A maximum-likelihood approach to removing radio sources from observations of the Sunyaev–Zel’dovich effect, with application to Abell 611


Astrophysics, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE

Accepted 2002 June 24. Received 2002 June 7; in original form 2001 March 1

ABSTRACT
We describe a maximum-likelihood technique for the removal of contaminating radio sources from interferometric observations of the Sunyaev–Zel’dovich (SZ) effect. This technique, based on a simultaneous fit for the radio sources and extended SZ emission, is also compared to techniques previously applied to Ryle Telescope observations and is found to be robust. The technique is then applied to new observations of the cluster Abell 611, and a decrement of $-540 \pm 125 \mu$Jy beam$^{-1}$ is found. This is combined with a ROSAT HRI image and a published ASCA temperature to give an $H_0$ estimate of $52^{+24}_{-16}$ km s$^{-1}$ Mpc$^{-1}$.

Key words: methods: data analysis – galaxies: clusters: individual: Abell 611 – cosmic microwave background – cosmology: observations – distance scale – X-rays: galaxies: clusters.

1 INTRODUCTION
This paper is concerned with the subtraction of radio sources that would otherwise contaminate or obliterate detections of the Sunyaev–Zel’dovich (SZ) effect (Sunyaev & Zel’dovich 1970, 1972) towards galaxy clusters. The work described here is in connection with the Ryle Telescope (RT; see e.g. Jones et al. 2001; Grainge et al. 2002b), but the issues are relevant to all centimetre-wavelength SZ observations with interferometers (see e.g. Reese et al. 2000). For a massive cluster at moderate or high redshift, the flux that the RT detects from the SZ effect at 15 GHz is typically $-500 \mu$Jy on its shortest baselines. This is sufficiently faint that radio sources will almost invariably be present with comparable or greater amplitudes. Thus removing the effects of radio sources is an essential step. We describe and compare two past methods of measuring SZ decrements in the presence of sources as well as a maximum-likelihood method. We then apply this to new RT observations of cluster Abell 611 (A611), which we combine with X-ray data to estimate $H_0$. All coordinates are J2000 and, except where otherwise stated, we use an Einstein–de Sitter world model.

2 REMOVING RADIO SOURCES FROM SZ OBSERVATIONS WITH THE RT
As the RT is an interferometer with a wide range of baselines, it can simultaneously measure the extended SZ flux and the fluxes and positions of the small angular size radio sources. Fig. 1 illustrates the variation of SZ flux with baseline for the RT when observing a massive Abell cluster at $z = 0.171$. We have assumed a cut-off in the King model of 10 core radii; in practice, the 6 arcmin primary beam of the RT means that the measured flux is effectively independent of this cut-off. The variation with redshift is slight over the range $z = 0.15–5$, as shown in Grainge et al. (2002b). The SZ effect has effectively been completely resolved out for baselines above $\approx 1.5 \lambda$, and so these ‘long’ baselines can be used to measure sources and so remove their effects from the SZ signal seen on the ‘short’ baselines. The measurements are simultaneous and, of course, at the same frequency, and so the spectral index is unimportant. Variability is unimportant if the telescope configuration does not change (for details see Grainge et al. 1996). By choosing an interferometer configuration such that there are more long baselines than short, it is possible to optimize the observations to achieve good signal-to-noise ratio for the SZ effect without it being dominated by noise from unsubtracted sources.

There are three methods of source subtraction that have been applied to RT data:

(i) using CLEAN, which has been used for the bulk of published SZ measurements from the RT (e.g. Grainge et al. 1996);

(ii) the matrix method, which was first used by Grainge (1996); and

(iii) the FLUXFITTER method, first used by Das (1999), and used here in modified form, based on the maximum-likelihood method used in, for example, very-long-baseline interferometry (VLBI; see e.g. Thompson, Moran & Swenson 1986).

© 2002 RAS
and positions are shown in Table 1, and the noise level was set to
see (for details see Grainge et al. 1996). The point-source
map with a central electron density of $10^4$ m$^{-3}$; this becomes somewhat shorter with projection.

The three methods can be explained and compared with an example
of a simulated data set containing both sources and an SZ effect. The
simulation is of a 54 $\pm$ 12-h long observation of a field at declination
44$^\circ$. The $uv$ coverage is based on a standard Cb configuration for the
RT; in this configuration four aerials are parked on an east–west rail-
track at locations 36, 72, 90 and 108 m from the closest fixed aerial
see (for details see Grainge et al. 1996). The point-source fluxes and
positions are shown in Table 1, and the noise level was set to
7 mJy visibility$^{1/2}$, corresponding to 200 $\mu$Jy d$^{-1/2}$, as expected
for a standard RT observation with five aerials. The SZ decrement
is based on Abell 2218, and is modelled as an isothermal ellipsoid
with a King electron density profile (King 1972) at the centre of the
map with a central electron density of $10^4$ m$^{-3}$, $\beta = 0.65$, a
temperature of $7.8 \times 10^7$ K and core radii of 60 and 40 arcsec on the
sky and 49 arcsec ($= \sqrt{60 \times 40}$) along the line of sight. The central
temperature decrement for this cluster is 0.82 mK and, as observed
with the RT, the cluster gives $-660 \mu$Jy on the shortest baseline.

2.2 The CLEAN method
For this method, a dirty (i.e. un CLEANed) map of all the baselines
longer than 1.5 $\lambda$ is produced within AIPS (Graisen 1994). Since we
require maximum sensitivity, natural weighting has been used, i.e.
we have not weighted the data to reduce the sidelobes of the syn-
thesized beam. The map is CLEANed in the standard way by placing
CLEAN boxes around the obvious sources. After deconvolution, the
source fluxes are measured, using the AIPS verb MAXFIT, which in-
terpolates the position and value of the maximum flux density. The
measured fluxes are then removed from the visibility data using the
task UVSUB.

The sources and the positions found in the simulated data set
are listed in Table 2, and the CLEANed map is shown in Fig. 2. The
noise in the map due to the system temperature is 38 $\mu$Jy beam$^{-1}$.
The six model sources are labelled; note that only four brightest
were found.

2.1 The test data
The three methods can be explained and compared with an example
of a simulated data set containing both sources and an SZ effect. The
simulation is of a 54 $\times$ 12-h long observation of a field at declination
44$^\circ$. The $uv$ coverage is based on a standard Cb configuration for the
RT; in this configuration four aerials are parked on an east–west rail-
track at locations 36, 72, 90 and 108 m from the closest fixed aerial
see (for details see Grainge et al. 1996). The point-source fluxes and
positions are shown in Table 1, and the noise level was set to
7 mJy visibility$^{1/2}$, corresponding to 200 $\mu$Jy d$^{-1/2}$, as expected
for a standard RT observation with five aerials. The SZ decrement
is based on Abell 2218, and is modelled as an isothermal ellipsoid
with a King electron density profile (King 1972) at the centre of the
map with a central electron density of $10^4$ m$^{-3}$, $\beta = 0.65$, a
temperature of $7.8 \times 10^7$ K and core radii of 60 and 40 arcsec on the
sky and 49 arcsec ($= \sqrt{60 \times 40}$) along the line of sight. The central

Table 1. The source fluxes and positions in the simulated data
set. The convention used in this and other tables is: positive
$\Delta$RA is an increase in the RA value, so the source is to the
east (to the left on conventional maps) and positive $\Delta$Dec. is
an increase in Dec., i.e. to the north.

<table>
<thead>
<tr>
<th>Source number</th>
<th>Flux ($\mu$Jy)</th>
<th>Offset from pointing centre (arcsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2960</td>
<td>-10, 10</td>
</tr>
<tr>
<td>2</td>
<td>910</td>
<td>35, 15</td>
</tr>
<tr>
<td>3</td>
<td>255</td>
<td>-50, -40</td>
</tr>
<tr>
<td>4</td>
<td>170</td>
<td>-120, 100</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>0, 0</td>
</tr>
<tr>
<td>6</td>
<td>80</td>
<td>60, -60</td>
</tr>
</tbody>
</table>

Table 2. Fluxes and positions as measured by MAXFIT
after cleaning with four CLEAN boxes. The noise on
each flux is 38 $\mu$Jy.

<table>
<thead>
<tr>
<th>Source number</th>
<th>Flux ($\mu$Jy)</th>
<th>Offset from pointing centre (arcsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3040</td>
<td>-9.6, 9.6</td>
</tr>
<tr>
<td>2</td>
<td>970</td>
<td>34.5, 15.8</td>
</tr>
<tr>
<td>3</td>
<td>270</td>
<td>-49.9, -45.8</td>
</tr>
<tr>
<td>4</td>
<td>170</td>
<td>-120.1, 91.9</td>
</tr>
</tbody>
</table>

Figure 1. The flux density as a function of baseline for a spherical King model
cluster with central electron density $= 10^4$ m$^{-3}$, $\beta = 0.65$, core radius
= 60 arcsec, $T_e = 7.8 \times 10^7$ K, $z = 0.171$ and $H_0 = 50$ km s$^{-1}$ Mpc$^{-1}$, as
observed with the RT (from Grainge 1996). The shortest RT physical baselines
are 870 $\lambda$; this becomes somewhat shorter with projection.

Figure 2. A CLEANed map of the simulated data made with baselines longer
than 1.5 $\lambda$. Source 1 has had 2500 $\mu$Jy of flux already removed to aid
identification of weaker sources. The contour levels are $-55 \times (1, \sqrt{2}, 2)$ $\mu$Jy beam$^{-1}$ (dashed) and $55 \times (1, \sqrt{2}, 2, \sqrt{3}, 4, \ldots)$ $\mu$Jy beam$^{-1}$ (solid). The
beam size full width at half-maximum (FWHM; $45 \times 26$ arcsec $^2$ at 4$^\circ$) is
shown in the bottom left. The noise in the map due to the system temperature
is 38 $\mu$Jy beam$^{-1}$.

© 2002 RAS, MNRAS 337, 1207–1214
A map made with baselines longer than 1.5 kλ was consistent with noise after these four sources had been subtracted. The SZ flux observed was $-675 \pm 72$ mJy, at a position ($-4, -2$) arcsec from the pointing centre. The SZ values were measured with MAXFIT from a map made with the baselines shorter than 1.0 kλ.

A comparison of the sources found (Table 2) and the sources actually in the model (Table 1) shows the how well the CLEAN method works. The most glaring problem is that only four out of the original six sources have been detected. As lower-frequency surveys such as FIRST, NVSS or optical images often provide information about the source environment at 15 GHz, the long-baseline map is rarely used in isolation. As such, a second test was performed, placing the CLEAN boxes as before, but adding small CLEAN boxes around the locations of the other two weaker sources. The positions and fluxes of the six sources are shown in Table 3. The noise on the long-baseline map is 38 µJy beam$^{-1}$. Thus additional radio or optical information leads to the detection of the two faintest sources.

### 2.2.1 Source finding

Source finding is done in a non-linear, iterative manner. A typical source subtraction process may involve many iterations of mapping, running the subtraction algorithm on the sources found, mapping residuals, finding another source and then adding that into the model.

There are tasks in AIPS to search for sources in both the image plane (e.g. IMFIT and JMFIT) and the visibility plane (UVFIT). Unfortunately, these are all limited to fitting up to four Gaussians. When used with the test data, both IMFIT and JMFIT failed to find a physically reasonable fit, i.e. fitting a negative flux source. UVFIT found a reasonable fit, but to its maximum of four sources.

Identification of sources in the presence of sidelobes is also difficult. A high signal-to-noise ratio source is easy to identify, and causes no problems. However, the process becomes subjective around signal-to-noise ratios of 4 to 3.5. This is the point at which the non-Gaussian, correlated statistics in the image plane conspire with the high RT sidelobes (when using a few antennas for source finding) to make source identification more difficult.

Radio sources in the clusters we observe have angular sizes smaller than the maximum resolution used. However, the number of CLEAN components is generally much larger than the number of sources in the field. With the test data, only four sources were identified, but AIPS produced a model with 28 CLEAN components. Since the real sources are evidently point sources, this is over-modelling the data, and potentially biasing. There is also a degree of subjectivity in the placing and sizing of CLEAN boxes.

### 2.3 The matrix method

The matrix method was initially developed by Grainge (1996), and we here describe it and assess its performance. In the matrix method, sources are first identified from an unCLEANed long-baseline map, and MAXFIT used to measure the positions and the fluxes of the sources. The convolution that occurs when observing with an interferometer means that the flux on the dirty map of the $i$th source, $S_{\text{dirty},i}$, is given by

$$
S_{\text{dirty},i} = \sum_{i=1}^{n} S_{\text{sky},i} B_{i,j} P_{/i},
$$

where $S_{\text{sky},i}$ is the true flux on the sky of the $i$th point source, $B_{i,j}$ is a factor due to the synthesized beam that depends on the displacement (on the sky) between the source in question and the $i$th source, and $P_{/i}$ is the primary (envelope) beam attenuation. The value of $B_{i,j}$ is directly measured from the dirty beam produced by HORUS. There is thus a matrix equation linking the measured dirty fluxes with the true sky fluxes, which is solved by inverting the matrix. This method has the advantage over CLEAN in that sources that are measured in the map to be point sources are modelled with one flux and position. With the simulated test data, the four sources identified with the CLEAN method were used. The resulting matrix was

$$
\begin{pmatrix}
2875 \\
473 \\
105 \\
585
\end{pmatrix}
= 
\begin{pmatrix}
1 & 0.13 & -0.03 & 0.16 \\
0.13 & 1 & 0.02 & -0.04 \\
-0.03 & 0.02 & 1 & 0.12 \\
0.16 & -0.04 & 0.12 & 1
\end{pmatrix}
\begin{pmatrix}
S_{\text{sky},1} \\
S_{\text{sky},2} \\
S_{\text{sky},3} \\
S_{\text{sky},4}
\end{pmatrix},
$$

where the vector on the left-hand side of the equation is a measurement of the dirty flux (in µJy) at each point. The value $S_{\text{sky},i}$ is the beam-attenuated flux on the sky of source $x$, using the same labelling as for sources found in the CLEAN method.

The matrix method is linear, an apparent advantage over the CLEAN method. If after the first subtraction attempt some sources are still present, then the additional terms for the matrix can then be measured and the solution recalculated.

Table 4 shows the fluxes found. A comparison with the sources known to be in the model (Table 1) shows that the flux of source 3 is significantly different. There are two reasons for this. First, not all the flux has been subtracted – see Fig. 3. Secondly, the positions used have been determined from the unCLEANed map, by searching for extrema with MAXFIT. Occasionally, the relative positions of sources are such that one source is in the steepest part of a sidelobe of another, and MAXFIT does not find an extremum that corresponds to that source. In this situation, MAXFIT is not used, and the position from the CLEAN method is used. The flux value from the unCLEANed map in that position is then used. This has in fact been done for source 4. The flux value from the unCLEANed map in that position is then used as the input to the matrix method. The reliance on the CLEAN method combined with positional problems and with difficulties automating the process for large numbers of sources means that the ‘matrix method’ has only been applied in simple cases with a few well-separated sources.

### Table 3. Fluxes and positions from the CLEAN method with six CLEAN boxes.

<table>
<thead>
<tr>
<th>Source number</th>
<th>Flux (µJy)</th>
<th>Offset from pointing centre (arcsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3020</td>
<td>$-9.6, 9.5$</td>
</tr>
<tr>
<td>2</td>
<td>950</td>
<td>$34.5, 15.8$</td>
</tr>
<tr>
<td>3</td>
<td>240</td>
<td>$-47.6, -42.0$</td>
</tr>
<tr>
<td>4</td>
<td>180</td>
<td>$-121.6, 98.4$</td>
</tr>
<tr>
<td>5</td>
<td>85</td>
<td>$12, -7$</td>
</tr>
<tr>
<td>6</td>
<td>82</td>
<td>$70, -70$</td>
</tr>
</tbody>
</table>

### Table 4. Source fluxes and positions from the matrix method.

<table>
<thead>
<tr>
<th>Source number</th>
<th>Flux (µJy)</th>
<th>Offset from pointing centre (arcsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2970</td>
<td>$-9, 9.6$</td>
</tr>
<tr>
<td>2</td>
<td>877</td>
<td>$34, 15$</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>$-47, -42.8$</td>
</tr>
<tr>
<td>4</td>
<td>142</td>
<td>$-122, 99.9$</td>
</tr>
</tbody>
</table>
2.4 The FLUXFITTER method

In an attempt to overcome the problems of the CLEAN and matrix methods, the FLUXFITTER algorithm was introduced. The algorithm that it uses is straightforward:

(i) An initial model of the sky is made, using a set of parameters that represent the positions and fluxes of each source, including the SZ decrement. This model is determined from the long-baseline RT maps (either raw or CLEANed) and, for the SZ parameters, from the X-ray image.

(ii) The flux that the RT would observe is calculated for every visibility point.

(iii) The misfit between the model and real \( uv \) data is then calculated as \( \chi^2 \).

(iv) The parameters are then varied to minimize \( \chi^2 \).

The best-fitting parameter values are then used to subtract the radio sources; this is done within AIPS. There are two advantages to this method. It works almost entirely in the aperture plane; only the source identification and approximate position finding is done in the image plane. Working in the aperture plane is preferred because the noise distribution is known to be Gaussian. The second advantage of this method is in point (i): a simultaneous fit to the positions and fluxes of the point sources and the SZ decrement is a clear improvement over either of the previous methods as there is no arbitrary ‘long’ and ‘short’ baseline split, and it allows full use of all visibility data, which increases the signal-to-noise ratio for point-source measurements. Note that this the same algorithm as used by the AIPS task UVFIT, but, unlike UVFIT, is not limited to a small number of sources and includes an SZ decrement.

The initial parameters for the sources are still estimated by iterating the CLEAN method. In a complex situation with both bright and faint sources, the map is CLEANed, and then positions measured. These sources are then subtracted and the subtracted data then mapped again. This loop can be performed many times to estimate the number of sources and their approximate positions and fluxes. In a less complex situation, CLEAN is not used, and sources are approximately subtracted and then the data are remapped. Again, this is just to provide an initial guess for FLUXFITTER.

Currently only the amplitude of the SZ decrement is varied. The other parameters that describe the decrement – position, core radius and \( \beta \) parameter for the cluster – are all fixed in advance from the X-ray data. This is done as the radio data do not constrain well the core radius or \( \beta \) value; with present telescopes, the X-ray measurements constrain the core radius, \( \beta \) value and position much better.

FLUXFITTER was run twice on the simulated test data. The fluxes and positions from the CLEAN method were used as an initial guess for the first run. The fluxes and positions from this run are shown in Table 5; the errors are those reported by FLUXFITTER – see below. After subtraction with \( UVSUB \), the map of baselines greater than 1.5 k\( \lambda \) was consistent with noise. For a second run, the fluxes and positions from the CLEAN method were used, and the positions of the additional two sources were also used. The fluxes and positions from this run are shown in Table 6. This table shows that the two additional sources are detected with good significance. That the overall noise level of 30 \( \mu \)Jy beam\(^{-1}\) is lower than that for the CLEAN method is not surprising as all the baselines are being used in the determination of the fluxes and positions. After subtracting the six reported sources, a map of baselines shorter than 1.0 k\( \lambda \) shows a decrement of flux \(-700 \pm 65 \) mJy at an offset (9, 4) arcsec. FLUXFITTER itself finds a central decrement of 0.87 \( \pm 0.08 \) mK, close to the set value of 0.82 mK.

As an additional check, a third run of FLUXFITTER was performed, and the four sources found from the CLEAN method and two random points were used as the initial guess. In this case, FLUXFITTER reported that both of the two random ‘sources’ had fluxes below the noise level and very large positional errors.

FLUXFITTER also reports error bounds. As Figs 4 and 5 show, the \( \chi^2 \) contours are elliptical and oriented along the variable axes, which shows that the parameters are independent. The error on each parameter is calculated by finding the parameter values at which the reduced \( \chi^2 \) increases by 1.

The error-bound reporting was checked by simulating a point source with differing signal-to-noise ratio. Some 500 observations

---

Table 5. Source positions and fluxes as reported by FLUXFITTER using the CLEAN model as an initial guess. The formal error on each flux is 30 \( \mu \)Jy.

<table>
<thead>
<tr>
<th>Source number</th>
<th>Flux (( \mu )Jy)</th>
<th>Offset from pointing centre (arcsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3031</td>
<td>(-9.5 \pm 0.1, 9.6 \pm 0.2)</td>
</tr>
<tr>
<td>2</td>
<td>966</td>
<td>(34.5 \pm 0.3, 15.3 \pm 0.5)</td>
</tr>
<tr>
<td>3</td>
<td>208</td>
<td>(-47.5 \pm 1.2, -42.8 \pm 2.8)</td>
</tr>
<tr>
<td>4</td>
<td>184</td>
<td>(-122.3 \pm 1.6, 98.2 \pm 2.8)</td>
</tr>
</tbody>
</table>

Table 6. Source positions and fluxes as reported by FLUXFITTER using additional information. The formal error on each flux is 30 \( \mu \)Jy.

<table>
<thead>
<tr>
<th>Source number</th>
<th>Flux (( \mu )Jy)</th>
<th>Offset from pointing centre (arcsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3024</td>
<td>(-9.9 \pm 0.1, 9.9 \pm 0.2)</td>
</tr>
<tr>
<td>2</td>
<td>970</td>
<td>(34.9 \pm 0.5, 15.6 \pm 0.7)</td>
</tr>
<tr>
<td>3</td>
<td>212</td>
<td>(-47.8 \pm 1.8, -40.8 \pm 3.6)</td>
</tr>
<tr>
<td>4</td>
<td>155</td>
<td>(-121.2 \pm 2.6, 98.8 \pm 4.1)</td>
</tr>
<tr>
<td>5</td>
<td>140</td>
<td>(7.9 \pm 3.8, -3.1 \pm 4.2)</td>
</tr>
<tr>
<td>6</td>
<td>132</td>
<td>(67.4 \pm 3.3, -64.8 \pm 4.6)</td>
</tr>
</tbody>
</table>

© 2002 RAS, MNRAS 337, 1207–1214
of a single point source were simulated; the signal-to-noise ratio was kept constant for groups of 10 simulations, and the position was held constant for all the simulations. The visibilities were simulated with Gaussian noise. The known position was then fed to FLUXFITTER as an initial guess, and the best-fitting position and flux recorded. It was found that the quoted error bar does enclose the position for 67 per cent of the simulations. It was also found that the uncertainty of the position, that is the size of $\sigma$, varies as the inverse of the signal-to-noise ratio. This result is shown in Fig. 6. Note that this relation holds down to very low signal-to-noise ratios. The result is useful for determining whether a tentative source found with the RT at low signal-to-noise ratio has a position coincident with higher-significance data, for example from NVSS or POSS.

### 2.4.1 Possible improvements

Source recognition is clearly the biggest problem that still remains. It is the only step that is still performed in the map plane rather than the aperture plane. There are computational issues involved here: producing a map and identifying sources ‘by hand’ is possible and fairly cheap in computer time, but the noise in the map plane is non-Gaussian. Minimizing the misfit between the data and a given number of sources is also cheap; but allowing the number of sources to vary vastly increases the complexity of the problem and the time required. It is possible that more advanced minimizing techniques such as simulated annealing (see e.g. Press et al. 1993) or using massive inference techniques will make this possible and robust in the aperture plane.

### 2.5 Comparison of results

Tables 3, 4 and 6 show the results of three different methods for source fitting. For both the CLEAN and FLUXFITTER methods, the results with six sources are considered. The positions and fluxes put into the model are shown in Table 1. The resultant position and depth of the SZ decrements after subtraction are shown in Table 7. The parameters are all measured from dirty maps, made with baselines shorter than 1 k$\lambda$. Note that the matrix method has a less deep SZ flux density, and that the FLUXFITTER and CLEAN method values are statistically consistent with the expected value for this model cluster.

#### Table 7. Parameters for the SZ decrements resulting from the simulated data.

<table>
<thead>
<tr>
<th>Method</th>
<th>SZ flux (µJy)</th>
<th>Position (arcsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLEAN</td>
<td>$-675 \pm 72$</td>
<td>$-4, -2$</td>
</tr>
<tr>
<td>Matrix</td>
<td>$-500 \pm 90$</td>
<td>$18, 0$</td>
</tr>
<tr>
<td>FLUXFITTER</td>
<td>$-700 \pm 65$</td>
<td>$9, 4$</td>
</tr>
</tbody>
</table>
(−660 \mu\text{Jy beam}^{-1}). Also, the positions of the SZ decrement are fully consistent with the model as the beam size is around 180 arcsec. It is not surprising that the flux of the SZ decrement from the matrix method is less deep as the central 100-\mu\text{Jy source has not been subtracted. Also, the matrix method does an incomplete subtraction of the sources it does find, resulting in more contamination of the SZ signal. Note that this does not imply that the matrix method will always give a lower SZ decrement if the source subtraction is incomplete.

All three methods benefit from prior knowledge of the source distribution on the sky. This can be estimated from looking at lower-frequency surveys such as NVSS or FIRST. Most falling-spectrum sources, i.e. with \alpha > 0 (where \alpha \propto \nu^{-\alpha} where \nu is the observing frequency and \alpha is the spectral index) will be detected in NVSS and/or FIRST. However, as shown by Cooray et al. (1998) in clusters and Taylor et al. (2001) generally, there are rising-spectrum sources, i.e. with \alpha < 0, that are present at 15 GHz and not detected in NVSS and FIRST.

3 OBSERVATIONS OF ABEll 611

3.1 X-ray

Abell 611 is a cluster at z = 0.288 (Crawford et al. 1995) originally identified by Abell (1957). It has a 0.1–2.4 keV luminosity of 8.63 \times 10^{44} \text{W} (Böhringer et al. 2000), with a temperature of 7.95^{+0.58}_{-0.72} \times 10^{7} \text{K} (White 2000). White derived this value from a 57-ks ASCA exposure by considering both a single-phase and two-phase cooling model. The temperature values found for the bulk of the gas are statistically equivalent, and a mass deposit rate of 0.177 M\odot \text{yr}^{-1} was found for the cooling model.

The 17-ks ROSAT HRI observation from 1996 April is shown with 8-arcsec binning in Fig. 7. The image contains two bright pixels, which, on comparison with the POSS image, are coincident with a large galaxy. These pixels are ignored whilst fitting a model to this observation. Inspection of the plots in Boese (2000) show that the point spread function (PSF) of the HRI has broad wings. We calculate that 15 per cent of the flux observed in these two pixels could be spread into these wings. This would bias the fitting by adding additional counts to what is assumed to be cluster-only emission. However, the 15 per cent in the wings adds, in the worst case, less than 1 per cent to the emission of the cluster. This is small in comparison to almost all the other errors in the measurement, and so we ignore it here. We follow the procedure discussed in Grainge et al. (2002a) and calculate an X-ray emissivity constant of 1.29 \times 10^{-59} \text{count s}^{-1} \text{from 1 m}^{3} of 7.95 \times 10^{7} \text{K gas of electron density 1 m}^{-3} at a luminosity distance of 1 \text{Mpc}, assuming a metallicity of 0.21 \pm 0.07 solar and an absorbing H column of 4.8 \times 10^{24} \text{m}^{-2}.

The best-fitting model parameters were \beta = 0.59, core radii of 26 and 24 arcsec with a position angle of the major axis of 101° and a central electron density of n_{0} = 11.6 \times 10^{3} \text{m}^{-3} [assuming a core radius along the line of sight of 25 = (24 \times 26)^{1/2} \text{arcsec and H_{0} = 50 \text{km s}^{-1} \text{Mpc}^{-1}}]. There is a degeneracy between the core radii fitted and \beta but this has no significant effect on H_{0} (see Grainge et al. 2002b; Jones et al. 2001).

3.2 RT observations

A611 was observed for 16 sets of 12 h between 1994 November and 1995 January with the RT in configuration Ch. Flux and phase calibration and overall data reduction strategy are described in Grainge et al. (2002a). Three days of data – taken in bad weather – were rejected after examining the 1-d maps and noise levels. A map of the combined 13 d of data using baselines longer than 1.5 k\lambda from prior knowledge of the source position was decontaminated using UVSUB, with only one source visible, with flux 299 \mu\text{Jy} at RA 8^h0^m57^s, Dec. +36°34′0″. This source was removed with UVSUB, using the flux and position from the dirty map. A long-baseline map of the subtracted data was consistent with noise, with no other sources in the field.

Table 8 lists the sources found in the FIRST catalogue around the pointing centre for A611. The NVSS catalogue contains no sources in this region. Neither of the two FIRST sources is detected at 15 GHz and the source that is present at 15 GHz is not detected at lower frequencies.

FLUXFITTER was then run using the X-ray data to provide a model of the SZ decrement and using all the baselines. Again, the initial guess was defined by the 1.5-k\lambda-only fitting. The source was found to be at RA 8^h0^m57^s1 ± 0.9, Dec. +36°33′5″±8, with a flux of 188 ± 65 \mu\text{Jy, which is lower than the long-baseline-only values. As the angular size of A611 is small, it is likely that the SZ signal was contaminating the ‘long’-baseline map. Note that this source is not detected in the FIRST survey, and so in this case prior knowledge from a lower-frequency survey has not helped. Fig. 8 shows a CLEANed map of baselines shorter than 1 k\lambda, and only one source was visible, with flux 299 \mu\text{Jy} at RA 8^h0^m57^s1, Dec. +36°34′0″. This source was removed with UVSUB, using the flux and position from the dirty map. A long-baseline map of the subtracted data was consistent with noise, with no other sources in the field.

Table 8. The radio sources in the FIRST catalogue within 400 arcsec of the RT pointing centre for A611.

<table>
<thead>
<tr>
<th>RA</th>
<th>Dec.</th>
<th>Flux (\mu\text{Jy})</th>
</tr>
</thead>
<tbody>
<tr>
<td>8^h1^m20^s248</td>
<td>36°5′9″3</td>
<td>1.25 ± 0.13</td>
</tr>
<tr>
<td>8^h0^m54′9″48</td>
<td>36°9′0″1</td>
<td>1.10 ± 0.13</td>
</tr>
</tbody>
</table>
3.3 $H_0$ determination

From the source subtracted data set and the X-ray parameters of the cluster, it is possible to estimate $H_0$, as described in Grainge et al. (2002b). The likelihoods for different $H_0$ values are calculated and are shown in Fig. 9. The best-fit $H_0$ for A611 is then $52^{+34}_{-28}$ km s$^{-1}$ Mpc$^{-1}$. The error quoted is due to the noise in the SZ measurement, and does not include any of the other sources of error in the determination. The additional sources of error are described fully in Grainge et al. (2002b). For A611, the error from the SZ measurement is by far the most important, and the final $H_0$ value is $52^{+24}_{-16}$ km s$^{-1}$ Mpc$^{-1}$ for an Einstein–de Sitter world model. This corresponds to an angular diameter distance of $1064^{+473}_{-336}$ Mpc at the redshift of A611, $z = 0.288$. Assuming a world model with $\Omega_\Lambda = 0.7$ and $\Omega_m = 0.3$, $H_0 = 59^{+27}_{-18}$ km s$^{-1}$ Mpc$^{-1}$ from this cluster.

4 CONCLUSIONS

The problem of radio source contamination in interferometric SZ observations and methods to remove it have been investigated, demonstrating the following.

(i) The non-linear CLEAN method can work well, but does not use all the available information and can overcomplicate the problem.

(ii) The matrix method, though linear, fails in typical situations such as the one simulated here. The failure is mainly due to the high sidelobes from the Ryle Telescope: they make it difficult to determine accurate positions, and so fluxes, for sources close to each other on a map.

(iii) FLUXFITTER uses all the available information and produces the simplest model. It solves simultaneously for sources and SZ decrement, and it works with the visibilities, where noise is known to be Gaussian, rather than in the map plane, where the noise is correlated.

(iv) All three techniques suffer from the problem of source identification, which is currently performed in the image plane where the noise characteristics are complex. Source identification can be aided by prior information, for example from lower-frequency surveys.

(v) The positional uncertainty, as determined in the aperture plane, is found to vary as uncertainty/arcsec $\propto$ (signal/noise)$^{-0.996\pm0.01}$ even at signal-to-noise ratios below nominal detection limits. The constant of proportionality will be a function of the interferometer used.

(vi) Observations of the cluster A611 with the Ryle Telescope give a 4.3$\sigma$ detection of an SZ decrement, and combination with X-ray data gives an estimate of the angular diameter distance to the $z = 0.288$ cluster of $1064^{+473}_{-336}$ Mpc. This corresponds to an $H_0 = 52^{+24}_{-16}$ km s$^{-1}$ Mpc$^{-1}$, assuming an Einstein–de Sitter cosmology, and $59^{+27}_{-18}$ km s$^{-1}$ Mpc$^{-1}$ using $\Omega_\Lambda = 0.7$ and $\Omega_m = 0.3$.

ACKNOWLEDGMENTS

We thank the staff of the Cavendish Astrophysics group who ensure the continued operation of the Ryle Telescope. We are grateful to the referee, Dr Patrick Leahy, for useful comments. Operation of the RT is funded by PPARC. WFG acknowledges support from a PPARC studentship. We have made use of the ROSAT Data Archive of the Max-Planck-Institut für Extraterrestrische Physik (MPE) at Garching, Germany.

REFERENCES

Greisen E., ed., 1994, AIPS Cookbook. NRAO, Green Bank, WV
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., Lloyd C.,
Rees P., 1993, Numerical Recipes in FORTRAN – the Art of Scientific
Thompson A. R., Moran J. M., Swenson G. W., 1986, Interferometry and

This paper has been typeset from a TEX/LATEX file prepared by the author.