Article

Lightning Interferometry with the Long Wavelength Array

Michael Stock 1,*, Julia Tilles 2, Greg B. Taylor 3, Jayce Dowell 3 and Ningyu Liu 4

Abstract: The Long Wavelength Array is a radio telescope array located at the Sevilleta National Wildlife Refuge in La Joya, New Mexico, well suited and situated for the observation of lightning. The array consists of 256 high-sensitivity dual polarization antennas arranged in a 100 m diameter. This paper demonstrates some of the capabilities that the array brings to the study of lightning. Once 32 or more antennas are used to image lightning radio sources, virtually every integration period longer than the impulse response of the array includes at least one identifiable lightning emitter, independent of the integration period used. The use of many antennas also allows multiple simultaneous lightning radio sources to be imaged at sub-microsecond timescales; for the flash examined, 51% of the images contained more than one lightning source. Finally, by using many antennas to image lightning sources, the array is capable of locating sources fainter than the galactic background radio noise level, yielding possibly the most sensitive radio maps of lightning to date. This incredible sensitivity enables, for the first time, the emissions originating from the positive leader tips of natural in-cloud lightning to be detected and located. The tip emission is distinctly different from needle emission and is most likely due to positive breakdown.

Keywords: lightning; interferometry; LWA

1. Introduction

Interferometers have been used quite successfully to study lightning since the 1970s [1–3]. These instruments proved to be exceptional tools for the study of lightning due to (1) their ability to capture the whole lightning event, rather than merely the portion of lightning exiting a cloud; and (2) their excellent temporal resolution and relatively good spatial resolution, which is capable of resolving fast lightning processes propagating at large fractions of the speed of light [4,5]. While the early interferometers were narrowband arrays where the correlation was performed with hardware mixers, the advent of faster digitizers enabled the construction of digital broadband interferometers that used digital Fourier techniques to interfere with the signals. These digital broadband interferometers were able to produce even more detailed pictures of lightning than their narrowband counterparts, especially once their memory buffers were large enough to hold a continuous record of an entire flash [6–8]. The broadband digital interferometers proved to be affordable and easy to build, resulting in many new scientific findings by multiple research groups [9–17].

Nearly all lightning interferometric studies performed to date have used relatively sparse arrays containing only a handful of antennas. Consequently, lightning at microsecond timescales has been universally modeled as a single source smaller than the angular resolution of the array (a point source). Sparse arrays are a limitation primarily driven by cost and complexity; larger arrays are certainly possible and capable of imaging non-point-like sources, as demonstrated by the astrophysical community [18]. In particular, two
radio telescopes already exist that use an appropriate frequency range and configuration for the study of lightning: the Low Frequency Array (LOFAR) and the Long Wavelength Array (LWA).

LOFAR is a very high frequency (VHF) radio telescope located in the Netherlands and surrounding countries, composed of 40 stations in the Netherlands alone, where each station contains 48 high-band (110–250 MHz) and 96 low-band (30–80 MHz) antennas [19]. LOFAR has been used to detect and map lightning VHF sources using the low-band antennas [20]. The many-kilometer baselines between stations were leveraged to be able to map VHF source locations in three spatial dimensions and time. Further, in Scholten et al. [21], Sterpka et al. [22] and Scholten et al. [23], VHF lightning sources were imaged without using a reference feature, with the imaging performed using beam steering. While it should be possible to detect and locate multiple simultaneously emitting lightning VHF sources with LOFAR, all the existing studies model each observation window as a single point emitter [24].

The LWA is a VHF radio telescope designed to study a wide variety of VHF sources, most, but not all, of which are extra-terrestrial [25,26]. Early studies with the LWA (specifically LWA1, located on the Planes of Agustin, NM, near the Very Large Array) successfully imaged lightning sources with the narrowband transient modes of the array [27–29]. However, the transient observing modes available in LWA1 only allowed for either 75 kHz of long-duration or 61 ms of full-band observations [25]. This greatly limited the utility of LWA1 in studying lightning; the very narrow-band observations were not capable of achieving the time resolution desired for lightning studies, while the full-band observation duration was not long enough to capture even the majority of a lightning flash. These limitations in lightning observation were intentionally rectified with the advent of the LWA station at Šešvileta (LWA-SV); with the inclusion of a new transient observation mode capable of recording 39.6 MHz of bandwidth for up to 5 s duration, the LWA-SV is capable of imaging multiple and extended lightning sources at sub-microsecond timescales [30,31].

This paper formalizes the processing technique for the use of the LWA-SV to perform aperture synthesis imaging of lightning at microsecond and sub-microsecond timescales. In addition, we explore the accuracy of the single point source model for lightning at microsecond timescales, investigate the improved sensitivity of synthesis imaging as the number of antennas in the array is increased, and demonstrate that the improved sensitivity allows the VHF emission from positive leader tips to be located.

2. Methods

2.1. Data

The LWA-SV consists of 255 dual horizontally polarized VHF antennas arranged in an array of roughly 100 m diameter, located at the Šešvileta National Wildlife Refuge in La Joya, New Mexico. In addition to the 255 primary antennas, there is a single out-rigger antenna (for 256 total antennas) located approximately 300 m west of the center of the array. Each antenna has an operating bandwidth of 3–88 MHz and a well-characterized gain pattern, which is maximum at the zenith and minimum at the horizon.

The LWA-SV includes a number of transient buffer data collection modes. None of the collection modes are ideal for lightning studies, as the LWA is not capable of recording the raw time-domain signals from all antennas for the duration of a lightning flash. However, the transient buffer frequency domain (TBF), while imperfect, was designed with lightning observations in mind. In this mode, signals from both polarizations of all 256 antennas are continuously sampled at 204.8 MS/s and 8 bits. The data are then converted to the frequency domain with an 8192 sample (40 µs) FFT on the acquisition board. Up to 1584 complex amplitudes (covering 39.6 MHz of bandwidth) can be transferred as 8-bit complex numbers in a 5 s circular buffer. The signals from a subset of the antennas are used to generate a trigger signal for the array. On triggering, an arbitrary pre- and post-trigger length is then written to disk for later analysis. Although the TBF buffer is approximately 5 s in size, the system takes approximately 3 min to rearm for every 1 s of data recorded and
is not capable of back-to-back triggers. The initial acquisitions presented herein were thus limited to ≤250 ms in duration to maximize the likelihood of a nearby lightning capture.

Images can be formed directly from the frequency domain observations recorded in the transient buffer, but these images will be limited to the time resolution of the FFT (40 µs). For lightning observations, a much higher time resolution is desired. To achieve this, the frequency domain observations for each 40 µs window are converted back into the time domain using an inverse FFT for each antenna. After the entire time series record for an antenna is reconstructed, a filter is applied to correct the frequency-dependent and sub-sample portion of cable delay, as well as the variations in the stand heights of the antennas. The resulting time series are then stored to a file for later use and imaging, with the added benefit that the time series data can be partitioned into arbitrary integration periods. The integration periods used for imaging still need to meet the Nyquist criterion and so should be longer than 25 ns.

An example time series from one antenna is shown in Figure 1c as a histogram, with the coloration showing the density of time–amplitude pairs. Depicting the time series in this way allows both the peak and the mean behavior to be seen. At the time of the flash, no other supporting observations were available at the site. Figure 1d shows the nearest available electric field change record, which was provided by the Earth Networks Total Lightning Network (ENTLN) sensor in Los Lunas, 50.9 km to the north. There is good correspondence between the pulses detected by the ENTLN sensor and the VHF data recorded by the LWA, consistent with the known time-keeping accuracy of the LWA-SV of 100 ns.

![Figure 1.](image-url)

**Figure 1.** (a) Interferometric map of the lightning flash shown on the image plane (direction cosines), colored by time. (b) Time–elevation angle plot of the interferometric map of the lightning flash, with the same coloring as panel (a). (c) VHF waveform recorded by one of the LWA antennas; color represents the density of time–amplitude pairs in the VHF data. (d) Electric field change signal recorded by an ENTLN sensor 50.9 km north of the Sevilleta site (timing corrected for propagation delay). Blue circle markers on panels (a,d) show direction to and time of ENTLN IC pulses located for the flash. Division lines are 10° in azimuth or elevation and 25 ms in time.
2.2. Projection Imaging

Once the time series is generated, the data can then be used to image the VHF lightning source using the imaging techniques applied in aperture synthesis (although the actual synthesis part will be less effective since the Earth rotates a negligible amount during a lightning flash). Imaging techniques have been used in previous lightning interferometric studies [32–34], but the methods used are not well documented and differ somewhat from standard imaging techniques [18]. Specifically, the image is formed by projecting the cross-correlation of antenna pairs onto the image plane, instead of aggregating the complex visibility observations in the diffraction plane and transforming back to the image. The study herein uses the same imaging technique as the previous lightning interferometry works, here referred to as ‘projection imaging’. The connection between projection imaging and more traditional interferometry methods has not been well established, so below we provide a brief formalization showing that projection imaging is mathematically equivalent to aperture synthesis.

The image of a source in two dimensions is given by inverting Equation (1-6) of Taylor et al. [18]:

\[
I(l,m) = \frac{1}{\sqrt{1 - l^2 - m^2}} \int \int V(u,v)e^{2\pi i (ul + vm)}dudv
\]

(1)

Here, \(I\) is the image of the incident radiation power at the \(l, m\) coordinates of the image plane (which are simply the direction cosines \(\alpha\) and \(\beta\)), and \(I_m = 1/\sqrt{1 - l^2 - m^2}\) is the modified intensity. In this study, as in previous studies, we will ignore the \(\sqrt{1 - l^2 - m^2}\) term and instead solve for the modified intensity. \(V\) is the complex visibility at the \(u, v\) coordinates of the diffraction plane. In terms of the baseline, \(b\), and the observed radio wavelength, \(\lambda\), the diffraction plane coordinates are \(u = bx/\lambda\) and \(v = by/\lambda\).

For practical radio interferometer arrays, the aperture is not completely sampled because there is a finite number of antenna elements, leading to the discrete form of the imaging equation:

\[
I_m(l,m) = \frac{1}{N} \sum_u \sum_v V(u,v)e^{2\pi i (ul + vm)} .
\]

(2)

where \(N\) is the number of \(u, v\) coordinates that the sum is computed over (this is very similar to Equation (7-3) of Taylor et al. [18]). From these equations, it is evident that the image of the source is a two-dimensional Fourier transform of the complex visibility \(V\).

The complex visibility is measured by correlating the signals arriving at pairs of antennas. For each pair of antennas and each wavelength, the complex visibility is sampled at a single \(u, v\) coordinate. For digital interferometers such as the LWA, \(V\) can be found for all frequencies using a Fourier transform of the input signal:

\[
V_{ij} = \mathcal{F}(E_i)\mathcal{F}(E_j)^* \quad (3)
\]

where \(E_i\) and \(E_j\) are the electric field signals incident on the \(i\)th and \(j\)th antennas, respectively. The image can then be obtained by computing the sums in Equation (2) for every \(l, m\) coordinate of interest by looping over the measured complex visibilities. These sums are commonly computed using a two-dimensional discrete inverse Fourier transform. Because the diffraction plane is sampled sparsely, the resulting image is the true image of the source convolved with the point-spread function determined by the antenna configuration.

Alternatively, and equivalently, we can re-order the sums to express the measured complex visibility in terms of frequency:

\[
I_m(l,m) = \frac{1}{N_b} \sum_{ij} \frac{1}{N_f} \sum_f V_{ij}(f)e^{2\pi i (ul + vm)} ,
\]

(4)

where \(ij\) indicates each antenna pair, and \(f\) is the frequency; similarly, \(N_b\) is the number of baselines, and \(N_f\) is the number of frequency components used in the Fourier transform.
The sum over $f$ is simply the cross-correlation of the signals arriving at the $i$ and $j$ antennas, rotated in the direction indicated by $u_f, v_f$, so we can write Equation (4) as

$$I_m(l, m) = \frac{1}{N_b} \sum_{ij} X_{ij}(\tau_{lm}),$$

(5)

where $X_{ij}$ is the cross-correlation between the $i$ and $j$ antennas, and $\tau_{lm}$ is the time delay for signals arriving from the $l, m$ direction. Conceptually, the way that this imaging method works is that the cross-correlation from each baseline is projected onto the image plane, and they are summed together. This works for any sky geometry in which there is a unique $\tau$ for each baseline for every direction in the sky; the cosine projection used in this paper is especially convenient because the transform is linear.

From Equation (5), it is evident that the amplitude of the image is the VHF energy arriving at the array (or the power if we normalize by the integration period). In general,

$$I_{\text{max}} \leq \frac{1}{N_a} \sum_{\tau} E_i^2$$

(6)

where $N_a$ is the number of antennas. Equality will happen when all of the VHF signal is due to a single point source in the center of the beam and in the absence of any noise.

Figure 2 shows the results of imaging various different lightning signals with 3, 32, and 249 antennas. How the antennas were selected, and why some antennas were omitted, is discussed in Section 3. All images were created using a 100 sample (488 ns) integration window with the beam steered at the center of the image. Panels a–c show the results of imaging a single point source that emerged close to the beginning of the flash. At this point in time, the flash had not grown enough for multiple sources to be likely. The location of the source is quite evident for each of the 3, 32, and 249 antenna images, although the side lobes decrease as the number of antennas increases. When 249 antennas are used to image the source, more short baselines are included in the array, causing the size of the primary lobe to be somewhat larger than the image produced with only 32 antennas. At the same time, including these short baselines also serves to reduce the side lobes of the image.

Many of the integration periods contain only a single source, but, even for single sources, the sensitivity is improved when more antennas are used to compute the image. This is shown in Figure 2d–f, where the image is still a single point source near the beginning of the flash but is less bright than the galactic background. In these cases, it is impossible to even identify that a source is present when imaging with only 3 antennas. The image is clearly improved but still noisy when computed with 32 antennas, and it is quite good when imaged with 249 antennas.

Figure 2g–i shows a case with two point sources of similar brightness. In the 3-antenna image, we can see that a source exists, but it is difficult to determine exactly where. If one of the sources were substantially brighter than the other, the 3-antenna case would start to resemble what is seen in panel a. With 32 antennas, both point sources can be clearly identified, but the side lobes are increased in comparison to panel b.

Figure 2j–l shows a case with many VHF sources. Again, it is impossible to determine whether there is a source at all in the 3-antenna image, although the signal amplitude suggests that there may be. The situation is improved in the 32-antenna image, but not significantly. The confusion added by the side lobes of the image makes it very difficult to identify where the sources are. Even in the 249-antenna image, the side lobes are quite high and not all sources are clearly distinct from one another.
Figure 2. Images of lightning VHF sources using 3, 32, and 249 antennas in the same projection format used in Figure 1a. For each image, the color map is normalized with red showing directions that are positively correlated and blue showing directions that are negatively correlated. (a–c) are of a single source near the beginning of the flash. (d–f) are also of a single source near the beginning of the flash, but a very weak one. (g–i) are of a period with 2 point VHF sources. (j,k) are of a period with many VHF sources. Divisions in all panels are 10°. Peak image power for panels (c,f,i,l) is 28.8, −10.4, 29.7, and 3.1 dB_{G}, respectively. Image time for panels (c,f,i,l) is 87.084, 87.792, 277.660, and 255.370 ms after 00:15:15 UT.

2.3. Deconvolution

The image computed above is the convolution of the antenna array response (or point spread function) with the true image of the sky, referred to as a dirty image. To recover the source image without the contribution from the antenna response, the dirty image is deconvolved. Unlike computing the dirty image, deconvolution does not have a unique solution, particularly for collections of closely spaced sources or extended sources. Indeed, in some cases, it is better to work with the dirty images directly. Even in cases
where deconvolution is desired, there are multiple methods to choose from, each of which performs best in different situations.

In previous interferometric studies of lightning, ‘deconvolution’ was achieved by modeling each integration period as a single radio point source located where the dirty image was brightest. This location was possible to calculate without explicitly computing the image of the source. It also had the benefit of reducing the sequence of images of the lightning sources into a sequence of point locations, or a map, of the flash. Maps in aggregate can appear similar to an image and make the analysis of lightning flashes much easier.

The LWA has enough antennas that it is easily capable of imaging multiple VHF sources in a single integration period. Therefore, instead of modeling each integration period as a single VHF point source, it is modeled as multiple VHF point sources. Finding the locations of these sources proceeds iteratively as follows. First, the brightest pixel of the dirty image is found. A sub-pixel location for the VHF point source is then calculated by fitting a paraboloid to the brightest pixel. Then, a Gaussian distribution is centered at the sub-pixel location and subtracted from the dirty image, leaving a residual image. To determine whether the sub-pixel location is likely to be from lightning, the brightest pixel amplitude is compared to the standard deviation of the residual image. If the brightness is more than a factor of 6 greater than the standard deviation of the residual, the sub-pixel location is appended to the map of the lightning flash; this is the order 1 solution. The factor of 6 threshold used here is a noise filter; lower threshold values will have more solutions and also more noise. The above procedure is repeated for increasingly higher-order solutions until no more sources can be identified. The procedure is very similar to the well-known CLEAN algorithm described in Högbom [35], except that instead of subtracting off a point spread function, we subtract off a Gaussian.

The size of the Gaussian used in this process is related to the angular resolution of the interferometer, $\theta \sim \lambda / D$, where, for a planar interferometer, $\theta$ is a measure on the image plane. A close enough estimate of the angular resolution for deconvolution is obtained if $\lambda$ is taken to be the shortest wavelength and $D$ is the diameter of the array. For the LWA, the diameter is larger in the north–south direction (111 m) than the east–west direction (89 m), so we use an elliptical Gaussian. If the standard deviation of the Gaussian is equal to the angular resolution of the array ($\sigma = \lambda / D$), the Gaussian will be slightly larger than the central lobe of the point spread function. An estimate that is closer to equality is $\sigma = \lambda / \sqrt{2}D$.

Figure 3a–f show a graphical example of deconvolution acting on simulated sources using both of these Gaussians (details of the simulation are provided below). When the simulated sources are far enough apart (as in Figure 3d–f), deconvolution results in accurate locations for both sources. When the simulated sources are closer together than the angular resolution of the interferometer (as in Figure 3a–c), their dirty images merge and artifacts can occur. The situation shown in Figure 3a is close to the worst-case scenario, where the sources have the same amplitude, causing the maximum brightness of the dirty image to occur in between them. In this situation, even using a single point source model for the source results in significant errors, and errors of this type have been discussed in previous lightning interferometer studies [10,11,24].

To test the errors that might result from the deconvolution method, a simulation with two point source emitters is developed. Each point source is modeled as long-duration Gaussian random noise source, to which a randomized filter is applied to soften the spectra. The first source is randomly positioned on the sky above 45° elevation; the second source is positioned at a specified distance but with random bearing from the first. In addition to the 2 point sources, 50 additional low-power point noise sources are modeled in the same way but distributed randomly on the sky. These noise sources are included because only a small fraction of the noise recorded at the real antennas is uncorrelated. Simulated antennas can then be configured in the same geometry as the LWA. The delay from each source (both signal and noise) to each antenna is then computed and applied to find the received signal.
at each simulated antenna. Next, uncorrelated Gaussian noise is added to the signal at each antenna. Finally, the effects of sampling are applied: the signal at each antenna is shifted in time by a random amount, and the signals are converted into 8-bit integers in the frequency domain. The power ratio of signal to correlated noise to uncorrelated noise in the simulation is 1:0.5:0.01; the standard deviation of the time jitter is 2% of a sample period (approximately 0.1 ns).

The distance between the simulated sources is varied, and the resulting signals are then run through the imaging and deconvolution pipeline. During imaging, a 200 sample (0.97 µs) integration time is used. The aggregate results from 50 trials for each distance are shown in Figure 3g,h using $\sigma = \lambda / \sqrt{2D}$ and $\sigma = \lambda / D$ for the Gaussian, respectively. The blue line in each figure shows the location error in the order 1 solution. This first solution is the same solution that we would obtain if a single point source model were used; as such, the error is independent of the specific Gaussian chosen for deconvolution. The orange line shows the error for the order 2 solution; here, the size of the Gaussian used for deconvolution does affect the error. In both cases, when the separation between the sources is small enough, no order 2 solution is found; this can be seen in the figure by the absence of the orange line and the mean number of solutions (grey dashed line) falling to 1. Similarly, when the sources are separated by a distance much larger than both the angular resolution of the interferometer and the standard deviation of the Gaussian, the location of both sources is determined with good accuracy. In between these two extremes, the specific Gaussian used does have an effect.

When $\sigma = \lambda / \sqrt{2D}$, the error for closely separated sources can be quite large, but, as can be seen by the mean number of sources found, the number of these is low. For a small range of separations, when the sources are of the same brightness, deconvolution will result in 3 solutions even though only 2 sources were simulated. The order 1 source will be found between the 2 simulated sources, with an order 2 and order 3 source on either side of it in a line. An example of this is shown in Figure 3b. As the source separation increases, the error in the order 2 source very quickly matches that of the order 1 source.

When $\sigma = \lambda / D$, the sources have to be separated by a larger distance before an order 2 source can be found. This also results in no situation in the simulation where more than 2 locations are produced by the deconvolution algorithm. However, while the error of the order 2 source still approaches that of the order 1 source, it does so more gradually.

From Figure 3, it may appear that using $\sigma = \lambda / \sqrt{2D}$ in the deconvolution will produce more accurate results than $\sigma = \lambda / D$. When the signals are imaged with 249 antennas, this can be true; however, the situation degrades when the number of antennas used is reduced. In these cases, the side lobes of the dirty image are higher, and the deconvolution algorithm is more likely to misidentify a side lobe as a true source when a smaller Gaussian is used. An example of this issue can be seen in Figure 2b,h, where the side lobes of the image resemble point sources. As a result, this study uses $\sigma = \lambda / D$ for the deconvolution Gaussian.

The deconvolution method described above is applied to the example flash used in this study, and the resulting map is shown in Figure 1. In this example, the deconvolution method is applied to images produced using 1000 sample (4882 ns) integration periods. Figure 1a shows the map in the LWA’s native image plane colored by time. This is a two-dimensional, angular depiction of the flash, where the zenith is at the origin and the horizon falls on a unit circle. The axes of the image plane are the direction cosines, which means that the distances measured on the image plane are unitless. In this projection, the angular resolution of the LWA is constant, but the distances between sources do not correspond to a constant angular distance; the azimuthal angle is compressed near the zenith, and the elevation angle is compressed near the horizon. Moreover, 10° increments in the azimuthal and elevation angles are indicated by the division lines. Figure 1b shows an angular elevation–time projection of the same map using the same coloring. The map depicts a fairly typical normal (positive) polarity intra-cloud flash. The positive and negative channels of the flash overlap in areas but are fairly easy to distinguish, with the negative leaders being fuzzier and generally at higher elevation angles [8,36].
Figure 3. Depiction of the effects of deconvolution determined through simulation. (a) is a dirty image of two simulated sources separated by 0.042 on the image plane (3.4° at 45° elevation). The black + marks the true location of the simulated sources. (b) is the residual image after deconvolving with a Gaussian whose standard deviation $\sigma = \lambda / \sqrt{2D}$. The red $\times$ show the locations determined by the deconvolution algorithm. (c) is the residual image after deconvolving the simulated dirty image with a Gaussian of standard deviation $\sigma = \lambda / D$. The red $\times$ show the locations determined by the deconvolution algorithm. (d–f) are the same as (a–c), except that the simulated sources are separated by 0.14 on the image plane (11.4° at 45° elevation). Images use the same projection format as Figure 1a; division lines in panels (a–f) are 10°. (g) shows simulated results of the location error using $\sigma = \lambda / \sqrt{2D}$ for the order 1 (blue) and order 2 (orange) solutions as the separation between the sources is varied. The grey dashed line shows the mean number of solutions found in each trial. (h) is the same as (g), but using $\sigma = \lambda / D$.

The lightning flash was also located by the ENTLN approximately 7.6 km to the south east. The blue dots in Figure 1a show the directions to the ENTLN-located pulses, and Figure 1d shows when they happened during the flash. The ENTLN detected, but did not locate, the initial breakdown pulses. Overall, the locations and distances of the located ENTLN pulses are very consistent with the LWA map and indicate that the lightning channels are approximately 6200 m above the ground (7600 m above sea level).
2.4. Calibration

The LWA antennas are calibrated reasonably well, but there can be small deviations that change day to day in the antennas. For the best results, it is reasonable to calibrate the cable delay for each antenna for each flash to be processed. This also helps to identify individual antennas that are not operating properly, so that they can be removed from the imaging routine.

Delay calibration is achieved by identifying one (or several) point sources in the flash to calibrate with. The initiation of the lightning flash is usually well suited for this purpose, since it is always a single source and frequently reasonably bright, but any point source in the flash can be used for calibration. The location of this single source is determined in the typical manner using the deconvolution method described in Section 2.3. A set of linear equations can then be set up in the following form:

\[ \tau_{ij}^\text{measured} - \tau_{ij}^\text{expected} = \sigma_i - \sigma_j \] (7)

where \(\sigma_i\) and \(\sigma_j\) are the delay correction for the \(i\)th and \(j\)th antennas, respectively. \(\tau_{ij}^\text{measured}\) is determined by the lag time of the closest peak in the cross-correlation, and \(\tau_{ij}^\text{expected}\) is calculated based on the location of the source (which is taken as truth):

\[ \tau_{ij}^\text{expected} = \frac{d_{ij}}{c} \alpha \sin(\theta_{ij}) - \frac{d_{ij}}{c} \beta \cos(\theta_{ij}) \] (8)

where \(\alpha\) and \(\beta\) are the direction cosines of the source, and \(d_{ij}\) and \(\theta_{ij}\) are the baseline length and orientation for the antenna pair, respectively.

For 256 antennas, each point source results in 32,640 linear equations with 256 unknowns (the \(\sigma_i\)s). However, because of the way in which the equations are set up, the unknowns are not fully constrained as we can add an arbitrary constant delay to all antenna corrections and not alter the system of equations. To account for this, a final equation must be satisfied:

\[ \sum_i \sigma_i = 0 \] (9)

The system of equations can be solved using normal least squares methods [37]. Although the system of equations for even a single source is heavily overdetermined, we still use several sources in the same region for the calibration. Since the initial locations of the calibration sources are assumed to be true, the calibration routine is somewhat iterative. With each iteration, the calibration sources will move, resulting in slightly different corrections on the next pass. This could be accounted for by solving for the location of the source as well as the antenna delay corrections (adding 2 more unknowns for each calibration source). In practice, the corrections converge quickly enough that the added complexity is not required.

Figure 4 shows the results of delay calibration based on 46 sources from the flash depicted in Figure 1. These sources are chosen such that each is in a different region of the sky, is relatively bright, and is the only source located at this time. The true location for these 46 point sources is not known; instead, the location determined when imaging with all 249 antennas is assumed as the truth. Then, the difference in the source location in the \(\alpha\) and \(\beta\) directions on the image plane is computed for all other antenna configurations. The red distribution shows the variation of the 46 point source locations before calibration, and the green after. The variation is measured on the image plane and so is unitless; a variation of 0.0025 corresponds to 0.20° at 45° elevation. In each distribution, the solid line shows the mean variation, the shaded region the standard deviation, and the dashed line the maximum deviation above and below.
Figure 4. Figure showing the effects of delay calibration on source location by examining how the locations of 46 sources vary as antennas are removed from the solution. Panel (a) (red) shows the variation before calibration, and panel (b) (green) after. In both, the solid line shows the mean, the shaded region the standard deviation, and the dashed line the maximum deviation above and below for the 46 solutions. Source variation is measured on the image plane; a variation of 0.0025 corresponds to 0.20° at 45° elevation.

Because the 249-antenna solution is assumed to be correct, the variation of both the calibrated and uncalibrated distributions converges to 0 when all 249 antennas are used to locate the source. This also means that the magnitude of the variation seen in Figure 4 cannot be interpreted as an estimate of the location error. However, the results do show that the calibration is at least self-consistent; the uncalibrated distribution shows a clear bias in the variation, whereas the calibrated results have variations centered around 0 as expected. In addition, the maximum variations (dashed line) for the calibrated source locations are smaller than those of the uncalibrated sources.

Of note, for both the calibrated and uncalibrated distributions, the variation in source location increases significantly when only 3 antennas are used to image the source. This indicates that the extremely sparse arrays of previous studies may be susceptible to significant systematic errors in source location.

One complication in antenna delay calibration for the LWA is that the expected delay correction for each antenna (a few nano-seconds) is approximately the same as the additional time delays caused by the source not being infinitely far from the antennas. This means that the peak accuracy of the calibration is limited to sources in the vicinity of the point sources used to perform the calibration. Alternatively, it is possible to account for the delays due to a near-field source if the distance (or height) of the source can be estimated, but doing so is beyond the scope of this study.

3. Results

In order to investigate the dependency of VHF lightning maps on the number of antennas used, the processing methodology above was applied to the example flash using increasing numbers of antennas. To allow the multiple processing runs to complete in a reasonable amount of time, the imaging window was incremented by 5000 samples (24.4 µs), with the start of each integration period being identical between runs. This resulted in 9665 windows being imaged and processed during the 235 ms duration of the lightning flash. Imaging for these surveys was done using a single beam steering at the approximate geometric center of the flash (coordinates 0.66, −0.38). Before imaging, the outrigger antenna and the six antennas identified as faulty during calibration were removed from the array, leaving 249 antennas to use for imaging. The antennas were ordered such
that imaging with the first \( N \) antennas would always result in a reasonable, but perhaps not ideal, antenna geometry. This was done by choosing the antennas sequentially such that the next antenna chosen always maximized the distance between itself and all antennas already selected. The resulting antenna geometries were somewhat heavily weighted towards long baseline lengths but were comparable to the sparse arrays used in previous lightning interferometer studies [11].

3.1. Multi-Source Detections

Figure 5a shows how the fraction of windows with identifiable VHF point sources changes as a function of the number of antennas used and the solution order. Solution order here refers to the order in which the solutions were identified during deconvolution. The order 1 solution is the brightest source in the integration window, and the higher-order solutions are increasingly fainter. For this survey, the integration period was held constant at 100 samples (488 ns), while the number of antennas was changed between runs.

The number of windows with at least one solution in them increases rapidly as the number of antennas increases. By the time 32 antennas are used to image the source, essentially all integration windows imaged have at least one source. The three-antenna case shown in Figure 5a is lower than expected based on the numerous previous studies conducted using three antenna arrays. The reason for this is that the LWA can only record 39.6 MHz of the VHF band, and this bandwidth is shifted towards the higher end of the 3–88 MHz sensing bandwidth to maximize the angular resolution. This results in the side lobes of the images being quite high (as seen in Figure 2) and dirty images that may not meet the 6\( \sigma \) deconvolution requirements specified in Section 2.3.

The number of higher-order solutions increases more gradually as more antennas are used to image the source. When 249 antennas are used to produce the image, 51% of the windows contain more than one VHF point source. A survey was also run to investigate the dependence of this fraction on the integration time. When the integration time was reduced from 100 samples (488 ns) to 50 samples (244 ns), the results were nearly identical. Integration periods shorter than this were not tested because a single beam steering would not fully cover the entire flash. The fraction of windows containing higher-order solutions does increase as the integration period grows longer, but it increases more slowly than expected, likely due to the limited ability of the deconvolution algorithm to detect multiple point solutions.

Some of the point source solutions in the maps are due to noise or side lobes, and a reasonable question is whether higher-order solutions are largely due to noise. Solutions due to noise and signals can be differentiated because the noise solutions are generally
scattered in space, whereas the real solutions are tightly clustered on channels. To compute the false alarm rate, differentiation was implemented by simply measuring the distance from each solution to all other solutions; if the 10th nearest other solution was farther than 0.02 on the image plane (1.6° at 45° elevation angle), the solution was identified as noise. For the 249-antenna case, less than 1% of the order 1 point locations were due to noise or side lobes. This increased to 3–5% for higher-order solutions, depending on the integration period.

3.2. Sensitivity

As shown in Equation (6), if we normalize by the integration period, the image amplitude is proportional to the average incident radio power at the antenna from the direction of the source. Similar to [21], the image amplitude can be expressed in terms of the galactic background power received by an antenna. The reference power was found by averaging the total power received in multiple integration windows from all antennas when no lightning was present.

The probability density for the received power of a number of different antenna configurations is shown in Figure 5b. Here, the probability density is normalized against the number of integration periods imaged, rather than the number of point sources located. This means that the cumulative sum of each distribution will be less than unity if, on average, less than one source is located for each integration window (when a small number of antennas is used) and greater than unity if, on average, more than one point source is located per window (when a large number of antennas is used). Unsurprisingly, the number of low-power sources located increases rapidly as antennas are added to the array.

With 32 antennas in the array, the distribution peaks around 12.5 dB\text{GB}; this peak remains constant as antennas beyond 32 are added to the array. This happens because essentially every integration window has at least one locatable VHF point source if 32 or more antennas are used to image the source, as shown in Figure 5a.

As more antennas are used to image the sky, the noise floor of the images becomes continually lower. Sources with received power below 0 dB\text{GB}, i.e., sources radiating less power than the galactic background, are routinely located. Indeed, there is a secondary peak in source powers less than 0 dB\text{GB} seen in the 249-antenna distribution. These low-power sources are not simply noise solutions from side lobes; approximately 85% of them are located on lightning channels. Sources this faint can be located because the galactic background is not a point source, and so the maximum image amplitude of the galactic background is significantly lower than the reference power.

3.3. Positive Leader Tips

Positive leaders are substantially less bright than their negative counterparts and also tend to extend along multiple channels simultaneously. Radio emission is frequently observed along the positive leader channel as either retrograde negative breakdown [4,8,38] or needle emission [14,39]. The positive leader tips should also radiate, but definitive observations of radio emissions are rare.

Figure 6 shows some of the positive leader activity seen in the example flash shown in Figure 1. Figure 6a,b show 0.25 ms snapshots of the positive leader activity in color. The grey scale shows the LWA map up to this point in the flash, where, at this time, in this field of view, two major channels can be seen. The channel to the left is the negative leader, which has propagated out of the field of view but is still active. The channel to the right is an actively developing positive leader, where the eventual path of the positive leader is shown with a thin line. In both cases, the polarity of each channel was identified visually based on its shape and behavior.

Needle activity can be seen towards the south end of the positive leader channel, extending off-frame towards the flash origin. This manifests as clusters of comparatively bright VHF sources along an extended portion of the channel, consistent with the observations published in Hare et al. [39] and Pu and Cummer [14]. In addition to the needle
activity, to the north along the channel line is a tight cluster of blue-colored sources, which are the positive leader tip.

![Depiction of positive leader activity](image)

**Figure 6.** Depiction of positive leader activity in the example IC flash. (a) 0.25 ms of activity starting 264.0 ms after 00:15:15 UT, in the same projection format as Figure 1a. In grey scale is the activity in the flash before this point, showing the channel structure. The thin grey line shows the path followed by the positive leader. Color shows the received power using the color bar at the bottom of panel (c). (b) Same as (a), but starting at 280.0 ms after the second. (c) Histogram of the source power for positive leader tips (blue) and needles (orange) observed during the ∼15 ms period between panels a and b. (d) Dirty image for a single 200 sample (976 ns) integration period starting 264.124 ms after the second (part of panel (a)). The locations identified by the deconvolution algorithm for this frame are marked +. The peak image power is 2.03 dB, with the color map being linear (not logarithmic) with respect to $E^2$. (e) Streak image showing the dirty image power along the positive leader channel path (shown in panels (a,b)), with color indicating received power using the scale in panel (c). The streak image was created using a 0.97 µs integration time and 4.88 µs between frames. When constructing the streak image, each column of the streak was obtained from 7 frames (34.2 µs), where each pixel was the maximum value of these 7 frames at this location along the positive leader path. The vertical axis has been converted to linear distance along the path by assuming a constant altitude of 6200 m above the ground (7600 m above sea level).
The coloration of these maps indicates the received power at the interferometer, with the color scale shown at the bottom of Figure 6c. The histogram in Figure 6c shows the distribution of the brightness for the VHF sources located in the needle emission region and in the tip. Unlike the needle emission, the positive leader tips are quite faint, with brightness values slightly below 0 dB.

At these faint brightness levels, the positive leader tip is frequently masked out by a brighter source and is rarely the only emission point in the sky. An example dirty image is shown in Figure 6d, including a source from the positive leader tip, where the + marks show the five VHF sources that were identified in the frame (including one of the positive leader tips). Although the deconvolution algorithm correctly identified much of the emission seen in the dirty image, there is still additional unlocated emission along other channels. The faintness of the positive leader tips’ emission, coupled with the rather crude deconvolution routine herein, causes the positive leader tips to come and go from the maps of the lightning flash.

By examining the dirty images directly, it becomes clear that the positive leader tip is a continuously radiating feature. Figure 6e shows a streak image of the dirty image brightness along the positive leader channel line depicted in Figure 6a,b. The color shows the indicated received power of the dirty image using the scale in Figure 6c. Portions of the dirty image that have negative intensity are not physical and are masked in white. Both the positive leader tip and the needle activity are identifiable, with the positive leader tip being a continuous feature in front of the needle activity along the channel. Because of the finite angular resolution of the LWA, both the positive leader tip and the needle emission cause brightness along an extended length of the positive leader path in the streak image, obscuring the separation between them. Careful examination indicates that the positive leader tip is 1–2 km in front of the needle emission.

The results also indicate that the positive leader propagates roughly 4.5 km in 48 ms at a nearly constant velocity of $9.4 \times 10^4$ m/s. This is significantly faster than the $2 \times 10^4$ m/s reported in previous radio studies of in-cloud positive leaders [4,8,39–41], but similar to studies of positive cloud-to-ground leaders [42,43]. The conversion to linear distance was achieved by assuming that the positive leader was propagating purely horizontally at a constant altitude, similar to the methods used in Stock et al. [33]. The altitude used was 6200 m above the ground (or 7600 m above sea level), based on the ENTLN pulse locations for this flash. The nearest sounding (85 km to the north) indicates a temperature of $-15^\circ\text{C}$ at this altitude, a reasonable height for the main negative charge layer for the storm. We estimate an error of around 1 km in the altitude of the positive leader, and the velocity estimate does not change greatly within this range. The unusually high velocity of this positive leader may be related to the detectability of the emission from the positive leader tip.

4. Summary and Discussion

In this study, we have presented a foundation for the use of measurements from the LWA to detect and locate lightning VHF sources. The resulting maps of lightning are possibly the most sensitive radio measurements of lightning to date. Sources can and are located with received power well below the galactic VHF background. It would be impossible to identify these VHF sources by examining the VHF signals received at only one antenna.

For the LWA, the ability to identify pulses in the time series is further hampered because the noise level of the time series is actually significantly higher when lightning is present than when it is not. This happens because, when the complex visibilities are compressed to 8 bits (split equally between real and imaginary components), the observation can and frequently does clip at ±7 counts. This clipping in the frequency domain causes additional noise for the entire 40 µs observation window for this antenna. Remarkably, simulations indicate that this clipping does not adversely affect source location retrieval.
The locations of VHF sources were achieved using a projection imaging routine that is fully equivalent to aperture synthesis imaging. The method is not the most computationally efficient way to image a source and is particularly inefficient when the number of baselines is large, as in this study. The current implementation takes several seconds to image a single integration period with all 256 antennas; imaging all integration periods in a flash requires a significant amount of time (days to months on a single computer). The primary motivation for using projection imaging versus any other imaging method was simply because we were familiar with it and so it could be implemented fairly quickly.

Nevertheless, there are benefits to projection imaging. The largest of these is that Equation (5) can be trivially extended to other coordinate systems. This allows imaging to be achieved in any coordinate system for which $\tau$ can be uniquely calculated for each baseline. For example, the image can be easily computed in the azimuth–elevation projection, as seen in Romero-Wolf et al. [44] and Tilles et al. [11]. If the baselines are long enough, the image can also be projected onto three spatial dimensions. This was achieved using VHF signals in Hare et al. [39] and with electric field change records in Zhu et al. [45]. While not used in the present study, it should be similarly possible to make near-field corrections for sources observed with the LWA.

For the flash examined in this study, the LWA was able to locate at least one VHF source in virtually every integration period, independent of the integration period used and with very low false alarm rates. In comparison, other lightning interferometer studies that mapped a significant portion of a lightning flash were able to locate $\sim 2$–14% of their observing windows [8, 39, 46].

Until now, lightning interferometer studies have universally assumed that if the integration period is small enough, lightning can be modeled as a single point radiator. A major advantage of using so many antennas to image lightning is that multiple VHF lightning sources can be identified in each integration window. In the flash in this study, around 51% of the integration windows examined had more than one VHF source, and this fraction was relatively independent of integration period. This estimate is almost certainly a lower bound as the deconvolution method employed in this study, which enabled multiple sources to be identified, was effective but crude. Much more sophisticated deconvolution algorithms exist and should allow sources to be identified even farther into the background noise. The fact that lightning emits VHF radiation from multiple locations at once should be no surprise; high-speed video observations of lightning routinely show multiple simultaneously active branches [47–51].

As interferometer observations increase in sensitivity and resolution, the single point radiator model becomes less and less accurate; this is shown in Figure 7. Here, statistics were computed for the 16 ms period of time between Figure 6a,b, although the results are very similar at other times in the flash. The solid line shows the probability that a source is the brightest in its integration window; for sources below 0 dB$_{GB}$, this happens less than 25% of the time. For very sparse arrays, the situation is even worse since, if there is another source close in power, the interference will make location unlikely, as illustrated in Figure 2g. The probability of there not being a second source close in power is indicated by the dotted line in Figure 7, specifically the probability that a source is at least 3 dB brighter than any other sources in the same integration window. In this study, had we not used a multiple-source model to deconvolve the dirty images, significantly less detail would be seen in the lower-power portions of the flash, such as the positive leader tip. Ultimately, using a single source model for the VHF emission from a lightning flash is only valid for the brighter VHF sources.

The improved sensitivity and multi-source capability of the LWA open up a new class of lightning observations. As an example of these abilities, we have provided observations of positive leaders in an IC flash. The streamer zone of this positive leader tip radiated at approximately the same brightness as the galactic background and propagated at an unusually quick $9.4 \times 10^4$ m/s. The fact that emissions such as this have not been seen in the past is unsurprising, as the received signals would be indistinguishable from noise.
when measured using a single antenna. The location of these positive leader tip sources is further hampered because sources this faint rarely appear isolated in time.

![Figure 7](image-url)

**Figure 7.** The relationship between the received power and source order. Solid line shows the probability that a source of a given power is the brightest source in the integration window (1 μs). Dotted line shows the probability that a source of a given power is at least 3 dB brighter than any other sources in the same integration window.

VHF emission on positive leader channels has been routinely observed, but, in almost all cases, the emission was due to needles [36,52,53], which have been interpreted as being caused by negative breakdown [14,39]. Very bright VHF emission from positive streamers, termed fast positive breakdown, has been seen at the beginning of some flashes [10,13,54,55]. Similar fast positive breakdown has also been been reported on lightning channels in the middle of a flash immediately following ground attachment or K/dart leaders, caused by a large potential difference between the leader tip and the cloud [33].

VHF emission from the tips of propagating positive leaders is much more rarely observed. Pu et al. [56] performed VHF interferometer observations of a very nearby positive leader to ground. The VHF emission from this leader tip was impulsive and comparably bright to the negative leader in the same flash—much brighter than the positive leader tip in this study, but substantially less bright than fast positive breakdown. The positive leader also propagated faster than the leader observed in this study, at $2 \times 10^8$ m/s as it neared the ground. The authors estimated that the electric field was the positive streamer region $0.9E_L$, substantially higher than that needed for steady positive leader propagation.

The leader propagation speed is dependent on the potential at the tip of the leader [57]. Even so, it is usually assumed that the field in the streamer zone is uniform and close to the stability field for positive streamer propagation [58]. The observations indicate that when the tip potential is higher, this may not be true, leading to VHF emission being produced. From this, we surmise that the tip potential of the positive leader in this study was higher than average, possibly because the observations were made during the active phase of the flash [59], when the negative leader maintained higher conductivity in the channel [4,60], or because the leader was propagating in a higher than average electric field. It is very likely that positive leaders propagating at a more typical $2 \times 10^4$ m/s will produce much less VHF emission and could even be virtually undetectable.

The LWA is capable of even more than what has been shown in this study. Because the antennas are well calibrated, it is possible to examine the spectral content of individual lightning VHF sources. Indeed, determining the spectra of lightning is a prerequisite for many of the more sophisticated deconvolution methods. In addition, the images and maps shown here used only one of the two polarizations recorded by the telescope. The VHF emission of lightning can be polarized [23,61], and the polarization information contains
important details about the physics of the breakdown process. Finally, the correlator on LWA1 and LWA-SV will soon be upgraded, allowing for more bandwidth to be captured. The number of LWA sites in general is also increasing, thereby increasing the number of lightning flashes that may be observed using LWAs.

**Author Contributions:** Conceptualization, M.S., J.T., J.D., G.B.T. and N.L.; methodology, M.S. and J.T.; software, M.S.; formal analysis and investigation, M.S. and J.T.; resources, J.D. and G.B.T.; data curation, J.D. and J.T.; writing—original draft preparation, M.S.; writing—review and editing, M.S., J.T., J.D., G.B.T. and N.L.; funding acquisition, J.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** Funding was provided by the NOAA/Office of Oceanic and Atmospheric Research under the NOAA–University of Oklahoma Cooperative Agreement NA21OAR4320204, U.S. Department of Commerce, and AFOSR Award FA9550-18-1-0358 to the University of New Hampshire. Supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525. Construction of the LWA has been supported by the Office of Naval Research under contract N00014-07-C-0147 and by the AFOSR. Support for the operations and continuing development of LWA1 is provided by the Air Force Research Laboratory and the National Science Foundation under grants AST-1835400 and AGS-1708855.

**Data Availability Statement:** The algorithms used to image lightning using LWA data are available at https://github.com/mikestock/lwali, accessed on 1 July 2023. The LWA observations for the example flash shown in this study are available at https://doi.org/10.5281/zenodo.8021798, accessed on 12 June 2023.

**Conflicts of Interest:** The authors declare no conflicts of interest.

**References**


20. Hare, B.M.; Scholten, O.; Buitink, S.; Dwyer, J.; Liu, N.; Sterpka, C.; ter Veen, S. Characteristics of recoil leaders as observed by LOFAR. Phys. Rev. D 2023, 107, 023025. [CrossRef]


47. Mazur, V.; Krehbiel, P.R.; Shao, X.M. Correlated high-speed video and radio interferometric observations of a cloud-to-ground lightning flash. *J. Geophys. Res. Atmos.* 1995, 100, 25731–25753. [CrossRef]


58. Dwyer, J.R. Terrestrial gamma-ray flashes initiated by positive leaders. *Phys. Rev. D* 2021, 104, 043012. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.