EPIC Imaging & Source Localization Comparisons

Craig Taylor, Jayce Dowell

November 27, 2023

1 Introduction

The purpose of this project has been to aid in commissioning and validating the data products of the E-field Parallel Imaging Correlator (EPIC) subsystem at the Long Wavelength Array Sevilleta (LWA-SV) Station. The goals of this work are to perform a direct image comparison using the standard LWA-SV all-sky FX-correlator mode, the "Orville Wideband Imager", and its direct-imaging analog in EPIC. This project developed two functional comparisons, one using live or "online" images from the array and the other using offline analysis of scheduled observations. The online comparisons would utilize standard captures with the two instruments, whereas offline captures use the Transient Buffer Frequency-domain (TBF) and Transient Buffer Narrowband (TBN) modes to record or simulate raw data to be correlated offline at a later time.

The basic outline of this effort is the following. Data was to be collected using the online captures on EPIC, commensal with Orville on the LWA-SV station, and the full-spectrum Orville data saved to the archive shortly after the observation. Based on previous EPIC publications for this correlation method, these two data feeds should produce identical data so long as we match EPIC's image size, shape, and bandwidth with that of Orville. Thus once data was collected using both machines, the EPIC dataset would be matched in the aforementioned aspects to create an analog to the Orville data. Next the normalized images are compared in total image subtraction to understand morphological differences (if any), and individual source localizations to ensure sources were in agreement. The output of these two tests should show effectively zero power in image subtraction and sources should have the same locations. Bright sky sources will be localized using a source finding algorithm on a window of pixels sufficiently encapsulating each source.

Originally we experimented with using almost an entire LWA sub-band (3.3 MHz) of frequency coverage , at the cost of having to downsize Orville images to 64×64 pixels to match what EPIC could produce at that bandwidth. However, we found this binning step too degrading on the images and instead opted to using a smaller bandwidth to allow EPIC to produce the same image sizes as Orville. For the online testing of this project, these considerations would provide us with image sizes of 128×128 pixels ($\sim 1.0156\ deg/pix$), total bandwidth of 300kHz, and an integration time on EPIC of $t = 1sec$. Since Orville images at a 5-second cadence, and EPIC at the time of this test does not place options for observation start time, this would give us a maximum timing error of $t/2$ or $0.5s$. After imaging, the data sets are aligned in time by minimizing the differences between timestamps,

then averaging the EPIC data between each minimum timing difference window. This results in some EPIC images having slightly more or less time averaged into them, but creates the best timing match for the currently online EPIC implementation. This method was used for investigating the autocorrelations and the preliminary stages of fitting for sources before we moved to almost exclusively offline data capture in §3.2.

In the following sections we will discuss the various tests we have performed to prepare EPIC for scientific research. §2 covers validating the autocorrelation calculation and excision in EPIC. In §3 we describe efforts to localize sources between the current implementation of EPIC and the Orville Wideband Imager or an equivalent. Lastly §4 summarizes the results of this study, outlines changes made to the current implementation of EPIC, and suggested future changes to EPIC operations.

2 Auto-correlation Removal

The most direct problem to solve during this process was to verify that auto-correlations were not removed properly by EPIC. It appeared that the auto-correlations were being propagated to the correct image size, but the scale was several orders of magnitude below what should be expected. Additionally, we noticed significant gridding errors that were effectively corrupting the images produced through the EPIC pipeline when auto-correlation removal was requested. These gridding errors were randomly distributed in time and frequency (see Figure 1), leading us to believe this was a software fault. Toggling autocorrelations on and off before imaging the data showed that without autocorrelations removed the images would not display gridding errors. Functionally isolating the gridding problem to the same code that calculates the autocorrelation magnitude. Since EPIC's autocorrelations are applied as a planar offset, with a series of test observations sampling the input datastream to the autocorrelation block, we were able to isolate that the issue was contained to a function used in gridding the auto-correlations called 'VGrid.cu'. Non-atomic operations in calculation of auto-correlations resulted in random memory conflicts within 'VGrid.cu' attempting to write to the same parcel of memory. Replacing a normal addition with an atomic addition, enforcing only a single write in memory at a given time, we were able to fix both the auto-correlation scaling as well as removed the corrupted (see Figure 2).

3 Phase Center Errors

The second major issue discovered in test observations was that Orville and EPIC were not producing matching images of the sky, despite polling the same data stream from the array to form images. Not only were point sources not in the same spot with the two imaging tools, but the galactic plane as well as the horizon were mismatched. The following sections will discuss two component of these errors individually, horizon effects and source positions.

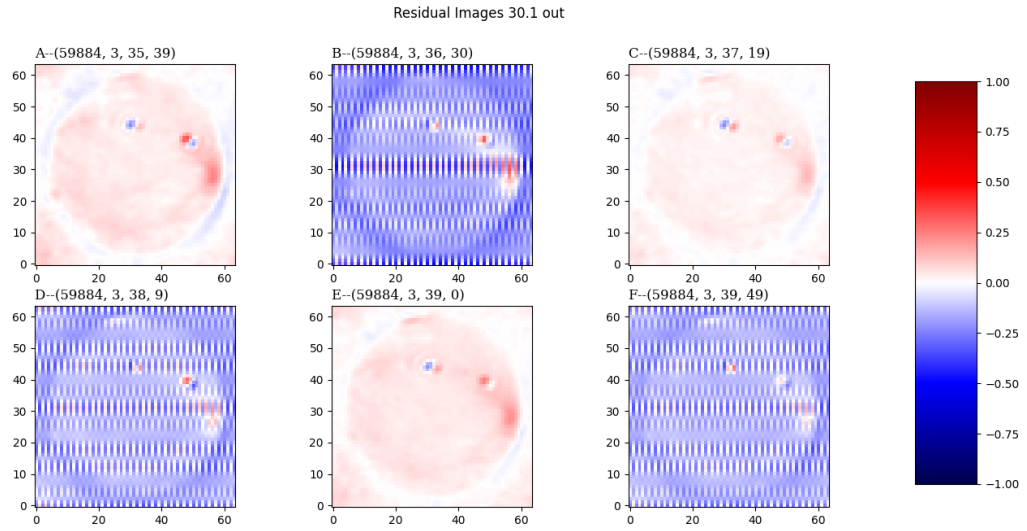

Figure 1: Residual images of 64×64 pixel imaging comparisons between EPIC and Orville. Images are timestamped in (MJD date, Hour, Minute, Second) format for a single track of images centered at 30.1 MHz (100kHz BW). The color bar represents normalized differential intensity between the two images with Bluer pixels indicating EPIC's image reporting more intensity, and Redder pixels for Orville respectively. Axial labels for individual images represent pixel coordinates. Note temporal irregularities of gridding errors presented plots B, D and F caused by memory errors in the EPIC gridder.

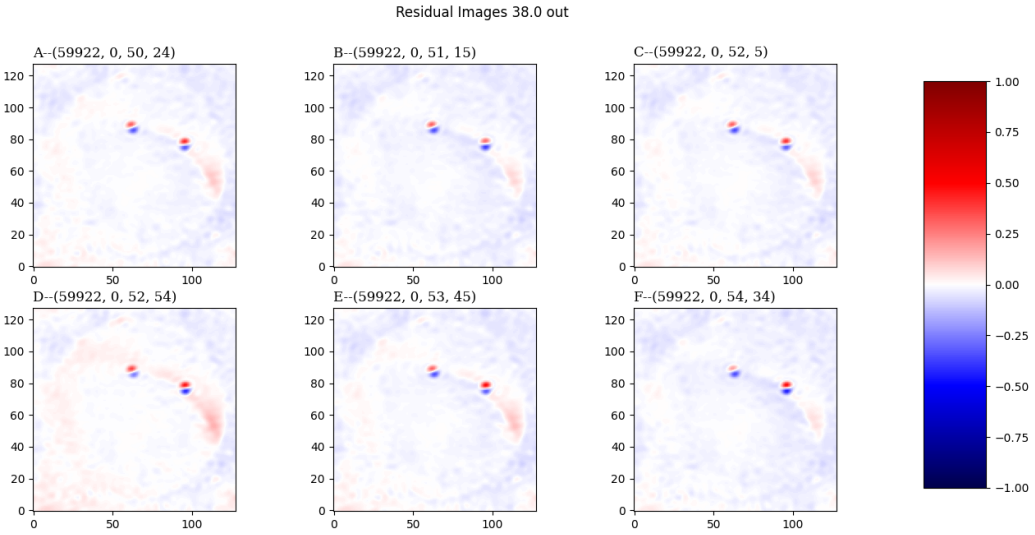

Figure 2: Residual images with the same description as Fig 1. Note that although this six pane figure was centered at 38.0 MHz rather than 30.1 MHz as above, it merely serves to illustrate that after gridding errors were rectified we are still left with pointing errors in EPIC.

3.1 Horizon Scaling

Comparing the horizon between Orville and EPIC showed a few irregularities such as improper scaling, different tapering applied at the horizon edge, and a mysterious shearing effect as though the sky in EPIC was slightly rotated compared to other imagers. We decided to focus on the scaling and shearing present in EPIC as tapering was currently under study by collaborators at Arizona State University as a part of a software redesign of EPIC. Among those considered here, the first to be fixed was incorrect beam scaling, which in turn also made the reported angular size of pixels and World Coordinate System (WCS) projections inconsistent with Orville. These corrections came in understanding the angular size of pixels and how that is related to the other functioning modes on the LWA system. Both the default setting for 'imageres' in the documented 'LWA_EPIC.py' script, and the version of the EPIC script circulating in use both had incorrect pixel resolution. The angular pixel scale calculated for the other LWA modes uses a base sky size of 130° divided by the frequency scaled number of pixels across the image to ensure good sampling (Ex. $130/128 \approx 1.0156$). This work led to several updates to the `lsl` standalone offline correlator and imaging tools. Simply running the EPIC code with an accurate pixel size to correspond with Orville effectively removed the size mismatch present at the horizon.

Shearing along the horizon was another irregularity we attempted to solve in this study. In all test observations for this project we noticed that residual images of subtracting `lsl` and EPIC showed point sources of radio frequency interference (RFI) associated with the horizon appeared to be rotated a few degrees counter-clockwise in EPIC images. Since discrete emitters seen in our images (point sources, the galactic plane, and various sources of RFI) also displayed some degree of offset and rotation, we decided to transition to localization of point sources to better understand and characterize the effect.

3.2 Discrete Sources and Horizon Shearing

To enable EPIC to make science ready observations in co-operation with the Orville subsystem, or any other LWA stations, we next needed to ensure point sources were seen in their correct positions. We took several transit observations using live data capture on Orville/EPIC, then moved to using raw voltage time-series captures with LWA's TBF capture mode, before finally settling on simulated model data. Using offline data allowed better control over our inputs down both pipelines, and also eliminated some degrees of freedom such as the known timing offsets when using live captures. These offline captures also allowed for better frequency resolution (25kHz channels over entire 19.8 MHz band) because we would no longer be limited to the memory constraints of either LWA subsystem in observing. To help with source localization we took a series of observations during the transit of Cygnus A and Virgo A, supplying instantaneous observation of sources as well as capturing the behavior of EPIC's imaging offsets over the shape of our sky.

In this offline configuration, the Orville subsystem is replaced by a correlator and imager analog already developed in the LWA Software Library [1] called *CorrelateTBF* and corresponding imager *ImageIDI*. Source localization was done using an island-finding algorithm similar to PyBDSF [3] or Aegean [2], performed over sufficiently large pixel windows that encompass target radio sources. This process provides us with a connected component object

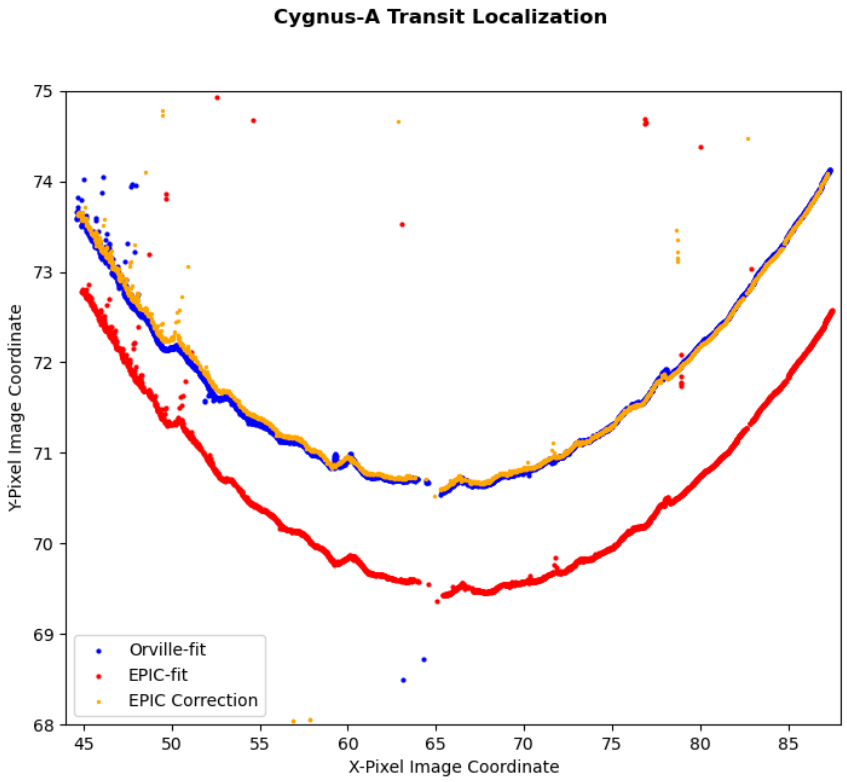

Figure 3: Example source fitting in pixel coordinates for Cygnus A during a full transit observation of approximately 12-hours. Blue and Red points are the positional coordinate fits for Cygnus-A using Orville and EPIC as the imager respectively. In orange are the corrected EPIC positions using least-squares optimization to solve for a rotational and linear offset in coordinates. After corrections for this observation the EPIC and Orville positions are in agreement.

for each source whose weighted centroid was used as the source location for better sub-pixel accuracy. This technique proves to be more effective at finding the most likely location of the source when compared to fitting a 2D-Gaussian to each point source. After fitting positions of sources in both EPIC and our `lsl` imager in a given set of transit observations, we solved for the offset between each transit curve.

Fitting this offset was done using least squares minimization tools in `scipy` and can be done using a simple Affine rotation matrix (in both cases) to find the rotation angle and linear pixel offset. An extra degree of freedom was included in the full image comparison to allow for the rotation center to be located anywhere within the image to help diagnose the issue. Results for this test are shown in Figure 3, where we can see a bulk offset of 1-2 pixels and a rotation component of roughly 0.5-2 degrees depending on the rotation center allowed in fitting. While this fitting scheme only used three degrees of freedom and does a poor job of properly transforming the data, it was simply used to get a gauge for the scale of incorrect positions. It should be noted that careful work was done to verify that the WCS projections for each set of images was identical and reflective of the phase center correction

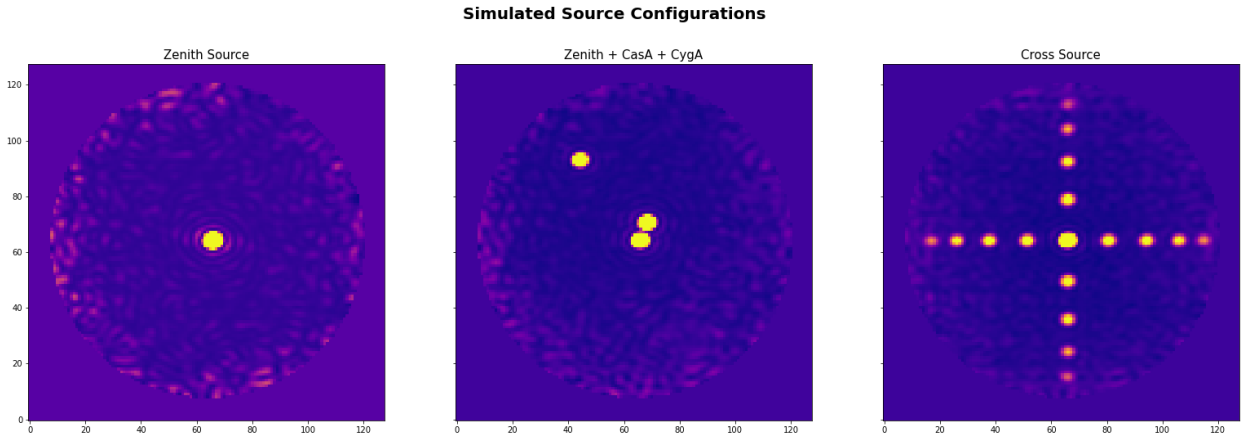

Figure 4: Example simulated point source orientations used in EPIC validation.

applied to each correlator. This way we could be certain that the pixel-to-celestial coordinate referencing was uniform.

Several attempts were made to identify these errors in EPIC such as verifying both pipelines reference the same phase center, better handling of gridding antenna positions and verifying their locations in the grid, cross-checking the phases for each antenna, confirming these effects scale with image size, and checking intermediate products out of the gridder and cross-multiplication modules. These tests proved fruitless in identifying a cause to these problems so we moved to modelling point sources and the array next.

We approached modelling this problem by creating an idealized LWA station with its phase center at the geometric zenith. This perfectly coplanar "LWA-SV" station was made by modifying the antenna positions and feeding in prescribed locations of simulated Gaussian sources in custom TBN files. These sources were placed in three configurations for the test: a single point source at zenith (Figure 4), a zenith source plus model sources for Cygnus A and Cassiopeia A, and a cross of identical point sources spaced at 15 degree altitude cuts. Simulated data was made in TBN format using adapted `lsl` tools, then each configuration was fed into both pipelines to be processed. Using all three of these configurations we were able to identify a systematic one pixel shift was the cause of a majority of the bulk offset seen in Figure 3. This was introduced by an inversion of the images immediately before saving to disk in order to orient the sky such that the East and West directions were appropriately located. We attempted to mitigate this by artificially inverting the horizontal coordinate of each antenna to remove the need to flip the images, but this proved less effective than simply rolling the image a single pixel to remove the offset.

These simulations did help to visualize the horizon shearing effects being introduced within the EPIC pipeline described in the previous sections. Seen in Figure 5, we can see the cross of point sources described above as imaged in both Orville and EPIC. Here we can better visualize the magnitude of the shearing effect originally noticed around the horizon, however these results show this is not contained to the horizon but rather an effect on the entire image. We thought the homogeneity of this error could be caused by inconsistencies in phase between the two imagers, but we were unable to locate any miscalculations of phase

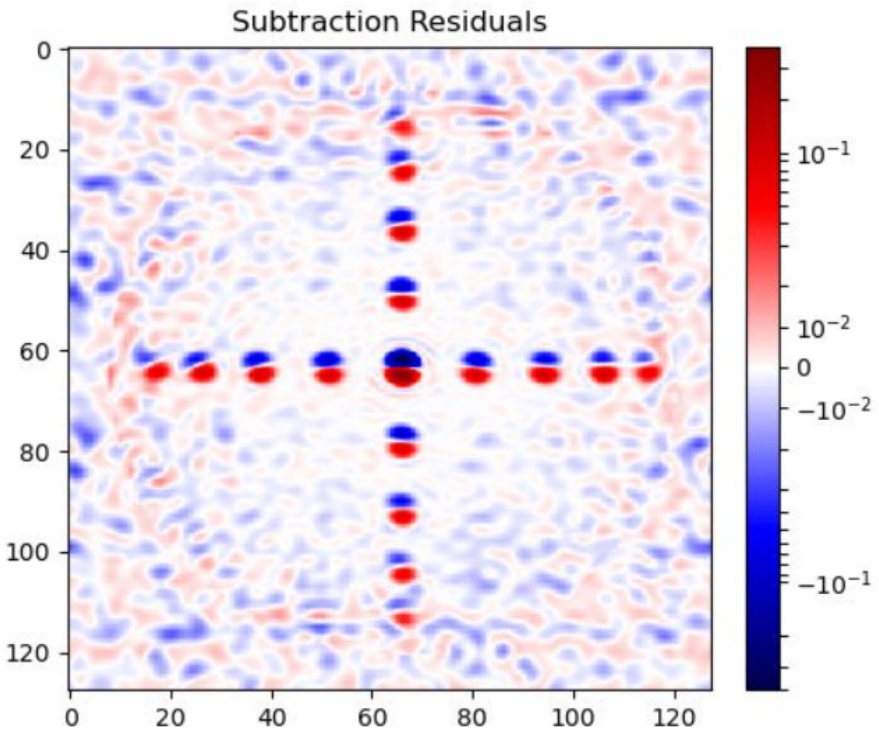

Figure 5: Subtraction using a cross of point sources as imaged by both EPIC and the offline `lsl` configuration. EPIC sources show the red end of the color bar while `lsl` sources represent the blue portions. These images are after rectifying the single pixel shift described in §2.2. Two effects can be observed in the figure, the rotational shearing of the entire image and a slight mismatch in flux as a function of altitude.

in EPIC. We also tried swapping the In-phase and Quadrature data packing in attempts to see if there was an accidental inversion somewhere in the EPIC pipeline, but testing showed this to be unlikely as images still contained the same artifacts.

To quantify the rotation seen in Figure 5, another series of TBF captures was taken to encompass a single transit of Cygnus A, then fit using the same algorithm as before. These TBF observations took place on September 17th 2023 from 00:00 UTC - 08:00 UTC, with a 15 minute observing cadence. Since it appears there is an altitude dependence on the rotational effects present in the EPIC images, we fit the data using a sliding box least squares regression. Using this method we find the average rotation across the sky to be approximately -1.278 degrees (negative indicating counter-clockwise or West-to-East) with a maximum rotation estimate near the horizon of -3.037 degrees. In future tests to diagnose this error it may be more useful to work in a simulation setting to acquire better estimates of this rotational offset using a better sampled grid of point sources than used in this work.

A clue to what may be causing these imaging inconsistencies was found when images are Fourier transformed back into the aperture plane. In the aperture plane, we find the amplitudes between the `lsl` and EPIC pipelines to be relatively consistent. However, the aperture phases in the EPIC pipeline appear to be rotated inside the main beam by approximately 25.2 degrees. Outside the horizon this effect is a bit more confounded but generally appears

to be similar in behaviour to inside the beam. Figures 6 & 7 both illustrate this effect with a simulated point source and cross of point sources centered on the LWA-SV zenith position, and on-going work is directed towards understanding where this effect is introduced within EPIC.

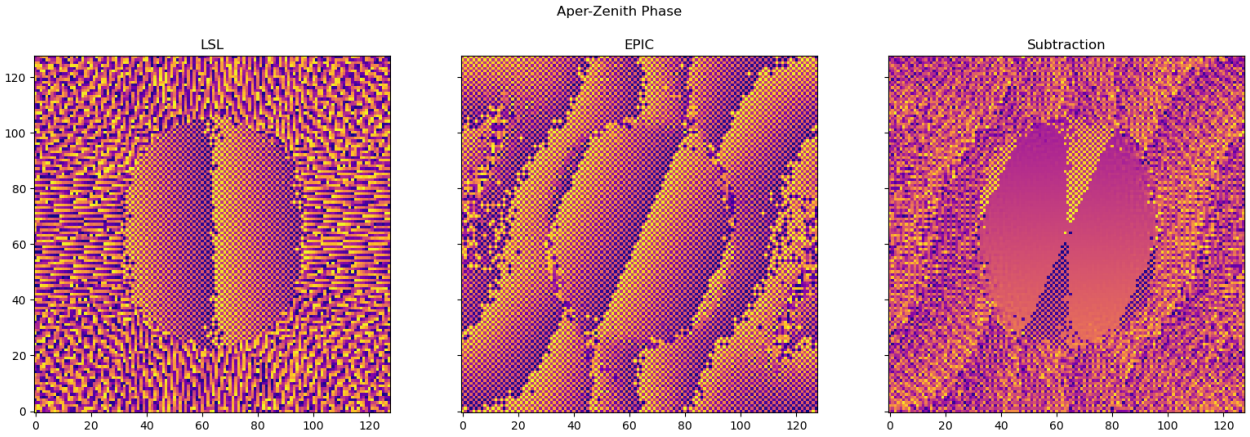

Figure 6: Aperture phase plot for a simulated point source located at the zenith as imaged by our `lsl` and EPIC pipelines respectively. The difference between these two normalized images is shown in the third panel. Note the ~ 25 degree rotation the to the EPIC images and the corresponding ripples outside the main lobe.

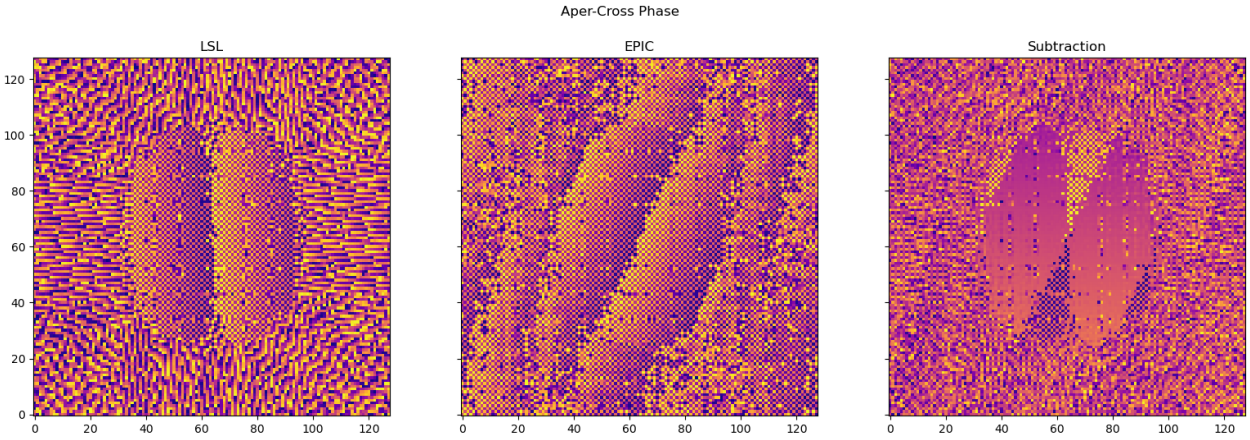

Figure 7: Same trio of plots as Fig 6, but seeded with a cross of simulated point sources arranged at the zenith. Interestingly the effects of adding more point sources does not change the structure of the underlying rotation.

4 Discussions and Suggested Improvements

This image analysis and testing has been essential in locating and rectifying several outstanding problems within the LWA-EPIC system, but still raises some questions about the reliability of this implementation. The autocorrelation features have been altered to use atomic operations to prevent data corruption and correctly offset images. We were also able to locate and correct several cosmetic changes to improve data reporting, such as an improved image headers and more accurate phasing to account for the offset zenith at LWA-SV. However, there are still critical issues present in the software causing a shearing effect on images that increases in severity as a function of angle from zenith. While we were able to remove the majority of observed linear offsets in images, this shearing can cause pointing errors near the horizon of several degrees. Thus this EPIC deployment cannot fully be considered science ready until the issues presented here can be rectified.

Changes made during this project will be reflected in the retired python implementation of EPIC and hosted on the LWA-EPIC github repository. The testing data and instructions to replicate the image products used are encapsulated in the EPIC file storage under the directory "/FileStore/epic0/memo12_Data". All of which can be done using basic functions within the LWA Software Library. The results of this testing also yielded the following suggestions to be implemented in future iterations of the LWA-EPIC system. EPIC currently is not equipped to tolerate imaging with odd image sizes which can be problematic when testing simulated data for example. There also does not exist a well documented user guide to operating the EPIC software, further raising the bar to entry on a theoretically very flexible system. This makes troubleshooting observations difficult without knowing what input values are important for achieving the desired image properties. Inputs such as image size, frequency resolution, or gulp size can significantly impact the reliability of the current EPIC implementation to produce science ready images. Moreover there is not currently a way to schedule observations via EPIC, regardless of the LWA-EPIC project scope of having a continuously monitoring instrument, it is still a valuable asset during these testing phases or for future targets of opportunity.

References

[1] Jayce Dowell et al. "The Long Wavelength Array Software Library". In: *Journal of Astronomical Instrumentation* 1.1, 1250006 (Dec. 2012), p. 1250006. DOI: 10.1142/S2251171712500067. arXiv: 1209.1576 [astro-ph.IM].

[2] P. J. Hancock et al. *Aegean: Compact source finding in radio images*. Astrophysics Source Code Library, record ascl:1212.009. Dec. 2012. ascl: 1212.009.

[3] Niruj Mohan and David Rafferty. *PyBDSF: Python Blob Detection and Source Finder*. Astrophysics Source Code Library, record ascl:1502.007. Feb. 2015. ascl: 1502.007.