%$ effect on the inferred parameters of the central model. The opposite effect—a putative bias in the virial model parameter estimates due to contamination by the central component—leads to a somewhat larger effect, with $\lesssim 25\%$ changes in best-fitting parameters. Such effects are discussed in §7.

In addition to the statistical errors associated with the real and control samples, propagated into the model-parameter estimates, there are also systematic effects associated with the model assumptions, the properties of the cluster catalogue, etc. Based on the study of different model and analysis variants, we find that the systematic uncertainties are typically no larger than the statistical errors. An exception is the radio signal normalisation, $\tilde{m}_\xi \xi_B \left( \frac{2+p}{4} \right)$, where the systematic error may be as large as a factor of $\sim 4$.

While replacing our nominal $\chi^2$ estimate in Eq. (13) by the more standard, i.e. non-stacked Eq. (12), does not significantly change the best-fitting values of most parameters, it does tend to lower the best-fitting normalisation $\tilde{m}_\xi \xi_B \left( \frac{2+p}{4} \right)$ of the virial signal by $\sim 20$–55%. Hence, we present results based on both equations in the following sections. The difference in the best-fitting $\tilde{m}_\xi \xi_B \left( \frac{2+p}{4} \right)$ can be attributed at least in part to the dispersion of excess significance among individual clusters, and the partial correlation between this excess and cluster mass.

### 6.3 Results for a strong virial shock ($p = 2$)

We begin with the simple case where the shock is strong. Here, CREs are injected with a flat, even non-stacked Eq. (12), it does not significantly change the best-fitting values of most parameters, it does tend to lower the best-fitting normalisation $\tilde{m}_\xi \xi_B \left( \frac{2+p}{4} \right)$ of the virial signal by $\sim 20$–55%. Hence, we present results based on both equations in the following sections. The difference in the best-fitting $\tilde{m}_\xi \xi_B \left( \frac{2+p}{4} \right)$ can be attributed at least in part to the dispersion of excess significance among individual clusters, and the partial correlation between this excess and cluster mass.
for the high-frequency (top and middle rows) and co-added low-frequency (bottom row) channels. Confidence contours in such figures are (unless otherwise stated) based on fitting $\chi^2$ as a fourth-order polynomial in the two free parameters, for better visibility. Models are shown both in the absence (dashed significance profiles) and in the presence (other dashings) of the virial shock, in both planar (dot-dashed) and shell (double dot-dashed) model variants. Table 2 demonstrates the best-fitting parameters and the implied TS significance levels of the virial shock, for different model variants (more details and models are provided in Table D1). In the low-frequency co-addition results shown in the table, parameter uncertainties are corrected using Eq. (15) and the TS values are based on $C\chi^2$ (where $C \approx 0.25$; see §5.2).

We discuss the results for the high-frequency channel in §6.3.1 and for the combined low-frequency channels in §6.3.2, obtained in both cases by simultaneously fitting the central and virial signals with $p = 2$.

6.3.1 High-frequency channel

The high-frequency models are shown in the top and middle rows of Fig. 10. The top row uses the nominal data, featuring in particular a $\sim 2\sigma$ excess around $\tau \approx 3.5$, shown above (see Fig. 4) to arise from two clusters with extended NVSS point sources. To reduce the sensitivity of the model to these and other structures outside the virial radius, the middle row of Fig. 10 shows high-frequency channel results when limiting the fitting radial range to $\tau_{\text{fit}} = 3$.

The nominal planar model (orange dot-dashed curves) fits the virial shock signal well, with $\chi^2_0(\tau \leq 5) \approx 1.0$, and is not sensitive to the inclusion (top row) or exclusion (middle row) of $\tau > 3$ data. The best-fitting parameters are (for $\tau_{\text{fit}} = 5$,

![Figure 10](https://example.com/f10.png)

**Figure 10.** Modelling both central and virial signals in the high-frequency (top row for $\tau_{\text{fit}} = 5$ and middle row for $\tau_{\text{fit}} = 3$) and co-added low-frequency (bottom, $\tau_{\text{fit}} = 5$) cleared data channels. The left panels show the significance profiles of LWA data (solid blue curve) and the best fit planar (dot-dashed orange), shell (double dot-dashed green), and no shock (dashed magenta) models. The control sample curves and vertical shock range are the same as in Fig. 4. The middle and right panels show the best-fit parameters (symbols) and the $1\sigma-3\sigma$ confidence level contours (same curves styles as in the left panels) for nominal analysis with beam (disks and thicker, short dashed contours) and cluster (squares and thinner, long dashed contours) co-addition. Switching to the standard $\chi^2$ of Eq. (12) (blue triangles with dashed or dotted confidence levels) tends to lower $m\xi L B^{(2+p)/4}$ and increase the confidence level; see §6.2. The contours in the top row are biased by the $\sim 2\sigma$ contamination around $\tau \approx 3.5$, but not when using the standard $\chi^2$ (see text for details).
Table 2. Best fit for joint central and virial shock nominal (cleared data; beam co-addition) modelling.

| Model | Channel | $F_\gamma$ | $\alpha_\xi$ | $\tau_\xi$ | $\dot{m}_\xi\xi B$ (10$^{-4}$) | $p$ | $\chi^2$ | $N$ | TS (σ) |
|-------|---------|-----------|-------------|----------|-----------------|-----|--------|-----|--------|
| planar | high    | 1.63$^{+0.23}_{-0.20}$ | — | 2.45$^{+0.11}_{-0.09}$ | 2.85$^{+0.66}_{-0.58}$ | 2 | 17.2 | 15 | 18.5 (3.9σ) |
| low   | 2.01$^{+0.22}_{-0.20}$ | 1.20$^{+0.27}_{-0.26}$ | 2.63$^{+0.09}_{-0.10}$ | 1.95$^{+0.36}_{-0.35}$ | 2 | 137.1 | 134 | 30.5 (5.2σ) |
| shell | high†  | 1.65$^{+0.33}_{-0.23}$ | — | 2.79$^{+0.31}_{-0.23}$ | 2.29$^{+1.17}_{-0.95}$ | 2 | 132.7 | 7 | 7.6 (2.3σ) |
| low   | 1.95$^{+0.25}_{-0.23}$ | 1.14$^{+0.28}_{-0.26}$ | 2.99$^{+0.13}_{-0.12}$ | 2.31$^{+0.49}_{-0.47}$ | 2 | 162.0 | 134 | 24.0 (4.5σ) |

Columns: (1) Shock model; (2) The high-frequency (high) or the seven co-added low-frequency (low) channels; (3) Flux of the central emission normalised to the 73MHz frequency, in units of 10$^{-23}$ erg s$^{-1}$ cm$^{-2}$ Hz$^{-1}$; (4) Spectral index of the extended central emission; (5) Shock radius normalised to $R_{500}$; (6) Virial shock normalisation, in 10$^{-4}$ units; (7) Injected spectral index of CREs; (8) $\chi^2$ values of the fit (without the C correction; see §5.2); (9) Number of degrees of freedom; (10) TS value (equivalent significance in parenthesis).

See text (§6.3) for technical details. Values in boldface are fixed rather than fitted.

† — Using $\tau_{fit} = 3$ to avoid a $\tau \approx 3.5$ fluctuation; see §6.3.1.

unless otherwise stated)

$\tau_\xi = 2.45^{+0.11}_{-0.09} (2.44^{+0.11}_{-0.09}) \quad (19)$

and

$\dot{m}_\xi\xi B = 2.85^{+0.66}_{-0.65} (2.79^{+0.70}_{-0.69}) \times 10^{-4} \quad (20)$

for beam (cluster; henceforth omitted) co-addition. The best-fitting shock position coincides with that inferred by RK18 based on the Fermi γ-ray signal. We find that $\tau_\xi \simeq 18.5$, corresponding to a 3.9σ detection of the virial shock for $N = 2$ degrees of freedom. If we were to use the $\tau_\xi = 2.4$ of RK18 as a prior, the significance would increase to $\tau_\xi \simeq 20.3$ (4.3σ for $N = 1$).

The above results are based on the nominal $\chi^2$ of Eq. (13). If we use, instead, the more standard, non-stacked least-squares of Eq. (12), the best-fitting shock position $\tau_\xi = 2.40^{+0.08}_{-0.06}$ is practically unchanged, but $\dot{m}_\xi\xi B = 1.48^{+0.27}_{-0.24} \times 10^{-4}$ is nearly halved (blue dashed contours), and the significance of the virial shock increases to $\tau_\xi \simeq 33.2$ (5.4σ for $N = 2$, or 5.8σ for $N = 1$ with the RK18 prior). Figure 11 presents the radial significance profiles based on Eq. (12). Also shown is the TS-based significance profile (thick red solid curve) corresponding to the nominal planar model, as a function of $\tau_\xi$.

Like the planar model, the shell model is not sensitive to the choice of $\tau_{fit}$, but only when using the standard least-squares Eq. (12). In this case, we find both $\tau_\xi = 2.70^{+0.15}_{-0.13}$ and $\dot{m}_\xi\xi B = 1.85^{+0.31}_{-0.30} \times 10^{-4}$ somewhat larger than their planar case counterparts in this method, with a high-significance, $\chi^2_0(\tau \leq 5) \simeq 1.2$ and TS $\simeq 41.1$ (6.1σ for $N = 2$; no priors used henceforth) detection of the virial shock. The nominal Eq. (13) is more sensitive to confusion with the $\tau \sim 3.5$ artefact, so using $\tau_{fit} \gtrsim 5$ leads to an exaggerated $\tau_\xi \simeq 4$. For $\tau_{fit} = 3$, the best fit is $\tau_\xi = 2.83^{+0.26}_{-0.22}$ and $\dot{m}_\xi\xi B = 2.55^{+0.88}_{-0.81}$ consistent with the above but with a weaker, $\chi^2_0(\tau \leq 3) \simeq 1.9$ and TS $\simeq 7.6$ (2.3σ for $N = 2$) detection due to the smaller radial range used.

We conclude that the virial shock signal is identified at a high confidence level in the high-frequency channel. Both planar and shell models give good fits to the data, and with similar normalisations $\dot{m}_\xi\xi B$. While the two models reproduce the same peak locations in flux, in the shell model this peak radius is smaller than the shock radius, due to the combined effect of projection and PSF convolution. Note that in the high-frequency channel, neither model reproduces the deep local minimum at $\tau \simeq 1.5$.

6.3.2 Low-frequency channels

The low-frequency model is demonstrated in the bottom row of Fig. 10 and in Table 2 (and in more detail in Table D1). The fitted models show similar behaviour to that found in the high-frequency channel, with comparable parameters and a similarly high-confidence detection of the virial shock.

The nominal planar model gives

$\tau_\xi = 2.63^{+0.09}_{-0.10} (2.62^{+0.10}_{-0.10}) \quad (21)$

and

$\dot{m}_\xi\xi B = 1.95^{+0.36}_{-0.35} (1.83^{+0.36}_{-0.36}) \times 10^{-4} \quad (22)$

for beam (cluster, henceforth omitted) co-addition. The $\chi^2_0(\tau \leq 5) \simeq 0.3$ fit yields TS $\simeq 30.5$ (5.1σ for $N = 2$). Again, replacing Eq. (13) by the standard Eq. (12) leaves $\tau_\xi$ unchanged, halves $\dot{m}_\xi\xi B = 0.84^{+0.15}_{-0.13} \times 10^{-4}$, and increases the significance to TS $\simeq 37.5$ (5.8σ for $N = 2$).

The shell model yield $\tau_\xi = 2.99^{+0.13}_{-0.11}$ and $\dot{m}_\xi\xi B = 2.31^{+0.49}_{-0.47} \times 10^{-4}$ with $\chi^2_0 \simeq 0.3$ and TS $\simeq 24.0$ (4.5σ for $N = 2$).

Figure 11. Modelling the high frequency data as in Fig. 10 (with the same notations), but using the least-squares estimate (12). The TS-based significance profile (thick solid red curve) of the nominal planar model is shown as a function of $\tau_\xi$. 

MNRA 000, 1–77 (2022)
2) when using the nominal Eq. (13). Like the high-frequency case, the shell model gives a somewhat larger $\tau_s$ compared to its planar counterpart. Using the standard Eq. (12) instead of Eq. (13) gives a slightly larger $\tau_s = 3.14^{+0.06}_{-0.09}$, a $\sim 30\%$ smaller $m_\xi = 1.58^{+0.19}_{-0.28} \times 10^{-4}$, and substantially raises the significance to $TS \approx 70.0$ (8.1σ for $N = 2$).

We conclude that the low-frequency channels support the high-frequency behaviour, again indicating the presence of the virial shock at a very high confidence level, and with parameter estimates and confidence levels consistent with their high-frequency counterparts. Co-adding the two virial signals, namely in the high-frequency channel and in the co-added low-frequency channels, with equal weights, the significance estimated by the standard Eq. (12) indicates a detection of the virial shock at confidence levels reaching $TS \approx 69.2$ (8.0σ for $N = 2$) for the planar model and $TS \approx 106.3$ (10.0σ for $N = 2$) for the shell model.

### 6.4 Results for virial shock with arbitrary $p$

We now relax the strong shock assumption, modelling the virial signal with $p$ as a free parameter rather than fixed to its strong-shock, $p = 2$ limiting value. For this purpose, we use the cleared data in all low-frequency channels, as in §6.3.2; the results are again summarised in Table 2 and in Table D1. Due to the significant uncertainties surrounding the virial shock emission, the spectrum estimated through forward modelling is more accurate than the direct estimate in §4.2.4.

To simplify the analysis, the central emission parameters are fixed to their best-fitting values inferred for the respective $p = 2$ case, leaving only the virial shock signal parameters $\tau_s$, and the generalised normalisation $m_\xi = (2p+4)/4$. The best-fitting shock radius again yields $\tau_s = 2.62^{+0.10}_{-0.09} (2.62^{+0.11}_{-0.09})$ for the planar (shell) model with nominal $\chi^2$, consistent with the results in §6.3.2, so we focus on the relation between the remaining two parameters, $p$ and $m_\xi$.

Figure 12 shows the confidence level contours of these parameters, for the planar (top panel) and shell (bottom) model. Here, contours are produced by interpolating and smoothing the $\chi^2$ data, instead of fitting it to a polynomial, as the latter would distort the contours in this particular case.

We find the spectral index $p = 2.02^{+0.17}_{-0.16} (2.04^{+0.12}_{-0.09})$,

\begin{equation}
(23)
\end{equation}

for the planar (shell) model, with $TS \approx 29.7$ (4.8σ for $N = 3$) for the planar model and $TS \approx 23.8$ (4.2σ for $N = 3$) for the shell model. The best-fit spectral index is consistent with the strong shock limit $p = 2$, justifying the assumption of a flat, $dN/dE \propto E^{-2}$ injected CRE spectrum. Consequently, the associated normalisation $m_\xi B = 2.27^{+0.72}_{-0.57} \times 10^{-4}$ ($2.73^{+0.63}_{-0.59} \times 10^{-4}$) for the planar (shell) model agrees with the findings in §6.3.2. By adopting the standard $\chi^2$ of Eq. (12) instead of the nominal Eq. (13), the best-fitting result for $p$ is again consistent with 2.0, and the other parameters and confidence levels are unchanged with respect to the $p = 2$ results.

The flat spectrum of the virial excess is robust, showing no sensitivity to the co-addition method or to cluster selection, nor any evidence for spectral curvature. For example, for the planar (shell) model, the four lowest frequency, 37–52 MHz channels yield $p = 2.04^{+0.15}_{-0.10}$, while the three higher frequency, 57–68 MHz channels give a similar $p = 2.04^{+0.29}_{-0.21}$, i.e. 70.2σ ($p = 2$) for $N = 3$). This robustness contrasts with the central-signal spectrum in our modelling, which strongly depends on assumptions, co-addition method, choice of channels, etc. While this sensitivity may reflect our inaccurate modelling of the cluster centre, it may well arise from a spatially-dependent spectral curvature of the combined point-like and extended components of the central excess. In such a case, the robustness of the virial excess spectrum provides further evidence that it does not arise from putative sidelobe artefacts.

### 6.5 Mass dependence

The synchrotron brightness of the virial shock is expected to be a strong, $I_v \propto M^{2/3}$ function of the cluster mass; see Eq. (16). We test whether massive clusters indeed show a stronger virial excess signal by splitting our sample into two sub-samples according to cluster mass. Choosing $M_{500} = 5 \times 10^{14} M_\odot$ as the approximate median gives a sub-sample of 23

---

**Figure 12.** Best fit (symbols and 1σ to 3σ confidence contours) for virial shock parameters $m_\xi B = (2p+4)/4$ and $p$, for the planar (top panel) and shell (bottom) models, using low-frequency cleared data. Each panel shows the contours and best-fitting parameter for nominal analysis with beam and cluster co-addition (same contour and symbol notations as in Fig. 10). The strong shock limit $p = 2$ is also shown (dashed line).
more massive clusters with a median $M_{500} \simeq 1.0 \times 10^{14} M_\odot$, and a sub-sample of 21 less massive clusters with a median $M_{500} \simeq 3.3 \times 10^{13} M_\odot$. We stack the two sub-samples in the same method as in §3.2, using nominal parameters. Figure 13 shows the significance of the excess brightness of the two sub-samples, using both high (top panel) and low (bottom) frequency cleared data with beam co-addition.

![Figure 13. Significance of excess brightness profiles in high-frequency (top panel) and low-frequency (bottom) cleared data, for the 23 high mass ($M_{500} > 5 \times 10^{13} M_\odot$; solid blue curves) and 21 low mass ($M_{500} < 5 \times 10^{13} M_\odot$; dashed orange) clusters. Other notations are as in Fig. 4.](image)

In both high and low frequencies, we find a significant, ~4σ excess in the stacked massive clusters, but only a ~2σ excess in the stacked low mass clusters, in qualitative agreement with the expected mass dependence. The high-mass excess is of comparable significance to that of the full sample, despite having only half the clusters. Due to the smaller size of the sub-samples, some spurious signals (such as a ~2.5σ excess around $\tau \simeq 3.5$ in the high-frequency channel) become stronger than in the full sample.

The estimated virial excess brightness per cluster in the high-frequency channel (stacked low-frequency channels) is $\Delta I \simeq 6.7 \times 10^{-12}$ erg s$^{-1}$ cm$^{-2}$ sr$^{-1}$ (3.6 $\times 10^{-12}$ erg s$^{-1}$ cm$^{-2}$ sr$^{-1}$) in the massive-cluster sub-sample, which is higher by a factor of 1.3 with respect to the full sample, and by a factor of ~2.0 (1.5) with respect to the low-mass sub-sample. These results are again consistent with a strong mass-dependence of the signal. For the low-mass sub-sample, the significance of the signal is too low for a forward modelling that could reliably test the $M^{5/3}$ dependence.

Fitting the central and virial models to the massive sub-sample alone yields results consistent with those derived for the full sample. The inferred spectral indices and shock position do not change appreciably, while the central flux $F_\gamma$ increases by $\lesssim 75\%$ ($\lesssim 45\%$) for the restored (clear) data of the massive sub-sample. The best-fitting virial shock normalisation $\dot{m}\xi_e\xi_B^{(2+p)/4}$ in the massive sub-sample is the same as in the full sample when using the standard $\chi^2$ of Eq. (12), but smaller by ~30% (~20%) for the planar (shell) model when adopting the nominal $\chi^2$ of Eq. (13), as expected from the dispersion among clusters and the partial correlation with mass. In conclusion, the results are consistent with massive clusters dominating the signal, but a larger sample is needed to quantify the mass dependence of the model parameters.

### 6.6 Magnetic field estimate

Synchrotron emission from the virial shock provides a new probe of the magnetic field in galaxy cluster peripheries. To estimate these magnetic fields, we need to first separate $\xi_B$ from the measured quantity $\dot{m}\xi_e\xi_B^{(2+p)/4}$. One method is to adopt the $\dot{m}\xi_e \simeq 0.6\%$ value inferred (RK18) by modelling the γ-ray emission from the virial shock. A second method is to estimate the accretion parameter $\dot{m}$ from a model for the gas distribution, and then break the degeneracy between $\xi_e$ and $\xi_B$ by assuming equipartition, i.e. $\xi_e = \xi_B$ (see §K).

In either method, a $\xi_B$ estimate gauges the magnetic field $B$ given a model for the gas distribution. We invoke a rudimentary $\beta = 2/3$ model for all 44 clusters in our sample, but focus on nine clusters for which an isothermal $\beta$-model was derived by Fukazawa et al. (2004). Table 3 provides the $\dot{m}$, $\xi_B$ and $B$ estimates for these seven clusters, obtained in both of the above methods. The $\beta$-model and virial shock parameters are also provided in the table. The table focuses on the nominal analysis — the planar model for the high-frequency, cleared data, with the $\tau_s$ of Eq. (19) and $\dot{m}\xi_e\xi_B^{(2+p)/4}$ of Eq. (20) — but the text provides a range of possible values covering all model variants.

In the first method, we adopt $\dot{m}\xi_e \simeq 0.6\%$, so our estimates of $\dot{m}\xi_e\xi_B^{(2+p)/4}$ translate into magnetisation efficiencies ranging from

$$\xi_B \simeq (1.6 \pm 0.2)\%$$

(24)

for the entire 44 cluster sample. Using these $\xi_B$ values to estimate the peripheral magnetic fields of the seven modelled clusters then gives, on average,

$$B \simeq 0.1$–$0.3 \mu G$$

(26)

(with smaller statistical uncertainties), where we assumed that density and temperature are both suppressed at $\tau_s$ by a factor $f_\beta = 1/3$ with respect to $\beta$-model estimates; see §K. For the rest of the clusters in the sample, we compute $B$ by assuming a $\beta = 2/3$ model and hydrostatic equilibrium (see §K). In this method, Eqs. (24) and (25) correspond to a mean $B \simeq 0.2$–$0.6 \mu G$ for all 44 clusters.

In the second method, by estimating $\dot{m}$ for the seven clusters and assuming equipartition between $\xi_e$ and $\xi_B$, we derive an average magnetisation efficiency $\xi_B$ in the range 0.8% to 1.8%, with statistical uncertainties $\lesssim 0.1\%$. Such efficiencies translate into $B \simeq 0.1$–$0.2 \mu G$, somewhat lower than in the first method, where $\xi_B$ was effectively found to be larger than $\xi_e$. 

MNRAS 000, 1–77 (2022)
Table 3. Magnetic field estimates for select galaxy clusters.

| Name      | β (°) | n₀ (10⁻¹⁵ cm⁻³) | rₑ (kpc) | kₜ₁T (keV) | z | R₅₀₀ (Mpc) | mξₑ (%) | ξₑ (%) | B (µG) | m (%) | ξ_B (%) | B' (µG) |
|-----------|-------|-----------------|----------|-------------|---|-------------|-----------|--------|--------|-------|--------|--------|
| A 376     | 0.57  | 5.83            | 119.1    | 4.69        | 0.05| 0.81        | 0.0       | 4.8 ± 1.1 | 0.24 ± 0.03 | 0.5 ± 0.1 | 1.0 ± 0.1 | 0.10 ± 0.00 |
| A 575     | 0.57  | 7.61            | 105.6    | 3.67        | 0.04| 0.83        | 0.0       | 7.6 ± 1.1 | 0.24 ± 0.03 | 0.9 ± 0.1 | 1.0 ± 0.1 | 0.10 ± 0.00 |
| A 779     | 0.34  | 10.35           | 11.5     | 1.76        | 0.02| 0.45        | 0.0       | 10.3 ± 1.1 | 0.24 ± 0.03 | 2.5 ± 0.2 | 0.6 ± 0.1 | 0.08 ± 0.00 |
| A 1795    | 0.63  | 29.55           | 86.4     | 5.76        | 0.06| 1.22        | 0.0       | 29.6 ± 1.1 | 0.26 ± 0.03 | 0.7 ± 0.0 | 1.1 ± 0.1 | 0.12 ± 0.01 |
| A 2065    | 0.57  | 12.11           | 113.6    | 5.15        | 0.07| 1.05        | 0.0       | 12.1 ± 1.1 | 0.31 ± 0.04 | 1.1 ± 0.1 | 0.9 ± 0.1 | 0.13 ± 0.01 |
| A 2256    | 0.93  | 5.05            | 417.6    | 6.45        | 0.06| 1.12        | 0.0       | 5.0 ± 1.1 | 0.23 ± 0.03 | 0.3 ± 0.0 | 1.7 ± 0.2 | 0.13 ± 0.01 |
| A 2634    | 0.38  | 4.11            | 78.1     | 3.45        | 0.03| 0.75        | 0.0       | 4.1 ± 1.1 | 0.35 ± 0.04 | 3.0 ± 0.2 | 0.5 ± 0.0 | 0.11 ± 0.01 |
| HCG 51    | 0.34  | 4.71            | 32.8     | 1.41        | 0.03| 0.44        | 0.0       | 4.7 ± 1.1 | 0.24 ± 0.03 | 3.7 ± 0.2 | 0.5 ± 0.0 | 0.08 ± 0.00 |
| NGC 6329  | 0.38  | 10.61           | 13.9     | 1.64        | 0.03| 0.48        | 0.0       | 10.6 ± 1.1 | 0.19 ± 0.02 | 1.7 ± 0.1 | 0.7 ± 0.1 | 0.07 ± 0.00 |

Results shown for the seven cluster with β-models in Fukazawa et al. (2004) based on results (19) and (20), using either \( m\ξₑ = 0.6\% \) inferred from the γ-ray stacking (columns 8–10) or by assuming \( ξ = ξ_B \) equipartition (columns 11–13).

Columns: (1) Cluster name; (2) β-model index; (3) Central particle number density converted from the central electron number density in Fukazawa et al. (2004).; (4) Core radius; (5) Temperature; (6) Redshift; (7) \( R₅₀₀ \); (8) \( m\ξₑ = 0.6\% \) from RK18; (9) Magnetisation efficiency \( ξ_B \) based on \( m\ξₑ = 0.6\% \); (10) Magnetic field \( B \) based on \( ξ_B \); (11) Accretion parameter [using Eq. (K2)] with \( f_J = 1/3 \), used to extrapolate the β-model to the shock position (see [K]); (12) Magnetisation efficiency \( ξ_B' \) based on \( m \) and assuming \( ξ = ξ_B' \); (13) Magnetic field \( B' \) based on \( ξ_B' \).

\( ξₑ \). Here, invoking \( β = 2/3 \) and hydrostatic equilibrium for all clusters yields \( B \approx 0.2–0.3 \) µG.

7 SUMMARY AND DISCUSSION

We stack and radially bin the low-frequency, (36–73) MHz LWA data around 44 massive, high-latitude, extended MCXC clusters, after rescaling each cluster to its \( R₅₀₀ \) radius. Overall, the results indicate a high-significance excess emission peaked around \( τ = r/R₅₀₀ \sim 2.5 \), i.e. near the virial radius and coincident with a similarly-stacked γ-ray signal previously found in Fermi-LAT data and identified as inverse-Compton emission from viral shock-accelerated CREs (RK18). The radius, spectrum, and brightness of the LWA signal are consistent with the anticipated (Waxman & Loeb 2000; Keshet et al. 2003) synchrotron emission from the same viral shock CREs, as they gyrate in the post-shock magnetic fields. The signal is also consistent with radio to γ-ray (Keshet et al. 2017; Keshet & Reiss 2018) and SZ (Hurier et al. 2019; Keshet et al. 2020b) signals detected in individual clusters. We show that the results are robust to analysis details, in particular point source removal, foreground and background modelling, and sidelobe residuals from central point sources and diffuse emission.

7.1 Stacking

The clusters, with mass \( M₅₀₀ > 10^{14} M_\odot \) and angular radii \( 0°.2 < θ₅₀₀ < 0°.4 \) (constrained by resolution from below and foreground structures from above), are selected from low-frequency data, including the Galactic plane, Loop-I, and the Fermi bubbles with high, \( δ > 20° \) declinations to ensure a well-behaved LWA PSF (see §2). The high, 73 MHz frequency data have a substantially better angular resolution than in the seven lower-frequency, 36–68 MHz channels, so we analyse the high-frequency channel and the combined low-frequency channels separately. To better understand the data and the impact of point source cleaning, we study both the cleared data (after CLEAN) and the restored data (after point sources identified by CLEAN were restored with an effective, Gaussian beam).

The data are stacked among rescaled clusters, and binned radially as a function of \( τ \). The stacking is carried out in two different methods (§3), giving equal weights to either flux (beam co-addition) or clusters (cluster co-addition), in order to better resolve the origin of the excess. Monte-Carlo simulations of control cluster samples are used to calibrate the noise level in each channel, estimate the correlations between the co-added low-frequency channels, and quantify the origins of the signals (see §3.4 and §E). These control samples consist of random points, referred to as control clusters, in the relevant LWA sky, which undergo the same pipeline applied to the real MCXC sample.

The stacked data show a strong signal from the centres of clusters, and a high-significance signal from their peripheries, near the expected position of the virial shock. This virial excess is strong enough to be seen in the stacked image without radial binning, at least when four-folded (Fig. 5). At low, \( Δτ = 0.5 \) radial resolution, both high- and low-frequency results show a strong central emission at \( τ \lesssim 1 \), and a peripheral, \( ~5σ \) excess peaked in the \( 2.25 ≤ τ < 2.75 \) bin. In our nominal, higher, \( Δτ = 0.25 \) resolution, the excess peak shifts somewhat between channels; a \( ~4σ \) signal in the \( 2.25 ≤ τ < 2.5 \) bin for the high-frequency data (referred to as the nominal virial excess; see Fig. 4 and §4.1), and a \( ~3σ \) signal in the adjacent, \( 2.5 ≤ τ < 2.75 \) bin for the co-added low-frequency channels (Fig. 8 and §4.2). This small shift in peak position is likely caused by the lower spatial resolution of the low-frequency channels, broadening the excess in each cluster and thus shifting the weighted average outwards. Testing a finer, \( Δτ = 0.167 \) resolution indicates that the shift is indeed small, with peaks found at the \( 2.33 ≤ τ < 2.5 \) bin in the high-frequency channel and at \( 2.5 ≤ τ < 2.67 \) in the low-frequency data (see §H).

We use several independent methods to verify that the virial excess is not an artefact originating from beam-sidelobe residuals of the strong central emission missed by CLEAN (as outlined in §4.1.4). These tests rule out such an artefact, because: (i) a similar virial excess is found both in clusters that show a central signal and in clusters that do not (Fig. 6); (ii) convolving the best-fitting models for the raw central emission with the PSF yields sidelobe residuals at cluster peripheries which are insignificant or inconsistent with the
observed signal (Fig. F1); (iii) similar virial-ring radii are inferred in multiple channels, whereas the respective point-like and extended components vary substantially (§4.2); (iv) clusters with bright central NVSS sources can be included in, or excluded from, the sample, with a major change in stacked central signal but no significant effect on the virial excess (Fig. G2); (v) multiple attempts (see §G) to artificially mimic the virial ring by stacking non-cluster sources or bright regions in the sky, all fail to produce a significant signal near the virial radius (Fig. G1); (vi) The robust power-law spectrum of the virial excess is unlikely to emerge from point-like residuals of the raw central excess, which varies among clusters and with radius and is not a single power-law (compare §6.4 and §F1); and (vii) the virial excess agrees with predictions for the synchrotron emission from virial shock CREs, showing a radius consistent with the similarly-stacked γ-ray signal (e.g., Figs. 4 and H6) and a normalisation consistent with a previous Coma signal (Keshet et al. 2017), not far from equipartition (see discussion below).

The synchrotron emission from the virial shock should be a strong, \( L_\nu \propto M^{5/3} \) function of the cluster mass (see §I). By splitting the cluster sample into two sub-samples of comparable sizes according to mass, with \( M_{500} = 5 \times 10^{14} M_\odot \) as the approximate median, we find a \( \sim 4\sigma \) virial excess in the massive sub-sample and a weaker, \( \sim 2\sigma \) in the low-mass sub-sample. The massive sub-sample dominates the signal, with model parameters and a confidence level comparable to those of the full sample, in spite of having only half the number of clusters. These results qualitatively agree with the expected mass-dependent virial signal and support our virial shock model. Note that inverse-Compton emission is expected to show a weaker, linear dependence on \( M \); indeed, no significant correlation was found by RK18 between γ-ray emission and cluster mass. The small size of our sample precludes a more accurate determination of the signal dependence upon the mass (§6.5) or dynamical state (§H) of the cluster.

### 7.2 Modeling

We model the LWA data as a superposition of a smooth background, central emission, and a virial signal. Different models are inspected for each of these components: different polynomials for the background (see §C), different combinations of point-like and extended emission for the central source (§5.1), and either a projected sphere (shell model) or a ring/cylinder aligned parallel to the line of sight (planar model) for the virial excess (§6.1 and §I). Monte-Carlo simulations are used to inject different models into random parts of the LWA sky, and analyse the outcome with the same pipeline used for the real sample. Least-squares minimisation, applied to the individual or stacked signals, is used to determine the best-fitting parameters and their uncertainties, and TS-tests are used to estimate the confidence level of incorporating each component in the model. In our nominal analysis, all three components are modelled simultaneously. However, the central and virial components are sufficiently separated spatially from each other to be studied independently, with only a modest bias to the model normalisation of each component.

We model the central signal as a combination of point-like and extended components (see §5). In the restored data without point-source masking, the central excess is dominated by point-like emission with a spectral index \( \alpha_p = 1.04^{+0.05}_{-0.04} \). In the cleared data, the residual, diffuse and fainter signal is dominated by an extended component with index \( \alpha_e = 1.08^{+0.24}_{-0.16} \), presumably arising from stacked diffuse emission and point sources below the CLEAN threshold. When modelling the cleared-data central emission and the virial excess simultaneously, the normalisation of the central model decreases by \( \lesssim 15\% \), while the spectral index does not change appreciably (see §J).

We model the virial excess as synchrotron emission, assuming that the shock deposits a fraction \( \xi_\nu \) of the downstream thermal energy in CREs, and a fraction \( \xi_B \) in magnetic fields. The model has three free parameters: the normalised shock radius \( \tau_e \), the CRE energy spectral index \( p \), and the brightness normalisation \( \xi_B \). The latter combining the acceleration rate with the acceleration and magnetisation efficiencies. Least-squares minimisation and TS-tests then indicate a \( \sim 4\sigma \) detection for CREs, showing a radius consistent with the similarly-stacked γ-ray signal (e.g., Figs. 4 and H6) and a normalisation consistent with a previous Coma signal (Keshet et al. 2017), not far from equipartition (see discussion below).

The synchrotron emission from the virial shock should be a strong, \( L_\nu \propto M^{5/3} \) function of the cluster mass (see §I). By splitting the cluster sample into two sub-samples of comparable sizes according to mass, with \( M_{500} = 5 \times 10^{14} M_\odot \) as the approximate median, we find a \( \sim 4\sigma \) virial excess in the massive sub-sample and a weaker, \( \sim 2\sigma \) in the low-mass sub-sample. The massive sub-sample dominates the signal, with model parameters and a confidence level comparable to those of the full sample, in spite of having only half the number of clusters. These results qualitatively agree with the expected mass-dependent virial signal and support our virial shock model. Note that inverse-Compton emission is expected to show a weaker, linear dependence on \( M \); indeed, no significant correlation was found by RK18 between γ-ray emission and cluster mass. The small size of our sample precludes a more accurate determination of the signal dependence upon the mass (§6.5) or dynamical state (§H) of the cluster.

### 7.2 Modeling

We model the LWA data as a superposition of a smooth background, central emission, and a virial signal. Different models are inspected for each of these components: different polynomials for the background (see §C), different combinations of point-like and extended emission for the central source (§5.1), and either a projected sphere (shell model) or a ring/cylinder aligned parallel to the line of sight (planar model) for the virial excess (§6.1 and §I). Monte-Carlo simulations are used to inject different models into random parts of the LWA sky, and analyse the outcome with the same pipeline used for the real sample. Least-squares minimisation, applied to the individual or stacked signals, is used to determine the best-fitting parameters and their uncertainties, and TS-tests are used to estimate the confidence level of incorporating each component in the model. In our nominal analysis, all three components are modelled simultaneously. However, the central and virial components are sufficiently separated spatially from each other to be studied independently, with only a modest bias to the model normalisation of each component.

We model the central signal as a combination of point-like and extended components (see §5). In the restored data without point-source masking, the central excess is dominated by point-like emission with a spectral index \( \alpha_p = 1.04^{+0.05}_{-0.04} \). In the cleared data, the residual, diffuse and fainter signal is dominated by an extended component with index \( \alpha_e = 1.08^{+0.24}_{-0.16} \), presumably arising from stacked diffuse emission and point sources below the CLEAN threshold. When modelling the cleared-data central emission and the virial excess simultaneously, the normalisation of the central model decreases by \( \lesssim 15\% \), while the spectral index does not change appreciably (see §J).

We model the virial excess as synchrotron emission, assuming that the shock deposits a fraction \( \xi_\nu \) of the downstream thermal energy in CREs, and a fraction \( \xi_B \) in magnetic fields. The model has three free parameters: the normalised shock radius \( \tau_e \), the CRE energy spectral index \( p \), and the brightness normalisation \( \xi_B \). The latter combining the acceleration rate with the acceleration and magnetisation efficiencies. Least-squares minimisation and TS-tests then indicate a \( \sim 4\sigma \) detection for CREs, showing a radius consistent with the similarly-stacked γ-ray signal (e.g., Figs. 4 and H6) and a normalisation consistent with a previous Coma signal (Keshet et al. 2017), not far from equipartition (see discussion below).

The synchrotron emission from the virial shock should be a strong, \( L_\nu \propto M^{5/3} \) function of the cluster mass (see §I). By splitting the cluster sample into two sub-samples of comparable sizes according to mass, with \( M_{500} = 5 \times 10^{14} M_\odot \) as the approximate median, we find a \( \sim 4\sigma \) virial excess in the massive sub-sample and a weaker, \( \sim 2\sigma \) in the low-mass sub-sample. The massive sub-sample dominates the signal, with model parameters and a confidence level comparable to those of the full sample, in spite of having only half the number of clusters. These results qualitatively agree with the expected mass-dependent virial signal and support our virial shock model. Note that inverse-Compton emission is expected to show a weaker, linear dependence on \( M \); indeed, no significant correlation was found by RK18 between γ-ray emission and cluster mass. The small size of our sample precludes a more accurate determination of the signal dependence upon the mass (§6.5) or dynamical state (§H) of the cluster.
in $F_\gamma$ and $\lesssim 25\%$ variations in $\tilde{m}\xi\xi_{\nu}^{(2+p)/4}$ (see §1). However, the parameter $\tilde{m}\xi\xi_{\nu}^{(2+p)/4}$ incurs a substantial, factor $\sim 4$ systematic uncertainty, as it depends on model assumptions and on analysis details. This systematic uncertainty exceeds the factor $\sim 2$ uncertainty previously attributed to the $\gamma$-ray normalisation parameter $\tilde{m}\xi$, due to the additional dependence on the magnetic field. In particular, the shell model gives $\tilde{m}\xi\xi_{\nu}^{(2+p)/4}$ estimates $\sim 20\%$–$90\%$ higher than its planar counterpart at low frequencies, as it distributes the virial signal over a wider $\tau$ range; in the high frequency channel, the shell and planar estimates agree within $\sim 20\%$. Applying the $\chi^2$ analysis after, rather than before, co-adding clusters, raises the $\tilde{m}\xi\xi_{\nu}^{(2+p)/4}$ estimates by a factor $0.5$–$1.3$, probably due to the dispersion of the virial excess strength among individual clusters, combined with a limited correlation between virial emission and cluster mass.

7.3 Discussion

Sensitivity and consistency tests are applied to each stage of the analysis, as demonstrated throughout the text and outlined in §C and §H. These tests include varying the parameters controlling cluster selection, data analysis, and stacking procedures. Cluster selection criteria tests include varying the range of angular sizes $\theta_{200}$, the range of masses $M_{200}$ (see §6.5), the galactic latitude cut $b$, and the declination $\delta$. Data analysis tests include Monte-Carlo simulations with control samples (§3.4), variations of the flux cutoff and masking radius of NVSS point sources, and tests of the background model, in particular variations of the order $N_0$ of the polynomial fit and the fitted angular radius $\theta_0$. Stacking procedure tests include testing beam vs. cluster co-addition (§4), least-squares computation both before and after stacking (§5 and §6), folding data 4–8 times (Fig. 5) instead of radial binning, and varying the size $\Delta \tau$ and offset of the radial binning. Overall, the results are found to be quite robust.

Nevertheless, while the general properties of the virial excess are robust, its precise stacked properties, inferred significance, and surrounding features are affected somewhat by the small size of our nominal, 44 cluster sample. In particular, as the projected signal of a virial shock is not expected to be exactly circular, some information and flux are lost when randomly rotated clusters are stacked or when the data are radially binned. Consequently, the peak of the virial excess shifts slightly as a function of frequency and selection criteria, and the normalisation $\tilde{m}\xi\xi_{\nu}^{(2+p)/4}$ may be somewhat underestimated. As an example of a surrounding feature, consider the $\sim 2.5\sigma$ extended, $\tau \approx 5.5$ excess that persists in all channels and effectively raises the background. This signal is robust to most of our sensitivity tests, but vanishes when radially binned if the $20^\circ$ lower limit on the declination $\delta$ is either raised or lowered (see §H); had we not chosen $\delta > 20^\circ$ as our nominal cut, this peripheral signal would vanish and the nominal significance of the virial excess would increase. As another example, the $\tau \approx 1.5$ dip, which may in the future distinguish between the shell (shallow dip) and planar (deep dip) models, is sensitive to our selection criteria.

The nominal data are consistent with both planar and shell models; fitting the model after (before) cluster co-addition, i.e. adopting Eq. (13) [adopting Eq. (12)], slightly favours the planar (shell) model. Interestingly, a similar situation was found in $\gamma$-rays by RK18 when stacking Fermi data around clusters. In principle, projection effects would naturally lead to the shell model, assuming that no preferential orientation survives the co-addition of multiple clusters. However, the signal could still appear planar if there is a substantial quenching of the radio background inside the virial radius (see for example Ando et al. 2023, and references therein). Such quenching could thus deepen the $\tau \sim 1.5$ dip to the observed $\sim 1.5\sigma$, if the former is comparable in magnitude to the virial shock signal. In contrast, if magnetic fields lie preferentially in the plane of the shock, then the $\sim \sin^2 \phi$ dependence upon pitch angle $\phi$ would render the shock signal somewhat less limb-bright, i.e. less planar-like. Note that in the Coma cluster, which is excluded from these stacking analyses due to its large $\theta_{200}$, the planar model provides a better fit, but the distribution of LSS around this particular cluster happens to lie preferably in the plane of the sky (Keshet & Reiss 2018).

A possibly related issue is the $\sim 2.5\sigma$, extended excess peaked at $\tau \approx 5.5$ (see Figs. 4 and 8). This signal passes most of our sensitivity tests (see §4.1.2, §4.2.2, and §H), and although the radially-binned excess is not robust to changes in the declination cut, this may be due to the poor statistics, as the feature appears to survive in 2D significance maps (see Fig. 5 and §M). Curiously, this excess coincides with a similarly stacked $\sim 2\sigma$ excess in $\gamma$-rays (RK18) and with the $\sim 2\sigma$ outer drop in SZ $y$-parameter (Anbajagane et al. 2022). These putative signals, if confirmed, would then suggest that virial shock are non-spherical, with minimal radii around 2.4$R_{200}$ and maximal radii that are a factor 2–3 larger, as found in the Coma cluster (Keshet et al. 2017, 2020b). In such a case, the excess brightness projected at $2.4 \lesssim \tau \lesssim 6$ would effectively raise the subtracted foreground estimate, leading to a deeper $\tau \sim 1.5$ dip and to a more planar appearance of the virial excess; indeed, such an effect would be difficult to distinguish from background quenching. Finally, note that a sidelobe origin can be ruled out for the $\sim 5\sigma$ dip and to a more planar appearance of the virial excess.

Synchrotron emission from the virial shock provides a new probe of the magnetic field in the periphery of galaxy clusters. With $\tilde{m}\xi \approx 0.6\%$ inferred form $\gamma$-ray stacking (RK18), we derive a magnetisation efficiency ranging from $\xi_B \approx (1.6 \pm 0.2\%)$% to $\sim (7.7 \pm 1.0\%)$%; see Eqs. (24) and (25). These estimates are broadly consistent with $\xi_B \approx 1\%$, inferred in the Coma cluster by correlating VERITAS and WMAP signals (Keshet et al. 2017). If we also adopt $\tilde{m} = 1.1$, inferred by in A2319 by combining galaxy counts and $\gamma$-ray signals (Keshet et al. 2020b), then we find $\xi_B/\xi_{\nu} \approx 10$, within an uncertainty factor of 3 or so. This result (which incurs a substantial systematic uncertainty due to the combination of estimates from different clusters), would suggest more energy in magnetic fields than in CREs (although not necessarily in CRIs). Extrapolating available isothermal $\beta$-models for seven clusters in the sample, and correcting for the lower peripheral pressure, we estimate mean magnetic fields $B \approx (0.1-0.3)\mu G$ inside the virial shocks; see §6.6 and Table 3. Note that for such magnetic fields, the radio-
emitting CREs cool over $\lesssim 0.4$ Gyr timescales, thus radiating over a radial range no larger than the extent of the detected virial signal, in agreement with our model.

Our results validate previous reports (see §1) of inverse-Compton and SZ signals from virial shocks, as well as a preliminary correlation between WMAP and VERITAS data in Coma, all inferring similar shock parameters. The detected level of synchrotron emission suggests an all-sky $\nu L_\nu \approx 10^{-11} (\xi, \dot{m}/0.01) (\xi_B/0.01) \text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1}$, dominating the extragalactic low-frequency radio background (Keshet et al. 2017). Synchrotron emission from virial shocks should surface in other present-day and future low-frequency interferometers, and — albeit with a relatively stronger foreground — at higher-frequency radio telescopes. Future work, incorporating additional tracers to support the signal and break the degeneracy between $\dot{m}$ and $\xi_e$ (e.g., using galaxy counts, see Keshet et al. 2020b), and free of our spherically symmetric binning, could map out the cosmic-web evolution. The $\xi_B \gtrsim 1\%$ magnetic fields injected downstream of virial shocks play an important role in the evolution of the ICM. As CREs are consistently found to carry nearly a percent of the energy downstream of virial shocks, their CRI counterparts should hold at least 10% of the energy, leading to a range of nonthermal phenomena as they permeate the ICM (Kushnir et al. 2009; Keshet 2010).

ACKNOWLEDGEMENTS

We are most grateful to Michael W. Eastwood for his key contributions. We thank I. Reiss, G. Ilani, A. Ghosh, E. Waxman, and the anonymous referee for helpful suggestions.

This research was supported by the Israel Science Foundation (grants No. 1769/15 and 2126/22), by the IAC-UPBC joint research foundation (grant 300/18), and by the Ministry of Science, Technology & Space, Israel, and has received funding from the GIF (grant I-1362-303.7/2016). The OVRO-LWA is supported by the National Science Foundation under grant number AST-1828784.

DATA AVAILABILITY

The data generated from computations are reported in the paper, and any additional data will be made available upon reasonable request to the corresponding author.

REFERENCES

Anbajagane D., et al., 2022, MNRAS, 514, 1645
Anderson M. M., et al., 2019, ApJ, 886, 123
Ando M., Shimasaku K., Ito K., 2023, MNRAS, 519, 13
Blandford R., Eichler D., 1987, Phys. Rep., 154, 1
Bona fade A., et al., 2014, MNRAS, 444, L44
Bona fade A., et al., 2018, MNRAS, 478, 2927
Brunetti G., et al., 2008, Nature, 455, 944
Cassano R., et al., 2019, ApJ, 881, L18
Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693
Dickinson C., 2018, Galaxies, 6
Dupke R. A., Mirabal N., Bregman J. N., Evrard A. E., 2007, ApJ, 668, 781
Eastwood M. W., et al., 2018, AJ, 156, 32
Eastwood M. W., et al., 2019, AJ, 158, 84
Fukazawa Y., Makishima K., Ohashi T., 2004, PASJ, 56, 965
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Haslam C. G. T., Kahn F. D., Meaburn J., 1971, A&A, 12, 388
Hoang D. N., et al., 2019, A&A, 622, A20
Hudson D. S., Mittal R., Reiprich T. H., Nulsen P. E. J., Anderson M. M., et al., 2019, MNRAS, 484, 2807
Loeb A., Waxman E., 2000, Nature, 405, 156
Mattox J. R., et al., 1996, ApJ, 461, 396
Miniati F., 2002, MNRAS, 337, 199
O’Hagan A., Leonard T., 1976, Biometrika, 63, 201
Osinga E., et al., 2020, arXiv e-prints, p. arXiv:2011.08249
Piffaretti R., Arnaud M., Pratt G. W., Pointecouteau E., Melin J.-B., 2011, A&A, 534, A109
Pratt C. T., Qu Z., Bregman J. N., 2021, ApJ, 920, 104
Proust D., Capelato H. V., Hickel G., Sodré L. J., Lima Neto G. B., Cuevas H., 2003, A&A, 407, 31
Reiss I., Keshet U., 2018, J. Cosmology Astropart. Phys., 2018, 010
Reiss I., Mushkin J., Keshet U., 2017, preprint, (arXiv:1705.05376)
Rybicki G. B., Lightman A. P., 1986, Radiative Processes in Astrophysics. John Wiley and Sons
Stewart G. C., Fabian A. C., Jones C., Forman W., 1984, ApJ, 285, 1
Sunyaev R. A., Zeldovich Y. B., 1972, Comments on Astrophysics and Space Physics, 4, 173
Totani T., Kitayama T., 2000, ApJ, 545, 572
Waxman E., Loeb A., 2000, ApJ, 545, L11
Wilks S. S., 1938, The Annals of Mathematical Statistics, 9, 60
Yuan Z. S., Han J. L., Wen Z. L., 2015, ApJ, 813, 77
A: SKY MAPS

Figure A1 shows a larger version of the cleared-data sky map of Fig. 1, in comparison to restored-data map.

B: STACKED INTENSITY

While flux co-addition involves a simple summation of the observed $\Delta T \delta \Omega$, stacking the intensity involves a normalisation by solid angle that can be carried out in more than one method. In analogy to the $S$ co-addition, here too we consider weighing the sum either by beams,

$$
\Delta I(\nu, \tau)^{(bm)} = \frac{\sum_{c=1}^{N_c} \sum_{h=1}^{N_p(\tau,c)} \Delta F(\nu, c, h)}{\delta \Omega \sum_{c=1}^{N_c} N_p(\tau,c)} = \frac{2\nu^3 k_B}{c^2} \sum_{c=1}^{N_c} \frac{\sum_{h=1}^{N_p(\tau,c)} \Delta T(\nu, c, h)}{N_p(\tau,c)},
$$

(B1)
or by clusters,

$$
\Delta I(\nu, \tau)^{(cl)} = \frac{1}{N_c} \sum_{c=1}^{N_c} \left[ \frac{\sum_{h=1}^{N_p(\tau,c)} \Delta F(\nu, c, h)}{\delta \Omega N_p(\tau,c)} \right] = \frac{2\nu^3 k_B}{c^2 N_c} \sum_{c=1}^{N_c} \left[ \frac{\sum_{h=1}^{N_p(\tau,c)} \Delta T(\nu, c, h)}{N_p(\tau,c)} \right].
$$

(B2)

Figure A1. Larger version of the sky map in Fig. 1, here showing both cleared (top panel) and restored (bottom) data.
The corresponding uncertainty estimates, accounting for antenna temperature noise, are then estimated as

\[
\sigma_I(\nu, \tau)^{(bm)} \simeq \frac{2\nu^3 k_B n(\nu, \tau) T_0(n)}{c^2} \left[ \sum_{c=1}^{N_c} N_{beam}(\tau, c) \right]^{1/2}
\]

\[
= \frac{2\nu^3 k_B n(\nu, \tau) T_0(n)}{c^2} \left[ \sum_{c=1}^{N_c} N_p(\tau, c) \right]^{1/2}
\]

and

\[
\sigma_I(\nu, \tau)^{(cl)} = \frac{2\nu^3 k_B n(\nu, \tau) T_0(n)}{c^2 N_c} \left[ \sum_{c=1}^{N_c} n^2 N_{beam}(\tau, c) \right]^{1/2}
\]

\[
\simeq \frac{2\nu^3 k_B n(\nu, \tau) T_0(n)}{c^2 N_c} \left[ \sum_{c=1}^{N_c} 1 \right]^{1/2}
\]

Note that due to our choice of random variables, \( S^{(bm)} = \Delta I^{(bm)} / \sigma^{(bm)} \) applies identically, whereas \( S^{(cl)} = \Delta I^{(cl)} / \sigma^{(cl)} \) hold only under the condition

\[
\sum_c (N_{p-1} \sum_n \Delta T) = (N_{c-1} \sum_c N_{p-1})^{1/2} \sum_c (N_{p-1}^{1/2} \sum_n \Delta T).
\]

We also estimate the mean intensity, co-added over the \( N_c \) low-frequency channels, as

\[
\Delta I(\tau) = \frac{1}{N_c} \sum_{\nu=1}^{N_c} \Delta I(\nu, \tau),
\]

for both beam and cluster co-addition. The associated uncertainty is estimated, for simplicity, as

\[
\sigma_I(\tau) = \frac{1}{N_c} \left[ \sum_{\nu=1}^{N_c} \sigma_I(\nu, \tau)^2 \right]^{1/2},
\]

taking into account the same effective correction for correlations between channels as used in Eq. (10), seen to provide the correct results within the limits of full or zero correlations. Again, due to our choice of random variables, the channel co-addition satisfies \( S(\tau) = \Delta I(\tau) / \sigma_I(\tau) \) only under some conditions, in particular

\[
\sum_\nu \Delta I = (N_{\nu-1} \sum_\nu \sigma_I^2)^{1/2} \sum_\nu (\Delta I / \sigma_I)
\]

for beam co-addition. We confirm that the estimate Eq. (B6) is consistent with our control samples.

C: BACKGROUND REMOVAL

An accurate estimate of the background, essential in order to study the weak virial excess signal, can be obtained on the relevant scales by approximating the background as a polynomial in the two sky coordinates (see §3.3). We test polynomials of different orders \( N_b \) determined from data within different outer radii \( \theta_b \), using Monte-Carlo simulations of mock clusters (see §3.4). The sensitivity of our analysis to these two parameters is then quantified.

Different choices of \( N_b \) are illustrated using the control samples in Fig. C1, and using the real MCXC sample in Fig. C2. The figures show the radially-binned significance profiles \( S(\tau) \) for cleared data in the high (top panel) and co-added low-frequency channels, using nominal parameters including \( \theta_b = 5^\circ \). The contribution of odd-order terms in the polynomial is integrated out by the radial binning, so the results for any odd choice of \( N_b \) are similar, although not identical, to those of \( N_b - 1 \).

Figure C1 illustrates the \( S(\tau) \) profile obtained after injecting our nominal, best-fitting central and virial planar models (see Table D1; first row for each frequency range) to random points in the relevant LWA sky. Namely, for each choice of \( N_b \) (see legend), the data around each mock cluster is fitted after injection as a polynomial, and the fit is removed to obtain \( S(\tau) \), using the same pipeline applied to the real sample. The figure also shows (blue disks with solid line to guide the eye) the 95% profile obtained after removing the real LWA background of the control sample, which was measured before injection. This profile is simply the stacked model, without any LWA data background, so no fitting is needed.

As the figure shows, for \( N_b \in \{0, 1\} \), the \( S(\tau) \) profile is very close to this stacked model, indicating that the real background on relevant scales is fairly uniform. For larger \( N_b \), the background fit incorporates an increasing fraction of the virial excess. Consequently, removing a background of higher \( N_b \) progressively lowers \( S(\tau) \). As the figure shows, \( S(\tau) \) outside the central \( \tau \sim 1 \) is thus diminished approximately by a
constant. The central signal is weakly affected by the choice of \( N_b \) because the latter is too sharp and compact to be captured by a low-order polynomial.

In the presence of the Galactic foreground and extended foreground and background sources, it is desirable to include a curvature term in the background fit, so orders \( N_b < 2 \) are insufficient. Moreover, we find that while \( S(\tau) \) approximately follows a normal, normalised distribution for \( N_b \geq 2 \), this is no longer true for \( N_b \leq 1 \). Figure C3 shows that the \( S(\tau) \) distribution for \( N_b = 1 \) (magenta disks and dashed curve) is not symmetric, does not match the normal confidence levels, and shows strong variations with \( \tau \) throughout the \( 0 < \tau < 10 \) range.

Therefore, orders \( N_b = 2 \) and \( N_b = 3 \), which are very

---

**Figure C2.** Same as Fig. C1, but for the real MCXC sample.

**Figure C3.** Statistics of data stacked around the control cluster sample, after removing a polynomial background of order \( N_b = 1 \) (magenta lines with disks), \( N_b = 2 \) (red lines with squares), \( N_b = 3 \) (blue lines with diamonds), or \( N_b = 4 \) (green lines with triangles), and calibrating the noise level using \( \bar{\eta} \), for each frequency channel (see legend) using the cleared data. Beam co-addition curves are shown with the same notations as in Fig. 2.

**Figure C4.** Same as Fig. C2, but for variations in the outer angular radius \( \theta_b \) of the fit. Results shown (see legend) for \( \theta_b = 4^\circ \) (dashed blue curve), \( 5^\circ \) (solid orange), \( 6^\circ \) (dotted green), and \( 7^\circ \) (dot-dashed red).
similar to each other, are the lowest orders possible for the analysis. As increasing $N_b$ shifts more of the virial excess into the background, we designate $N_b = 3$ as our nominal choice. Comparing the blue disks and green diamonds in Fig. C1 indicates that had we known the true background, we could have reported a somewhat more significant virial excess.

Figure C2 shows the results for the real, MCXC sample, with polynomial orders ranging from $N_b = 0$ to $N_b = 5$. The results are very similar to those anticipated using the control sample. As expected, the virial excess is found to be more significant for orders $N_b = 2$ and 3 than it is for higher orders. Repeating the least-squares procedure for different orders, we find that most model parameters are not sensitive to the choice on $N_b$. However, the normalisation of the virial excess does diminish with an increasing $N_b$; conversely, it should be slightly higher if we could use the real background. In particular, the nominal, best-fitting normalisation $\hat{m}_v \xi_B^{(2+p)/4}$ of the virial excess at order $N_b = 3$ is higher than found for $N_b \geq 4$. For example, for $N_b = 4$, the nominal $\hat{m}_v \xi_B^{(2+p)/4}$ is $\sim 20\%$ lower than its $N_b = 3$ counterpart in the result (20), i.e. for the planar model with nominal $\chi^2$ in the high-frequency channel.

Figure C4 shows the effect of changing the radius $\theta_b$ used to estimate the background polygonal. We find that increasing $\theta_b$ beyond the nominal 5° does not have a significant effect on the results. However, a smaller, $\theta_b = 4°$ region is not sufficiently large for a good handle on the background, and is strongly affected by the central excess, thus diminishing the $S(\tau \gtrsim 1)$ profile.

D: CONTROL SAMPLES AND BEST-FIT RESULTS

Our nominal analysis is based on the cleared data of the high-frequency channel. This choice combines a high resolution with the removal of point-source contamination using the modified CLEAN algorithm. The corresponding control sample, shown in the top left panel of Fig. 2, approximately follows a normal distribution, at least out to the accessible $\pm 3\sigma$ confidence level. Good results are obtained for these nominal data even with no calibration at all, as evident from $\sigma \sim 0.9$ (see Table 1) being close to unity in this channel. The well-behaved control data support the application of the estimates (8) and (9) to the nominal, real sample.

As pointed out above, the restored data (bottom panels in Fig. 2) show a somewhat skewed distribution, due to the effect of the restored point sources. In the high-frequency channel, the distribution of $S$ is fairly independent of $\tau$ even in the restored data. We find the confidence intervals here to be consistent with a normalised skew normal distribution, e.g., $O'Hagan & Leonard (1976)$, $S \propto e^{-x^2/2} [1 + erf(\alpha x/2)]$, where erf is the error function. The best-fitting skew parameter $\alpha \simeq 0.12$ is obtained for the log-histogram in the $\tau > 0.5$ range. One could in principle correct
the skewed distribution to a normal one. Without such a correction, Eqs. (8) and (9) can still be used, but small, $\lesssim 10\%$ errors are introduced in the restored-data $S$. When plotting $S$, in addition to the nominal values based on these equations, we also present the confidence levels extrapolated from the control samples assuming a normal distribution. Such an extrapolation provides a fairly accurate, $|\Delta S| \lesssim 0.2$ description of the cleared data, but has only $|\Delta S/S| \lesssim 10\%$ accuracy for the restored data due to the skewness.

As §E shows, individual low-frequency channels show a similar behaviour: an approximately symmetric, $\tau$-independent, normal distribution of $S$ for the cleared data, and a somewhat asymmetric, approximately skew-normal distribution for the restored data, with $\alpha_\tau$ in the range $\sim 0.08$–0.12 for different channels. In the lower frequencies, where the resolution generally deteriorates, some variations with $\tau$ become more apparent. Similar results are found after the co-addition of low-frequency channels in Eq. (10), as shown in the right panels of Fig. 2. Here, after $\bar{\eta}$ is calibrated in each channel, we calibrate also the correlation parameter $\psi$, by demanding that the standard deviation of $S$ in the control samples be unity, after channel co-addition and averaging over $\tau$. The results show an approximately symmetric, normal and normalised distribution of $S$ for the cleared data, and a slightly skew-normal distribution in the restored data with $\alpha_\tau \approx 0.11$.

With the $\bar{\eta}$ and $\psi$ parameters calibrated above (see Table 1), we use the same method to substantiate the standard deviations of the stacked brightness $\Delta I$, in its different co-addition variants. Namely, we test if the control samples are consistent with the stacked brightness in Eqs. (B1), (B2), and (B5) having the uncertainties estimated in Eqs. (B3), (B4), and (B6), respectively, and with the same calibration. To do so, we compare the distribution of $\Delta I$ among the control samples to a normal distribution with the standard deviation $\sigma_I$ given by the above calibrated equations. We find a good agreement, within $\lesssim 2\%$, out to the accessible, $\pm 3\sigma$ confidence level for the cleared data.

Table 2 shows the best-fit results, based on the calibrated control samples, for our nominal models. Table D1 provides results for a range of model variants, demonstrating the robustness of the results.

E: CALIBRATING NOISE AND CORRELLATION FACTORS

Although the LWA antenna noise is approximately known in each channel, a more accurate determination of the noise level and any inter-correlations between channels is necessary in order to correctly determine the significance level of any excess signal. We use the statistical distribution of the control samples, first to calibrate the noise level in each channel, and then to estimate the correlations between channels, as shown in §3.4.

Figure E1 demonstrates the control sample distribution in the same manner as Fig. 2, but for each channel separately. The figure shows that after calibrating the noise levels, good agreement is obtained between the nominal confidence level $S$ (values on the $y$-axis), the corresponding median and $68\%$, $95\%$, and $99.7\%$ containment levels (dashed curves), and the corresponding normal-distribution significance levels (the mean offset by integer multiples of the standard deviation; dotted curves). The figure adopts the same notations as Fig. 2, but for simplicity, only includes beam co-addition results.

The noise calibration factors $\eta(\nu)$ are properties of the telescope, and so, should not depend on the analysis details. Indeed, the figure shows that the factors needed to obtain the above agreement do not vary significantly with $\tau$ (except at small radii, where the statistics is poor). We thus adopt the $\tau$-averaged value $\bar{\eta}(\nu)$ in each channel, as summarised in Table 1.

As another illustration, we examine the statistical distribution of control samples after the same noise calibration, with different choices of $N_0$. Figure C3 demonstrates the confi-
dence level estimates for \( N_b = 1, 2, 3, \) and \( 4, \) for each frequency channel, using the cleared data. We find similar distributions of \( S(\tau) \) for \( 2 \leq N_b \leq 4, \) agreeing with a normal, normalised distribution across the explored \( \tau \) range, when using the same calibration factors. For \( N_b = 1, \) the \( S(\tau) \) distribution becomes skewed and is neither normal nor normalised; however, as mentioned in §C, this results from the low-order fit and is unrelated to the noise calibration. More generally, we find that the same noise calibration works for all reasonable analysis variations.

Our procedure of effectively correcting the linear regression for correlations among the low-frequency channels is described in §3.2 and §5.2. Table E1 demonstrates how replacing \( \chi^2 \) by a corrected \( C\chi^2 \) in the nominal (cluster but not channel co-added) least-squares Eq. (13) effectively offsets these correlations. Namely, the table illustrates how applying \( C \approx 0.25 \) to Eq. (13) yields the same confidence intervals as those obtained with the channel co-added Eq. (14), i.e. by fitting the data after co-adding channels with the correlation correction in Eq. (10). Moreover, introducing this \( C \) correction is shown in §5.2 to be equivalent to a generalised Eq. (10) with the same parameter \( \psi \). The best-fitting values are unaffected, of course, by this correction.

**F: CENTRAL EXCESS**

For the central emission, \( \chi^2 \) is computed using the radial range \( 0 \leq \tau \leq 1.5 \). Notice that a fit to the extended component cutoff parameter \( \tau_{\text{cut}} \) may still fall outside the fitting region, as \( \tau > 1.5 \) emission can contribute to the \( \tau < 1.5 \) signal after PSF convolution. We verify that the results are not sensitive to the precise \( \tau \) range used for the fitting, including larger ranges reaching \( 0 \leq \tau \leq 2 \).

The best-fitting models for the central emission are derived and studied using the control samples. The results are presented in §F1 and Table F1 based on the raw (restored without NVSS point-source masking) and in §F2 and Table F2 based on the cleared data.

**F1 Restored data**

We first model the central signal in the raw data. To simplify this modelling, we use an average FWHM value for the restoring beam in each channel, in the declination range \( 20^\circ < \delta < 60^\circ \) relevant to the (41 out of 44) clusters in our sample. One way to quantify this FWHM is by fitting the raw data to the convolution of a circular Gaussian with many
(say, 300) bright \( F_1.4 \text{GHz} > 1 \text{ Jy} \) and compact (major-axis \(< 0.25^\circ \)) known (NVSS) sources. The resulting FWHM values (see Table 1) are fairly constant in this declination range, showing for example \( < 3\% \) differences between \( 20^\circ < \delta < 40^\circ \) and \( 40^\circ < \delta < 60^\circ \) fits. It is challenging to fit the central cluster emission with our oversimplified model, composed of only two, point-like and extended, components, each with a different power-law spectrum.

The central signal is reasonably fitted in the seven low-frequency channels by the Gaussian-convolved point-source model (\( \Delta k = 0 \)), with a \( \chi^2 \) per DOF of \( \chi^2_n(\tau \leq 1.5) \approx 2.8 \) or equivalently \( \chi^2/\Delta k(\tau \leq 1.5) \approx 1.3 \). However, such a point-source model does not provide a good fit in the high-frequency channel, where \( \chi^2_n(\tau \leq 1.5) \approx 12.2 \). Modelling point and extended components simultaneously provides a good fit both in the high-frequency channel, with \( \chi^2_n(\tau \leq 1.5) \approx 0.8 \), and at low frequencies, where \( \chi^2/\Delta k(\tau \leq 1.5) \approx 0.6 \); see best-fit parameters in Table F1. Here, we assume that the extended component is bright and compact enough to be mostly picked-up by CLEAN, and hence convolve both point-like and extended components with the Gaussian kernel. This procedure is somewhat inaccurate, because \( \sim 1/4 \) of the total raw central excess is not picked up by CLEAN (see §F2). At both high and low frequencies, we find that the point-like and extended components make equal contributions to the raw data, i.e. \( F_\ell \approx 0.5 F_{7\tau} \). When excluding the cluster MCXCJ2338.4+2700, whose centre harbours bright NVSS point-sources, including an \( F_{1.4 \text{GHz}} \approx 2.8 \text{ Jy} \) source, the contribution of the extended component increases to 80%.

Figure F1 shows that the central signal (for the 44-cluster sample; solid red) is fairly well fit by this Gaussian-convolved two-component model in each channel separately (double-dot-dashed green). The best-fit spectral index is

\[
\alpha_\ell = 1.16^{+0.10}_{-0.10} (1.22^{+0.13}_{-0.14}) \quad \text{(F1)}
\]

for the point-like component, and

\[
\alpha_e = 0.76^{+0.09}_{-0.09} (0.95^{+0.12}_{-0.17}) \quad \text{(F2)}
\]

for the extended component, using beam (cluster) co-addition. Differences between beam and cluster co-addition might be due to the aforementioned cluster MCXCJ2338.4+2700. Indeed, if we exclude this cluster, the spectral indices become \( \alpha_\ell = 1.20^{+0.20}_{-0.20} (1.13^{+0.16}_{-0.16}) \) and a more consistent \( \alpha_e = 1.07^{+0.23}_{-0.14} (1.05^{+0.18}_{-0.13}) \) for beam (cluster) co-addition. The spectra slightly softens when excluding this source, which itself shows an \( \alpha \approx 0.75 \pm 0.5 \) spectrum based on its catalogued GMRT (Giant Metrewave Radio Telescope; Intema et al. 2017) and NVSS fluxes.

These spectra are sensitive not only to the co-addition method, but also to the choice of channels. Fitting the four lowest-frequency channels, focusing for example on beam co-addition, gives a slightly harder \( \alpha_\ell = 1.05^{+0.09}_{-0.09} \) but a similar \( \alpha_e = 0.82^{+0.27}_{-0.26} \); using the three higher frequency channels gives a softer \( \alpha_e = 0.95^{+0.13}_{-0.14} \) and a much softer \( \alpha_\ell = 2.14^{+0.20}_{-0.20} \). Overall, we find spectral indices in the range \( 1.0 \lesssim \alpha_\ell \lesssim 2.2 \) and \( 0.7 \lesssim \alpha_e \lesssim 1.1 \) for our best-fit \( \zeta \) and \( \tau_{\text{cut}} \) values. We cannot determine the spectrum more precisely, due to our weak constraints on \( \zeta \) and \( \tau_{\text{cut}} \) (see Table F1) and to the sensitivity of the results to assumptions.

As expected, modelling the data with central components alone does not provide a good fit to the data beyond \( \tau \sim 2 \), and in particular, does not reproduce the virial excess. Nevertheless, although the sidelobe pattern induced by the strong central signal should ideally be entirely removed by CLEAN, some residual may remain in the data due to PSF inaccuracies, noise, and faint sources below the CLEAN threshold. The worst-case putative residual artefacts can be crudely estimated by a PSF convolution of the best-fit to the raw data (itself obtained by convolving the model with the Gaussian restoring beam), shown as dashed curves in Fig. F1. Such a PSF convolution leads to a significantly worse fit to the raw data, especially in the \( 1 < \tau < 2 \) region, suggesting that CLEAN has been rather successful in removing the full PSF-convolved central signal. For example, our best fit to the central region, convolved with the PSF, gives \( \chi^2_n \approx 134.5 \) for the high-frequency channel and \( \chi^2/\Delta k \approx 7.1 \) for the low-frequency channels in the \( \tau \leq 1.5 \) region.

Even a putative failure of CLEAN only at large angular separations, corresponding to \( \tau > 2 \), does not provide a good explanation for the virial excess. In terms of amplitude alone, this scenario would require CLEAN to remove essentially all of the sidelobe flux inside \( \tau < 2 \), but little to none of it beyond \( \tau > 2 \), and even that would not suffice to explain the excess. In terms of the angular position of the ripples, the \( \tau \)-dependence does not match the data either, except for a few chance combinations of individual clusters in specific channels (see Fig. F1). Quantitatively, the PSF-convolved models derived based on the \( \tau \leq 1.5 \) data provide a poor fit in the \( 2 < \tau < 5 \) region, with \( \chi^2_n(2 < \tau < 5) \approx 10.2 (\chi^2/\Delta k \approx 3.1) \) for the high (low) frequency channels. This should be compared to \( \chi^2_n \approx 5.9 (\chi^2/\Delta k \approx 2.9) \) with the Gaussian convolution. As the figure shows, neither a PSF nor a Gaussian convolution of the model shows a visually reasonable fit in the \( 2 < \tau < 5 \) range. These conclusions are not sensitive to the precise range of fitted radii.

Indeed, while putative PSF sidelobe residuals of the central signal could, in principle, generate some excess rings, we find that they cannot mimic the virial excess: the first ripple does not reach the virial radius, while the second ripple is both too weak and typically too extended to reproduce the virial signal. Consider the low-frequency channels first. As Fig. F1 shows, in the two lowest-frequency channels, the PSF residuals are negligible. Higher-frequency channels show stronger residuals, but misaligned with the virial excess of most or all clusters. In the highest of these channels, 67.9 MHz, the first ripple is expected to lie, for different clusters, at \( 1.3 \lesssim \tau \lesssim 2.5 \) for different clusters, while the second ripple lies at \( 2.0 \lesssim \tau \lesssim 4.0 \); the higher \( \tau \) values are more relevant due to more numerous compact clusters. As the figure shows, the co-added result gives a first ripple peaked at \( \tau \sim 2.2 \), slightly inside the low-frequency virial excess, and a very weak second ripple at \( \tau \sim 3.7 \), far outside the virial signal. In any case, the mismatch between the restored data and the PSF-convolved central excess (compare solid and dashed curves in the figure) shows that sidelobes are removed quite accurately by CLEAN.

For the high-frequency channel, PSF-convolution shows a peculiar flat excess in the \( 1 \lesssim \tau \lesssim 2.5 \) range. This excess can be traced to a positive average of the PSF outside the main beam in this channel. As mentioned above, our results indicate that CLEAN efficiently removed this residual. But even if this were not the case, one can show that this residual cannot mimic the virial excess. For example, the bottom-right panel of Fig. F1 shows that this flat residual does not even

\[ \sum_{\text{data}} \]
Figure F1. Best-fit to the raw (restored without NVSS point-source masking) data using the central model only, shown in each individual channel (labelled frequencies) and stacked over the seven low-frequency channels (labelled “low ν”). The seven low-frequency channels are jointly fit with six free parameters \( (F_p, F_e, \alpha_p, \alpha_e, \zeta \text{ and } \tau_{\text{cut}}) \), while the high-frequency channel is fitted separately and without spectral parameters. Each panel shows the beam co-added significance profiles of the raw data (solid red curve), the best-fit model with a Gaussian restoring beam (using Eq. (13) for \( \chi^2 \); double dot-dashed green) and the same model convolved with the PSF (dashed blue). In the 73 MHz channel (bottom-left panel), PSF convolution yields a flat, \( 1 \lesssim \tau \lesssim 2.5 \) excess due to the more compact clusters; the bottom right panel shows that this feature does not reach the virial radius for the larger clusters (shown for the eight clusters of \( \theta_{500} > 0.3 \)), which alone still show a virial excess in cleared data (cyan dot-dashed curve). The virial shock range of Fermi is highlighted (notations as in Fig. 4).
reach the virial radius for the more extended (θ_{200} > 0.3) clusters in our sample.

### F2 Cleared data

Fitting the cleared data with both components of the central emission model, we find that the point-like component makes a negligible contribution, indicating an efficient CLEAN. Fitting the point-like component alone fails to reproduce the central excess in the cleared data (compare the solid blue and dot-dashed green curves in Fig. F2), with fairly large \( \chi^2(\tau \leq 1.5) \approx 3.5 \) for the high-frequency channel and \( C\chi^2(\tau \leq 1.5) \approx 1.2 \) for the co-added low-frequency channels. In contrast, the extended emission model nicely reproduces the central significance profiles (compare the solid blue and double dot-dashed purple curves in the figure), with \( \chi^2(\tau \leq 1.5) \approx 0.2 \) for the high-frequency channel and \( C\chi^2(\tau \leq 1.5) \approx 0.1 \) for the co-added low-frequency channels.

![Figure F3](image-url) Best fit (symbols and 1σ to 3σ confidence contours) for central extended-source parameters \( F_e \) (at 73 MHz) and \( \alpha_e \), for the low-frequency cleared data. Results are shown for nominal analysis with beam (circle with purple solid contours) and cluster (square with magenta dashed contours) co-addition.

With this extended component alone, we find that the cleared data has a total central flux \( (F_{73}) \) which is about a quarter of its raw-data counterpart (compare Tables F1 and F2). The spectral index of the extended source for the cleared data is

\[
\alpha_e = 1.08^{+0.24}_{-0.24} \quad (1.11^{+0.25}_{-0.26})
\]  

(F3)

for beam (cluster) co-addition; see Fig. F3. This spectrum is consistent with the backward estimate of the cleared-central emission in §4.2.4.

The \( \zeta \) and \( \tau_{cut} \) values obtained using the cleared data are broadly consistent with their raw-data counterparts, but the large statistical uncertainties in these two parameters and the different values obtained for high vs. low frequencies indicate that the model is too crude (see Table F2). However, the best-fitting results for \( \alpha_e \) are not sensitive to the choice of \( \zeta \) and \( \tau_{cut} \). For simplicity, we fix \( \zeta = 0.9 \) (1.2) and \( \tau_{cut} = 0.9 \) (1.6) for the high (low) frequency data when simultaneously modelling the central and virial excess signals in §6. Adopting the standard \( \chi^2 \) of Eq. (12) instead of our nominal Eq. (13) gives more consistent best-fitting values for \( \zeta \) and \( \tau_{cut} \) among high and low frequencies (see Table F2), without changing the spectrum significantly; namely, giving \( \alpha_e = 1.13^{+0.27}_{-0.27} \) consistent with Eq. (F3).

The resulting model, convolved with the PSF before fitting to the data, shows (see Fig. F2) some minor sidelobe ripples, but again, these artefacts cannot account for the virial access in terms of neither amplitude nor position. This conclusion pertains to each of the channels separately, and to both extended and point-like models. We conclude that the extended central excess in the cleared data has a very minor effect near the virial radius. This strengthens our conclusion that any central excess sidelobes missed by CLEAN cannot mimic the virial excess.

### G: MORE TESTS RULING OUT VIRIAL PSF RIPPLES

As the centres of clusters are often radio bright and the stacked virial signal assumes the shape of a concentric ring, it is important to test if the virial signature could be an artefact arising from PSF sidelobes of the central emission. Several tests were outlined in §4.1.4 and reviewed in §7, indicating that the virial signal is not such an artefact.

Here, we present two such tests. In the first, we attempt to mimic the virial ring using PSF sidelobes, by stacking bright sources or bright regions in the sky that are unrelated to clusters. In the second, we selectively exclude from our sample clusters with a particularly strong central excess, to see if this
but for cleared data. The model (11) is used with a single parameter $F_{73}$ for the high-frequency channel, while the seven low frequencies are jointly fit with the same two parameters, $F_{73}$ and $\alpha$. The parameters $\zeta$ and $\tau_{\text{cut}}$ are fixed. Each panel shows the beam co-added significance profiles of the data (solid blue curve) and of the PSF-convolved point (dot-dashed green) and extended (double dot-dashed purple) central emission models.

**Figure F2.** Same as Fig. F1 but for cleared data. The model (11) is used with a single parameter $F_{73}$ for the high-frequency channel, while the seven low frequencies are jointly fit with the same two parameters, $F_{73}$ and $\alpha$. The parameters $\zeta$ and $\tau_{\text{cut}}$ are fixed. Each panel shows the beam co-added significance profiles of the data (solid blue curve) and of the PSF-convolved point (dot-dashed green) and extended (double dot-dashed purple) central emission models.

**Table F2.** Best-fit results for the central emission in the cleared data.

| Channel | Central source model | Weights | $F_{73}$ | $\alpha_c$ | $\zeta$ | $\tau_{\text{cut}}$ | $\chi^2$ | $N'$ | $\chi^2$ Eq. (10) |
|---------|----------------------|---------|---------|------------|--------|---------------------|---------|-----|------------------|
| high    | extended $\times$ PSF| bm      | 1.83$^{+0.24}_{-0.22}$ | $-0.94^{+0.49}_{-0.32}$ | 0.95$^{+0.32}_{-0.21}$ | 0.7 | 3 | (13) |
| low     | extended $\times$ PSF| bm      | 2.13$^{+0.28}_{-0.20}$ | $1.08^{+0.24}_{-0.24}$ | 1.24$^{+0.14}_{-0.19}$ | 1.61$^{+0.37}_{-0.37}$ | 19.0 | 38 | (13) |
| high    | extended $\times$ PSF| bm      | 1.87$^{+0.18}_{-0.18}$ | $-0.9$ | $0.9$ | 0.9 | 5 | (13) |
| low     | extended $\times$ PSF| bm      | 2.21$^{+0.22}_{-0.22}$ | $1.01^{+0.22}_{-0.22}$ | $1.2$ | $1.6$ | 19.5 | 40 | (13) |

**Columns:** (1) The high-frequency (high) or the seven co-added low-frequency (low) channels; (2) Central source model, specifying point or extended components in Eq. (11) with Gaussian or PSF convolution; (3) Stacking method: beam co-addition (‘bm’ for brevity) or cluster co-addition (‘cl’ for brevity); (4) Flux of the central emission at 73 MHz in units of $10^{-23}$ erg s$^{-1}$ cm$^{-2}$ Hz$^{-1}$; (5) Spectral index $\alpha = -d\ln I_\nu/d\ln \nu$ of the central emission; (6) Radial power law of the extended central emission in Eq. (11); (7) Normalised cutoff radius for the extended central emission in Eq. (11); (8) $\chi^2$ value of the best fit, uncorrected for correlations; (9) Number of degrees of freedom in the fit; (10) The $\chi^2$ equation. Values in **boldface** are fixed, not fit, parameters. In the low-frequency channels, the uncertainties are corrected using Eq. (15).
may have any effect on the virial excess. Both tests, like the others, summarised in §7, do not support a putative PSF sidelobe origin for the virial excess.

**G1 PSF sidelobes fail to mimic virial excess**

The procedure is identical to that used for control clusters in §3.4, but instead of stacking random regions in the allowed part of the sky (§2.2), here we stack local brightness peaks in the restored data, around which one may find sidelobe emission that was not fully removed by CLEAN. We test different prescriptions for the scale and compactness of these peaks, and examine the cleared-data results using the same nominal stacking and analysis pipelines described in §3.

Figure G1 demonstrates the resulting high-frequency (top panel) and low-frequency (bottom) significance profiles, for different choices of brightness peak samples. These include peaks in the LWA sky on small (smoothed on radius 5′) or medium (radius ∼ θ500 ∼ 15′) scales, after large-scale sky structures (obtained by smoothing with a Gaussian of radius 40′) were removed. We also examine sharp vs. shallow peaks on 15′ scales, defined by comparing the maximal gradient at the smoothing radius to its ensemble median.

As the figure illustrates, in all our attempts, we were unable to obtain a significant excess at scales corresponding to cluster peripheries, even though we choose a control sample with a stacked central signal stronger than in the real sample. Some ripples can be seen, but in addition to being weak, they are also misaligned with the real virial excess.

**G2 Virial excess unaffected by changing central excess**

If the virial excess were a sidelobe artefact of emission from the centres of clusters, then lowering the latter would diminish the former. Such a scenario is readily tested, as the central excess is dominated by a few clusters harbouring bright NVSS sources near their centres. Figure G2 demonstrates that if one or more of these clusters are excluded from our nominal cluster sample, the raw (restored and unmasked; see inset) central excess is nearly halved, while the nominal (cleared and masked) virial excess is barely affected. This test alone proves that the virial signal does not depend on most, if not all, of the central emission.

In particular, only one cluster harbours $F_{1.4\text{GHz}} > 1$ Jy NVSS sources. This cluster alone contributes $\sim 50\%$ ($\sim 45\%$) of the central excess at high (low) frequencies, as two very bright, $\gtrsim 2.4$ Jy sources are found within $0.2'$ of its centre. As the figure shows, removing this cluster alone dramatically

---

**Figure G1.** Excess significance profiles of the control samples purposely chosen around random local brightness maxima, for the high (top panel) and co-added low (bottom) frequency channels, using the cleared data. Shown are results using the smoothing scale 15′ (blue disks with dashed lines to guide the eye), 5′ (orange squares with dotted lines to guide the eye), 15′ with sharp peaks (green triangles with dot-dashed lines to guide the eye), 15′ with shallow peaks (red stars with double dot-dashed lines to guide the eye). Other notations are as in Fig. 4.

**Figure G2.** Excess significance in the high (top panel) and co-added low (bottom) frequency channels, before (blue disks with solid line to guide the eye) and after (curves) removing the one cluster with $F_{1.4\text{GHz}} > 1$ Jy (dashed orange), or the four clusters with $F_{1.4\text{GHz}} > 0.5$ Jy (dot-dashed green), NVSS sources within $0.2'$ from their centres. Other notations are as in Fig. 4. While removing these clusters from the nominal sample substantially lowers the central excess in the restored, unmasked data (insets), the effect on the nominal virial excess in the cleared, masked data (main figures) is minute.
demonstrates variations in the angular range. Further excluding clusters harbouring $F_{1.4\text{GHz}} > 0.5$ Jy NVSS sources within 0.2$ from their centres removes three additional clusters from the sample. The remaining, 40 cluster sample shows a raw central excess fainter by a factor of $\gg 2.3$ than in the nominal sample, yet the virial excess is still barely modified.

**H: SENSITIVITY AND CONSISTENCY TESTS**

We apply sensitivity and consistency tests to each stage of the analysis, in order to examine its robustness and identify optimal control parameter values. Below we provide tests of parameters controlling the cluster selection (range of angular size $\theta_{500}$, the cut on galactic latitude |b|, and the cut on the declination $\delta$), the data analysis (the cut on the flux $F_{1.4\text{GHz}}$ of NVSS point sources and the masking radius around them), the and stacking procedure (choice of radial bin size $\Delta \tau$).

For simplicity, in the following figures, each parameter is varied individually, leaving the remaining parameters nominal: galaxy clusters with angular dimensions in the range $0^\circ.2 \leq \theta_{500} < 0^\circ.4$, galactic latitudes $|b| > 20^\circ$, and declination $\delta > 20^\circ$, after masking the data within 0.3$ from NVSS point sources of flux exceeding 1 Jy, removing a background polynomial model of order $N_b = 3$ estimated within $\theta_b = 5^\circ$, and radially binning the results with resolution $\Delta \tau = 0.25$.

Each figure shows both high (top panel) and co-added low (bottom) frequency channel results, but focuses for brevity on the beam co-added, cleared data.

Figure H1 demonstrates variations in the angular range $\theta_{500}$ of clusters in our sample. Alternative ranges shown are $0^\circ.18 < \theta_{500} < 0^\circ.4$ (leaving $N_c = 55$ clusters instead of the nominal $N_c = 44$), $0^\circ.22 < \theta_{500} < 0^\circ.4$ ($N_c = 33$), $0^\circ.24 < \theta_{500} < 0^\circ.4$ ($N_c = 21$), $0^\circ.2 < \theta_{500} < 0^\circ.5$ ($N_c = 49$), and $0^\circ.15 < \theta_{500} < 0^\circ.4$ ($N_c = 89$); see label. As the figure shows, the virial signal is not sensitive to small changes near the nominal range. Lowering the minimal $\theta_{500}$ admits more clusters at marginal or insufficient resolution; therefore, a substantially smaller $\theta_{500}$ washes out the excess. Raising the maximal $\theta_{500}$ admits highly extended clusters, thus contaminating the virial signal due to two effects: (i) substantial foreground and background structures on these scales are introduced to the stacking; and (ii) the extended clusters are bright, and thus offset the background removal themselves.

Figure H2 shows the effect of varying the cut on $|b|$. Alternative cuts shown in the figure are $|b| > 0^\circ$ ($N_c = 57$), $|b| > 10^\circ$ ($N_c = 51$), $|b| > 30^\circ$ ($N_c = 34$), and $|b| > 40^\circ$ ($N_c = 18$); see label. As the figure shows, raising the cut from the nominal $20^\circ$ to $30^\circ$ does not significantly change the virial excess, in spite of the smaller sample size, suggesting that Galactic contamination is not entirely negligible at $20^\circ < |b| < 30^\circ$. The signal starts dropping if the $|b|$ cut is raised further, due to the smaller $N_c$. Lowering the latitude cut to $10^\circ$ also does not significantly change the virial excess, but does increase the noise. Lowering the cut further does di-

---

**Figure H1.** Significance profiles with various $\theta_{500}$ selections: $0^\circ.18 < \theta_{500} < 0^\circ.4$ (dashed orange curve), $0^\circ.2 < \theta_{500} < 0^\circ.4$ (default; solid blue), $0^\circ.22 < \theta_{500} < 0^\circ.4$ (dotted green), $0^\circ.24 < \theta_{500} < 0^\circ.4$ (dotted-dashed red), $0^\circ.2 < \theta_{500} < 0^\circ.5$ (double dot-dashed purple), $0^\circ.15 < \theta_{500} < 0^\circ.4$ (long dashed magenta), for the high-frequency (top penal) and low-frequency data (bottom), using the cleared data. The number of clusters ($N_c$) in each sample is shown in the legend. Other notations are as in Fig. 4.

**Figure H2.** Same as Fig. H1, but varying the latitude cut: $|b| > 0^\circ$ (dashed orange curve), $|b| > 10^\circ$ (dotted green), $|b| > 20^\circ$ (default; solid blue), $|b| > 30^\circ$ (dot-dashed red), and $|b| > 40^\circ$ (double dot-dashed purple).
minish the virial signal, although it remains detectable even with no latitude cut.

Figure H3 demonstrates variations in the minimal declination threshold of the cluster sample. Alternative thresholds shown are \( \delta > 0 ^\circ \) \((N_c = 79)\), \( \delta > 10 ^\circ \) \((N_c = 59)\), \( \delta > 20 ^\circ \) \((default; \ N_c = 44)\), \( \delta > 30 ^\circ \) \((N_c = 29)\), and \( \delta > 40 ^\circ \) \((N_c = 16)\); see label. As shown in the figure, the virial excess is not sensitive to the choice of the cut. A high \( \delta \) cut leaves fewer clusters in the sample, but these clusters have better PSF properties.

\[\text{Figure H3. Same as Fig. H1, but varying the declination cut: } \delta > 0 ^\circ \text{ (dashed orange curve), } \delta > 10 ^\circ \text{ (dotted green), } \delta > 20 ^\circ \text{ (default; solid blue), } \delta > 30 ^\circ \text{ (dot-dashed red), and } \delta > 40 ^\circ \text{ (double dot-dashed purple).}\]

Figure H4 demonstrates the effect of varying the cut on the NVSS point-source flux, in the range 100 mJy to 2 Jy. Masking pixels within 0.3 around NVSS point sources with \( F_{1.4GHz} > 200 \) mJy or \( F_{1.4GHz} > 500 \) mJy does not significantly affect the virial shock signal, but it does diminish the central signal. The approximate independence of the virial excess upon the NVSS flux cut indicates either that CLEAN has successfully removed point sources down to the \( F_{1.4GHz} \sim 200 \) mJy level within a radius 0.3, or that point sources neither contaminate the virial excess nor contribute to it; the change in central excess would favour the latter. The virial excess is still present when the cut is lowered further, to 100 mJy, but becomes fainter, probably due to the excessive masking of pixels: here, only \( \leq 60\% \) of the initial pixels in the virial bin remain unmasked, compared to \( \sim 85\% \) \((\sim 90\%)\) for the \( F_{1.4GHz} = 200 \) (1000) mJy cut.

Figure H5 demonstrates the dependence of the results upon the radius within which HEALPix pixels are masked around NVSS point sources. Varying the masking radius in the range 0.1 to 0.6 has a negligible effect on the results, but further raising the masking radius removes enough pixels to start diminishing the virial excess, especially at low frequencies.

\[\text{Figure H4. Same as Fig. H1, but varying the point-source flux cut: } F_{1.4GHz} > 100 \text{ mJy (dot-dashed orange curve), } F_{1.4GHz} > 200 \text{ mJy (dotted green), } F_{1.4GHz} > 500 \text{ mJy (dashed red), } F_{1.4GHz} > 1 \text{ Jy (default; solid blue), and } F_{1.4GHz} > 2 \text{ Jy (double dot-dashed purple).}\]

Figure H6 demonstrates variations in the radial bin size, from \( \Delta \tau \sim 0.167 \) (six bins inside per \( \tau \)) to \( \Delta \tau = 0.5 \) (two bins per \( \tau \)). The \( S(\tau) \) profiles of different resolutions are consistent with each other, and show that both central and virial excess signals are extended, as their significance increases as the bins are enlarged beyond the beam size. The peak of the excess slightly shifts between high and low frequencies at nominal and low resolutions, but only between adjacent radial bins, so we cannot substantiate any frequency dependence.

Out of the 44 clusters in our sample, 12 can be presently classified as perturbed (due to long cooling times, an absent cool-core, or an identified merger), and two as relaxed. One expects the peripheral, virial excess to be fairly indifferent to the dynamical state of the cluster, whereas the central signal should be modified, for example become more extended, in perturbed clusters. To examine the effect of the dynamical state of the cluster, Fig. H7 shows the radial significance profiles obtained separately for the 12 disturbed clusters and for the two relaxed clusters. Both sub-samples show central and virial signals, but due to the small sample sizes, we cannot identify any significant differences between sub-samples.

I: VIRIAL SHOCK MODEL

We generalise the virial shock model of RK18 for radio emission, using the same underlying assumptions but incorpo-
rating also the magnetic field and the resulting synchrotron emission, recovering Eq. (16) used in §6.

To determine the synchrotron luminosity of the virial shock, we first derive an approximation for the accretion rate through the shock and its downstream conditions. The thermal plasma is assumed neutral, with a particle number density following an isothermal $\beta$-model,

$$n(r) = n_0 \left[ 1 + \left( \frac{r}{r_c} \right)^2 \right]^{-3\beta/2},$$ (11)

where $n_0$ is the central number density, $r_c$ is the core radius, and $\beta$ is the slope parameter. Approximating the cluster as in hydrostatic equilibrium, the total (gravitating) mass inside a radius $r$ can be written as

$$M(r) \simeq \frac{3\beta kT}{Gm} \left( 1 + \frac{r^2}{r_c^2} \right)^{-1},$$ (12)

where $G$ is Newton’s constant. As the typical core radius $r_c \sim 0.1R_{500}$ is much smaller than the virial shock position at $r \gtrsim 2R_{500}$, to study the distribution near the shock one can neglect the core and approximate $n(r) \simeq n_0(r/r_c)^{-3\beta}$.

As $r \gg r_c$, the mass can be approximated as

$$M(r) \simeq \frac{3\beta kT}{Gm} r.$$ (13)

This approximation yields the enclosing radius

$$R_s = \left[ \frac{9\beta kT}{4\pi p_c(z)\beta Gm} \right]^{1/2}$$ (14)

and the mass-temperature relation

$$M_s = \frac{9}{2\sqrt{\pi}p_c(z)\delta} \left( \frac{\beta kT}{Gm} \right)^{3/2},$$ (15)

where $\delta$ is the over-density parameter defining a radius $R_s$, and an enclosed mass $M_s$, such that the mean enclosed mass density $\delta \equiv \rho_c R_s/\rho_m R_m^3$ is the present critical mass density, $\rho_m \equiv \rho_0 H_0^2/(8\pi G)$, describes the evolution of the Hubble constant, and $\Omega_m$ is the matter fraction of the Universe.

The MCXC catalogue provides $M_{500}$ for all clusters, but for only some of them, so we use the mass–temperature relation (15) to estimate $T_d$, the downstream temperature. The downstream particle number density,

$$n_d = (1 - \beta) \frac{f_0 p_c(z)}{m} \delta_s,$$ (16)

is estimated by assuming that the total baryon mass $f_0 M_s$ is given by the spatial integral of the isothermal $\beta$-profile (11), and quantities with subscript $s$ are evaluated at the shock radius. The magnetic field can then be parameterised by assuming that magnetic energy is a fraction $\xi_B$ of the downstream thermal energy,

$$\frac{B^2}{8\pi} \equiv \xi_B \frac{3}{2} n_d kT_d,$$ (17)

where we assumed that the adiabatic index is $5/3$.

Assuming that a fraction $\xi$ of the thermal energy density downstream of the shock is deposited in CREs of spectral index $p$, solving for the steady-state CRE distribution yields

![Figure H5](image1.png)

**Figure H5.** Same as Fig. H1, but varying the point source masking radius: $0'1$ (dashed orange curve), $0'3$ (default; solid blue), $0'6$ (dotted green), and $1'$ (dash-dotted red).

![Figure H6](image2.png)

**Figure H6.** Same as Fig. H1, but varying the radial bin size: $\Delta r \simeq 0.167$ (dashed orange curve), $\Delta r = 0.25$ (default; solid blue), and $\Delta r = 0.5$ (dotted green). A bin size $\Delta r = 0.5$ with a 0.25 offset from the origin (dot-dashed red) is also shown.
For the steady state CRE spectral index $p + 1$, the luminosity of synchrotron radiation is \cite[e.g.,][]{Rybicki1986}
\begin{equation}
L_\nu = \alpha_e h\nu C\Phi (p) \left( \frac{3\nu_B}{\nu} \right)^{1 + \frac{p}{2}} \text{erg s}^{-1},
\end{equation}
where $\alpha_e = e^2/(hc)$ is the fine-structure constant, $h = 2\pi\hbar$ is Planck’s constant, $e$ is the electron charge, $m_e$ is the electron mass, $\nu_B = eB/(2\pi mc)$ is the cyclotron frequency, $\Gamma(y)$ is the gamma function, and the $p$-dependent numerical factor
\begin{equation}
q(p) \equiv \frac{\Gamma(\frac{p+1}{2})}{\sqrt{p+2}} \frac{\Gamma(\frac{p+3}{2})}{\Gamma(\frac{p+5}{2})}.
\end{equation}
The dependence $\Phi \equiv \sin^{(p+2)/2}\phi$ upon the pitch angle $\phi$ is isotropically averaged to give
\begin{equation}
\langle \Phi \rangle = \frac{\sqrt{\pi} \Gamma(\frac{p+3}{2})}{2 \Gamma(p+2)}.
\end{equation}
The specific energy flux can now be estimated as
\begin{equation}
F_\nu = \frac{L_\nu}{4\pi d_L^2} = \frac{\theta^2_0 L_\nu}{4\pi(1+z)^4 R_s^2},
\end{equation}
where $d_L$ is the luminosity distance. The specific brightness is then given by
\begin{equation}
I_\nu = F_\nu f(\tilde{\theta}) = \frac{F_\nu \tilde{f}(\tilde{\theta})}{\theta^2_0} = \frac{L_\nu \tilde{f}(\tilde{\theta})}{4\pi(1+z)^4 R_s^2},
\end{equation}
where $\tilde{\theta} \equiv \tilde{\theta}/\theta_s$. The functions $f(\theta)$ or equivalently $\tilde{f}(\tilde{\theta}) = \theta^2 f(\theta)$ describe the brightness distribution across the cluster, and are normalised to unity,
\begin{equation}
\int f(\theta) d\Omega = 2\pi \int_0^\infty \tilde{f}(\tilde{\theta}) \tilde{\theta} d\tilde{\theta} = 1.
\end{equation}
In the limit of emission in a thin ring in the plane of the sky, one has $\tilde{f} = \delta(\tilde{\theta} - 1)/2\pi$. The function $\tilde{f}(\tilde{\theta})$ was derived for the projection of a finite shell with an evolving CRE spectrum in \cite{Keshet2018}; in the thin shell limit, the result reduces to $\tilde{f} = (1 - \theta^2)^{-1/2}\Theta(1 - \theta^2)/(2\pi)$.

For completeness, we combine Eqs. (14)–(15) to derive the flux in an explicit form. In the strong shock limit $p = 2$, we find
\begin{equation}
\nu F_\nu \simeq 1 - \frac{\beta}{12\beta^2} \left( \frac{\pi}{6} \right)^{1/2} \left( \frac{f_0 G}{} \right) \frac{\delta_0 H_0 H_{\delta}^{1/2}}{\alpha_{\text{cmb}}^2 \ln(\gamma_2/\gamma_1)} \frac{\xi_e \xi_B M_s^{4/3} \theta_0^{2/3}}{(1+z)^4} \frac{M_s^{4/3} \theta_0^{2/3}}{(1+z)^4} \text{erg s}^{-1} \text{ cm}^{-2},
\end{equation}
where we defined $\delta_{100} \equiv \delta_0/100$ and $\theta_{0.2} \equiv \theta_{0.0}/0.2$, and in the second line we took the typical $\beta = 2/3$. For an arbitrary
p \neq 2$, we obtain
\begin{align}
\nu F_\nu \simeq & \frac{(1 - \beta \frac{\delta m}{m c}) (p - 2) \Gamma \left(\frac{p}{2} + 1\right) \Gamma \left(\frac{p}{2} + 1\right) \frac{\delta m}{m c} (p - 1) \Gamma \left(\frac{p}{2} + 1\right)}{(\beta/3) \frac{\delta m}{m c} \Gamma \left(\frac{p}{2} + 1\right) (p - 1) \Gamma \left(\frac{p}{2} + 1\right)} \\
& \times \left(\frac{e}{m c}\right) \frac{\delta m}{m c} (f_{\nu c} G)^{1/2} r_0^2 H_0 \xi_{13/2} \frac{\delta m}{m c} \xi_{B}^{1/2} \\
& \times \frac{\delta m}{m c} \frac{H_{13/2}}{\nu^{2/2} (1 + z)^4} \nu_m \xi_{B}^{1/2} \\
\simeq & 7.4 \times 10^{-12} (p - 2) \Gamma \left(\frac{p}{2} + 1\right) \Gamma \left(\frac{p}{2} + 1\right) \frac{\delta m}{m c} (p - 1) \Gamma \left(\frac{p}{2} + 1\right) \\
& \times \frac{0.047^2 (p + 2) (p - 1) \Gamma \left(\frac{p}{2} + 1\right)}{(1 + z)^4 (\gamma_{12} - \gamma_{22}) \nu^{12/2} \nu_m \xi_{B}^{1/2}} \text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1},
\end{align}

(119)

again taking $\beta = 2/3$ in the second equality.

The last result can be combined with Eq. (116) in the form
\begin{align}
\nu I_\nu \simeq & 2.1 \times 10^{-6} (p - 2) \Gamma \left(\frac{p}{2} + 1\right) \Gamma \left(\frac{p}{2} + 1\right) \frac{\delta m}{m c} (p - 1) \Gamma \left(\frac{p}{2} + 1\right) \\
& \times \frac{0.047^2 (p + 2) (p - 1) \Gamma \left(\frac{p}{2} + 1\right)}{(1 + z)^4 (\gamma_{12} - \gamma_{22}) \nu^{12/2} \nu_m \xi_{B}^{1/2}} \text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1},
\end{align}

(120)

which is used in Eq. (16). In the strong shock limit, this result reduces to
\begin{align}
\nu I_\nu \simeq & 2.9 \times 10^{-7} \frac{M_{14}^3}{(1 + z)^4} \ln(\gamma_{12}/\gamma_{22}) \nu^{12/2} \nu_m \xi_{B}^{1/2} \text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1},
\end{align}

(121)

J: MUTUAL EFFECTS OF CENTRAL AND VIRIAL MODELLING

For better accuracy, the results in §6.3.1 and 6.3.2 were obtained by simultaneously modelling the central and virial signals. However, as the spatial overlap between the two signals is limited, our qualitative conclusions are robust to the analysis details. In particular, modelling each signal separately yields similar results, with only a small enhancement in the individual signal normalisations.

For example, one can approximately neglect the effect of the central signal at large radii beyond, say, $\tau = 1.5$. We thus fit the virial shock model (in the strong shock limit $p = 2$) alone, focusing only on data outside $\tau \simeq 1.5$. The high-frequency nominal planar model gives
\begin{align}
\tau_s = & 2.40^{+0.08}_{-0.08} \text{,} \label{eq:tau_s}
\end{align}
consistent with Eq. (19), and
\begin{align}
\tilde{n}_B \xi_B = & 3.99^{+0.62}_{-0.62} \times 10^{-4} \text{,} \label{eq:nB_xi_B}
\end{align}
\sim 40\% higher than Eq. (20), with TS $\simeq 41.5$ (6.1$\sigma$ for $N = 2$). The same model in the low-frequency channels yields
\begin{align}
\tau_s = & 2.50^{+0.09}_{-0.09} \text{,} \label{eq:tau_s_low}
\end{align}
consistent with Eq. (21), and
\begin{align}
\tilde{n}_B \xi_B = & 2.35^{+0.36}_{-0.36} \times 10^{-4} \text{,} \label{eq:nB_xi_B_low}
\end{align}
\sim 20\% higher than Eq. (22), with TS $\simeq 42.4$ (6.2$\sigma$ for $N = 2$). Similar results are obtained for the shell model, with normalisation enhancements varying between 20\% and 40\% for different model variants.

An alternative, more accurate possibility is to freeze the central model parameters to the values inferred in §F, and then model only the virial signal. The resulting parameter values are consistent with those inferred above, within the statistical uncertainties. In particular, in the high-frequency channel, $\tau_s$ shifts by $\lesssim 2\%$ (5\%) and $\tilde{n}_B \xi_B$ decreases by $\lesssim 20\%$ (10\%) for the planar (shell) model. In low-frequency channels, $\tau_s$ shifts by $\lesssim 1\%$ (1\%) and $\tilde{n}_B \xi_B$ diminishes by $\lesssim 10\%$ (25\%) for the planar (shell) model.

The opposite effect, of the virial excess biasing the central model parameters, can be inferred for the cleared data by comparing the best-fitting parameters obtained with (Table D1) and without (Table F1) the virial shock modelling. When modelling the central and virial signals simultaneously, we obtain central model parameters comparable to those obtained while neglecting the virial signal in §F, with central flux normalisations smaller than those of Table F1 by $\lesssim 10\%$ (15\%) for the planar and (shell) model in both high and low frequencies, and a central spectrum
\begin{align}
\alpha_c = 1.20^{+0.27}_{-0.26} \text{,} \label{eq:alpha_c}
\end{align}
consistent with Eq. (F3).

K: SHOCK PARAMETER DEGENERACY

In the limit of a strong shock, where $p = 2$, the radio signal provides a measurement of $\tilde{n}_B \xi_B \xi_B$. The degeneracy between the three terms in this product can be broken under some assumptions on the gas profile. If the parameters of the isothermal $\beta$-model are known, and the model can be extrapolated (with some correction) to the cluster periphery, then the density $\rho_d$ and temperature $T_d$ downstream of the virial shock can be inferred. The downstream velocity is then
\begin{align}
v_d \simeq \left(\frac{k_B T_d}{3 m}\right)^{1/2} \text{,} \label{eq:v_d}
\end{align}
so the accretion parameter can be estimated from the definition (111),
\begin{align}
\tilde{m} \simeq \frac{4\pi R_d^2 \rho_d m_{\text{v, d}}}{f_0 m H_0 H} \text{,} \label{eq:m_d}
\end{align}
facilitating a measurement of the product $\tilde{n}_B \xi_B$. A similar approach was used when combining the galaxy distribution around A2319 (Hurier et al. 2019) with the product $\tilde{n}_B \xi_B$ inferred from the leptonic signal, giving $\xi_B \simeq 0.5\%$ and $\tilde{m} \simeq 1.1$ (Keshet & Reiss 2018).

One way to break the degeneracy between $\xi_B$ and $\xi_B$ is to assume equipartition, i.e. $\xi_B = \xi_B$. An alternative with no additional assumption (in the one-zone model) is to use the $\tilde{n}_B \xi_B \simeq 0.6\%$ estimates based on modelling the leptonic signals from the virial shock (Keshet et al. 2017; Keshet & Reiss 2018; Reiss & Keshet 2018). After this degeneracy is broken, giving a measure of $\xi_B$, one may equivalently estimate $B$, as $n_d$ and $T_d$ were already determined; see Eq. (I7). For clusters with available $\beta$-model parameters, typically estimated from the X-ray distribution at $r < R_{500}$, we extrapolate the model to the virial shock radius but assume that by $r_s$, both
temperature and density are diminished by an additional factor $f_\beta \simeq 1/3$. Alternatively, and without a $\beta$-model, we may estimate $B$ by assuming a $\beta = 2/3$ model and substituting $T_d$ using Eq. (15) and $n_d$ using Eq. (16) (no $f_\beta$ correction needed), in which we invoke hydrostatic equilibrium and assume $r \gg r_c$.

L: CLUSTER SAMPLE

The clusters in our sample are listed in Table L1, based on the MCXC catalogue supplemented by the estimated $\theta_{500}$ values and the dynamical state of the cluster (when available).

M: 2D SIGNIFICANCE MAP

Figures M1–M3 demonstrate 2D significance maps in a large, $|\tau| < 8$ radial range, for different choices of the declination cut used for cluster selection: $\delta > 20^\circ$ (nominal), $10^\circ$, and $30^\circ$.

Figure M1. Same as Fig. 5, but with a larger $\tau$ range.

Figure M2. Same as Fig. M1, but with a more relaxed declination cut of $\delta > 10^\circ$ (leaving 59 clusters).

Figure M3. Same as Fig. M1, but with a more restrictive declination cut, $\delta > 30^\circ$ (leaving 29 clusters).
Table L1. Cluster sample.

| MCXC Name | Other Name | $b$ | $l$ | $M_{500}$ | $R_{500}$ | $\theta_{500}$ | $L_X$ | $z$ | Dyn. state |
|-----------|------------|-----|----|--------|--------|---------|------|----|-----------|
| MCXCJ1204.1+2020 | NGC4066 | 77.25 | 242.53 | 0.32 | 0.48 | 0.26 | 4.880 | 0.0252 |
| MCXCJ2336.5+2108 | A2626 | -38.44 | 100.45 | 1.81 | 0.84 | 0.21 | 262.87 | 0.0565 |
| MCXCJ1109.7+2145 | A1177 | 66.28 | 220.43 | 0.46 | 0.54 | 0.24 | 9.028 | 0.0319 |
| MCXCJ1122.3+2419 | HCG51 | 69.76 | 215.92 | 0.25 | 0.44 | 0.24 | 3.302 | 0.0258 |
| MCXCJ0036.5+2544 | ZwCl193 | -37.01 | 118.75 | 0.49 | 0.55 | 0.22 | 9.881 | 0.0341 |
| MCXCJ1348.8+2655 | A1795 | 77.18 | 33.82 | 5.53 | 1.22 | 0.28 | 547.807 | 0.0622 |
| MCXCJ0058.9+2611 | RXJ0058.9+2657 | -35.89 | 124.99 | 0.82 | 0.65 | 0.23 | 23.388 | 0.0451 |
| MCXCJ1522.4+2742 | A2065 | 56.62 | 42.84 | 3.51 | 1.05 | 0.28 | 262.787 | 0.0723 |
| MCXCJ1658.0+2751 | AWM5 | 35.93 | 49.02 | 0.71 | 0.62 | 0.26 | 18.244 | 0.0357 |
| MCXCJ1359.2+2758 | A1831 | 74.95 | 40.07 | 0.71 | 0.62 | 0.26 | 18.244 | 0.0337 |
| MCXCJ1206.6+2811 | NGC4104 | 80.02 | 204.27 | 0.42 | 0.52 | 0.26 | 7.641 | 0.0283 |
| MCXCJ0228.1+2811 | RXJ0228.2+2811 | -30.01 | 147.57 | 0.66 | 0.61 | 0.24 | 16.276 | 0.0353 |
| MCXCJ0150.7+3305 | A260 | -28.16 | 137.01 | 0.63 | 0.60 | 0.23 | 14.929 | 0.0363 |
| MCXCJ1320.2+3308 | NGC5098 | 78.15 | 204.87 | 0.42 | 0.52 | 0.24 | 7.641 | 0.0283 |
| MCXCJ0110.9+3308 | NGC0410 | -29.54 | 127.63 | 0.22 | 0.42 | 0.32 | 4.025 | 0.0168 |
| MCXCJ0919.8+3345 | A779 | 44.40 | 191.08 | 0.26 | 0.46 | 0.38 | 4.025 | 0.0168 |
| MCXCJ0933.4+3403 | UGC05088 | 47.24 | 191.05 | 0.21 | 0.41 | 0.21 | 2.385 | 0.0269 |
| MCXCJ1334.3+3441 | NGC5223 | 45.67 | 35.82 | 0.67 | 0.61 | 0.27 | 16.810 | 0.0315 |
| MCXCJ0933.4+3403 | A0671 | 33.15 | 192.75 | 1.15 | 0.73 | 0.21 | 41.316 | 0.0503 |
| MCXCJ0919.8+3345 | A0671 | 33.15 | 192.75 | 1.15 | 0.73 | 0.21 | 41.316 | 0.0503 |
| MCXCJ0907.8+4936 | VV196 | 42.12 | 169.27 | 0.40 | 0.51 | 0.22 | 7.641 | 0.0283 |
| MCXCJ0751.3+5012 | UGC04052 | 29.84 | 168.39 | 0.42 | 0.52 | 0.32 | 7.650 | 0.0228 |
| MCXCJ1134.8+4903 | A1314 | 63.56 | 151.79 | 0.46 | 0.54 | 0.22 | 9.025 | 0.0341 |
| MCXCJ0907.8+4936 | VV196 | 42.12 | 169.27 | 0.40 | 0.51 | 0.22 | 7.641 | 0.0283 |
| MCXCJ1715.3+4240 | A2192 | 43.90 | 64.83 | 0.39 | 0.51 | 0.23 | 6.744 | 0.0301 |
| MCXCJ1755.8+6236 | NGC6521 | 30.22 | 91.82 | 0.31 | 0.47 | 0.25 | 4.684 | 0.0266 |
| MCXCJ1736.3+6803 | Zw1745.6+6703 | 32.00 | 98.27 | 0.24 | 0.43 | 0.24 | 3.050 | 0.0248 |

Columns: (1) MCXC cluster name; (2) Other cluster name; (3) Galactic latitude in degrees; (4) Galactic longitude in degrees; (5) Mass within $R_{500}$ in $10^{14} M_\odot$ units; (6) Radius enclosing an overdensity $\delta = 500$ in Mpc; (7) Angular scale of $R_{500}$ in degrees; (8) X-ray luminosity in the 0.1 - 2.4 keV energy band, in $10^{42}$ erg s$^{-1}$, out to the $R_{500}$ radius; (9) Redshift; (10) Dynamical state. R — relaxed; sample has a cooling time less than the Hubble time or is classified as a cool-core cluster; D — disturbed; sample has a cooling time greater than the Hubble time, is classified as an non cool-core cluster or merger, or has an observed radio relics. References: $^a$ — Stewart et al. (1984); $^b$ — Johnson et al. (2011); $^c$ — Laganá et al. (2019); $^d$ — Jones & Forman (1999); $^e$ — Proust et al. (2003); $^f$ — Dupke et al. (2007); $^g$ — Yuan et al. (2015). †: Reported as a cool-core cluster in Hudson et al. (2010).