MNRAS **000**, 1–10 (2015) Preprint 26 October 2015 Compiled using MNRAS LaTeX style file v3.0

A Generic and Efficient E-field Parallel Imaging Correlator for Next-Generation Radio Telescopes

Nithyanandan Thyagarajan,[1★] Adam P. Beardsley,[1] Judd D. Bowman[1] and Miguel F. Morales[2]
[1]*Arizona State University, School of Earth and Space Exploration, Tempe, AZ 85287, USA*
[2]*University of Washington, Department of Physics, Seattle, WA 98195, USA*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
Modern radio telescopes are favouring densely packed array layouts consisting of a large number of antennas ($N_{\rm a} \gtrsim 1000$–$\gtrsim 10000$). Since the cost of traditional correlators scales as $O(N_{\rm a}^2)$, there will be a steep cost for realizing the full imaging potential of these powerful instruments. Through our generic and efficient E-field Parallel Imaging Correlator (EPIC), we present the first software demonstration of a generalized direct imaging algorithm known as the Modular Optimal Frequency Fourier (MOFF) imager. It takes advantage of the multiplication-convolution theorem of Fourier transforms. Not only does it bring down the cost to $O(N_{\rm a} \log N_{\rm a})$ but can also image from irregularly arranged heterogeneous antenna arrays. EPIC is highly modular and parallelizable. It is implemented in object oriented Python and is publicly available. We have verified the images produced to be mathematically identical to those produced using traditional techniques. We have also validated our implementation on data observed with the Long Wavelength Array. Antenna systems with a dense filling factor consisting of a large number of antennas such as LWA, SKA, and HERA will gain a significant advantage by deploying EPIC. Inherent availability of calibrated time-domain images on timescales roughly equal to the writeout timescale of the digitizer will make it a prime candidate for transient searches of primarily Fast Radio Bursts (FRB) as well as slower planteary and exoplantenary phenomena.

Key words: instrumentation: interferometers – techniques: image processing – techniques: interferometric

1 INTRODUCTION

Radio astronomy is entering an era in which interferometers of hundreds to thousands of individual antennas are needed to achieve desired survey speeds. Nowhere is this more apparent than at radio frequencies below 1.4 GHz. The study of the history of hydrogen gas throughout the universe's evolution is pushing technology development towards arrays of low-cost antennas with large fields of view and densely packed apertures. Similarly, the search for transient objects and regular monitoring of the time-dependent sky is driving instruments in the same direction. A number of new telescopes are under development around the world based on this new paradigm, including the Murchison Widefield Array (MWA; Tingay et al. 2013; Bowman et al. 2013), the Precision Array for Probing the Epoch of Reionization (PAPER; Parsons et al. 2010), the Hydrogen Epoch of Reionization Array[1] (HERA), the LOw Frequency ARray (LOFAR; van Haarlem et al. 2013), the Canadian Hydrogen Intensity Mapping Experiment (CHIME, Bandura et al. 2014), the Long Wavelength Array (LWA, Ellingson et al. 2013), and the low frequency component of the Square Kilometer Array Low Frequency Aperture Array (SKA-Low Mellema et al. 2013).

This paradigm shift requires a fundamentally new approach to the design of digital correlators (Lonsdale et al. 2000). Modern correlators calculate the cross-power correlation between all antenna pairs in many narrow frequencies, forming *visibilities*, the traditional fundamental measurement of radio interferometers. The computational requirements for a modern FX correlator scale with the number of antenna pairs, or the square of the number of antennas $\sim N_{\rm ant}^2$ (Bunton 2004). For this reason traditional correlators have difficulty scaling to thousands of antennas. As an example, the full HERA correlator for 352 dishes with 200 MHz of bandwidth requires 212 trillion complex multiplies and adds per second (TMACS). Future arrays with thousands of collecting elements will require orders of magnitude more computation, making the correlator the dominant cost.

For certain classes of radio arrays there is an alternative to

★ E-mail: t_nithyanandan@asu.edu
[1] http://reionization.org

the FX correlator that can lower the computational burden by directly performing a spatial fast Fourier transform (FFT) on the electric fields measured by each antenna in the array at each time step, removing the cross-correlation step. This relieves the computational scaling from the harsh $N_{\rm ant}^2$ to the more gentle envelope of $N_{\rm pix}\log N_{\rm pix}$, where $N_{\rm pix}$ is the number of pixels in the Fourier transform (e.g. Morales 2011; Tegmark & Zaldarriaga 2009; Tegmark & Zaldarriaga 2010). This architecture is often referred to as a "direct imaging" correlator because it eliminates the intermediary cross- correlation data products of the FX and earlier lag correlators, but instead directly forms images from the electric field measurements.

Direct imaging correlators have begun to be explored on deployed arrays including the Basic Element for SKA Training II (BEST-2) array (Foster et al. 2014), the Omniscope (Zheng et al. 2014), and an earlier incarnation at higher frequencies with the intent of pulsar timing (Otobe et al. 1994; Daishido et al. 2000). However, each of these examples make assumptions about the redundancy of the array layout, and require the collecting elements are identical. On the other hand, the MOFF algorithm achieves the same $N_{\rm pix}\log N_{\rm pix}$ computational scaling without placing any restriction on antenna placement, can accommodate non-identical beam patterns, and is a provably optimal mapping (Morales 2011). This algorithm uses the antenna beam patterns to grid the electric field measurements to a regular grid in the software holography/A-transpose fashion (Morales & Matejek 2009; Bhatnagar, S. et al. 2008; Tegmark 1997a) before performing the spatial FFT. This process has been shown to theoretically produce a data product identical to images produced from the traditional FX correlator.

Here we present the first software implementation of the MOFF correlator, and announce the public release of the E-field Parallel Imaging Correlator (EPIC) code. We begin with a technical description of the algorithm in §2, then discuss our particular implementation in §3. We then verify the output data quality from our code in §4 by presenting simulated images from both the EPIC correlator and comparing to a simulated FX correlator. We also demonstrate the performance with real-world data from the LWA. In §5 we analyze the scaling relationships of the algorithm. We identify specific array design classes where the EPIC correlator is computationally more efficient than the FX algorithm. We conclude and discuss future research prospects in §6.

2 MATHEMATICAL FRAMEWORK

We provide a brief summary of the amthematical equivalence of the MOFF and FX correlators detailed in Morales (2011). We first relate the dirty image produced from visibilities to the electric fields of astrophysical sources, then show that operations can be reordered to produce the same images at a lower computational cost.

Electric fields from astrophysical sources, $E(\hat{\mathbf{s}})$, in the sky coordinate system denoted by sine-projected unit vector $\hat{\mathbf{s}}$, propagate towards the observer as:

$$\widetilde{E}(\mathbf{r}) = \int E(\hat{\mathbf{s}})\, e^{-i2\pi\mathbf{r}\cdot\hat{\mathbf{s}}}\, \mathrm{d}^2\hat{\mathbf{s}}, \tag{1}$$

where, $\mathbf{r}$ denotes the observer's location (measured in wavelengths relative to some arbitrary origin) and $\widetilde{E}(\mathbf{r})$ is the propagated electric field. Thus the propagated electric field is a linear superposition of the electric fields emanating from astronomical sources with appropriate complex phases. It can also be described as a Fourier transform of the electric fields in the sky coordinates.

An antenna, a, measures a phased sum of these propagated electric fields over its effective collecting area with an additive receiver noise:

$$\begin{aligned}
\widetilde{E}_a &= \int \widetilde{W}_a(\mathbf{r}-\mathbf{r}_a)\, \widetilde{E}(\mathbf{r})\, \mathrm{d}^2\mathbf{r} + \widetilde{n}_a && (2)\\
&= \int \widetilde{W}_a(\mathbf{r}-\mathbf{r}_a) \left[\int E(\hat{\mathbf{s}})\, e^{-i2\pi\mathbf{r}\cdot\hat{\mathbf{s}}}\, \mathrm{d}^2\hat{\mathbf{s}}\right] \mathrm{d}^2\mathbf{r} + \widetilde{n}_a && (3)\\
&= \int W_a(\hat{\mathbf{s}})\, E(\hat{\mathbf{s}})\, e^{-i2\pi\mathbf{r}_a\cdot\hat{\mathbf{s}}}\, \mathrm{d}^2\hat{\mathbf{s}} + \widetilde{n}_a && (4)
\end{aligned}$$

where, $\widetilde{W}_a(\mathbf{r})$ is the aperture electric field illumination pattern of the antenna and its Fourier transform, $W_a(\hat{\mathbf{s}})$, is the directional antenna voltage response.

Interferometers measure *visibilities* – the degree of coherence between electric fields measured by a pair of antennas (van Cittert 1934; Zernike 1938; Thompson et al. 2001). A visibility, $\widetilde{V}_p$, can be written as:

$$\begin{aligned}
\widetilde{V}_p &= \left\langle \widetilde{E}_a \widetilde{E}_b^\star \right\rangle_t && (5)\\
&= \left\langle \left[\int W_a(\hat{\mathbf{s}})\, E(\hat{\mathbf{s}})\, e^{-i2\pi\mathbf{r}_a\cdot\hat{\mathbf{s}}}\, \mathrm{d}^2\hat{\mathbf{s}} + \widetilde{n}_a\right]\right. \\
&\qquad \times \left.\left[\int W_b^\star(\hat{\mathbf{s}}')\, E^\star(\hat{\mathbf{s}}')\, e^{i2\pi\mathbf{r}_b\cdot\hat{\mathbf{s}}'}\, \mathrm{d}^2\hat{\mathbf{s}}' + \widetilde{n}_b^\star\right]\right\rangle_t && (6)\\
&= \iint W_a(\hat{\mathbf{s}}) W_b^\star(\hat{\mathbf{s}}') \left\langle E(\hat{\mathbf{s}}) E^\star(\hat{\mathbf{s}}')\right\rangle_t\, e^{-i2\pi(\mathbf{r}_a\cdot\hat{\mathbf{s}}-\mathbf{r}_b\cdot\hat{\mathbf{s}}')}\, \mathrm{d}^2\hat{\mathbf{s}}\, \mathrm{d}^2\hat{\mathbf{s}}', && (7)
\end{aligned}$$

where we have brought the time average into the integral under the assumption that the aperture illumination pattern does not change over the time-scale of the averaging. This expression can be further simplified with the sky brightness, $I(\hat{\mathbf{s}}) = \left\langle E(\hat{\mathbf{s}}) E^\star(\hat{\mathbf{s}}')\right\rangle_t \delta(\hat{\mathbf{s}}-\hat{\mathbf{s}}')$, and defining the antenna pair sky power response function (or the primary beam), $B_p(\hat{\mathbf{s}}) \equiv W_a(\hat{\mathbf{s}})\, W_b^\star(\hat{\mathbf{s}})$. The result is the visibility expressed in terms of the sky brightness, the primary beam, and uncorrelated noise terms which we group into $\widetilde{n}_p$,

$$\widetilde{V}_p = \int e^{-i2\pi\mathbf{u}_p\cdot\hat{\mathbf{s}}}\, B_p(\hat{\mathbf{s}})\, I(\hat{\mathbf{s}})\, \mathrm{d}^2\hat{\mathbf{s}} + \widetilde{n}_p, \tag{8}$$

where the baseline coordinate $\mathbf{u}_p = \mathbf{r}_a - \mathbf{r}_b$ is the vector separation between the two antennas. This signifies that the visibility ($\widetilde{V}_p$) measured between a pair of antennas (p) is obtained by the multiplying the sky brightness $I(\hat{\mathbf{s}})$ by the antenna power response $B(\hat{\mathbf{s}})$ and Fourier transforming from the directional coordinates ($\hat{\mathbf{s}}$) to uv coordinates, which are then sampled at the locations of the antenna spacings (or baselines), namely, $\mathbf{u}_p$, and added to the receiver noise n_p.

This can be equivalently re-written as:

$$\widetilde{V}_p = \int \tilde{B}(\mathbf{u}'-\mathbf{u}) \times \left[\int e^{-i2\pi\mathbf{u}.\hat{\mathbf{s}}}\, I(\hat{\mathbf{s}})\, \mathrm{d}^2\hat{\mathbf{s}}\right] \mathrm{d}^2\mathbf{u} + n_p, \tag{9}$$

where, $\tilde{B}(\mathbf{u})$ denotes the uv-space antenna power response obtained by a Fourier transform of $B(\hat{\mathbf{s}})$. Effectively, the multiplication in image space by $B(\hat{\mathbf{s}})$ has been replaced by a convolution with $\tilde{B}(\mathbf{u})$ in uv-space. This is the software holographic equivalent of traditional FX correlator output.

From here we adopt the matrix notation of Morales (2011), where vectors are represented with single coordinates, and matrices are represented by two coordinates denoting the spaces the operator transforms between. In this notation, the above measurement equation can be expressed as:

$$\mathbf{m}(\mathbf{v}) = \widetilde{\mathbf{B}}(\mathbf{v},\mathbf{u})\, \mathbf{F}(\mathbf{u},\hat{\mathbf{s}})\, \mathbf{I}(\hat{\mathbf{s}}) + \mathbf{n}(\mathbf{v}), \tag{10}$$

where the sky brightness $\mathbf{I}(\hat{\mathbf{s}})$ is Fourier transformed using $\mathbf{F}(\mathbf{u},\hat{\mathbf{s}})$

and the resultant spatial coherence function is weighted and summed using the antenna power response, $\widetilde{\mathbf{B}}(\mathbf{v},\mathbf{u})$ in uv-space sampled at the baseline location to obtain the measured visibilities:

$$\mathbf{m}(\mathbf{v}) = \left\langle \widetilde{\mathbf{E}}^{\star}(\mathbf{a})\,\widetilde{\mathbf{E}}(\mathbf{a}')\right\rangle_t , \tag{11}$$

where $\mathbf{m}(\mathbf{v})$ denotes visibilities measured by cross-correlating measured antenna electric fields over all possible pairs of $\mathbf{a}$ and $\mathbf{a}'$. It is the same as equation 5 written in matrix notation.

Using the optimal map-making formalism (Tegmark 1997b; Tegmark 1997a), a software holography image is formed using (Morales & Matejek 2009):

$$\mathbf{I}'(\hat{\mathbf{s}}) = \mathbf{F}^{\mathrm{T}}(\hat{\mathbf{s}},\mathbf{u})\,\widetilde{\mathbf{B}}^{\mathrm{T}}(\mathbf{u},\mathbf{v})\,\mathbf{N}^{-1}(\mathbf{v},\mathbf{v})\,\mathbf{m}(\mathbf{v}) \tag{12}$$

where the measured visibilities are weighted by the inverse of the system noise, followed by a gridding process using the holographic antenna power response as the gridding kernel, followed by a Fourier transform to create an image $\mathbf{I}'(\hat{\mathbf{s}})$. This is the optimal estimate of the true image $\mathbf{I}(\hat{\mathbf{s}})$ given the visibility measurements.

The intermediate step of gridding with the antenna power response can be expressed as a convolution of a data vector generated by gridding the electric fields directly with the antenna illumination pattern.

$$\widetilde{\mathbf{B}}^{\mathrm{T}}(\mathbf{u},\mathbf{v})\,\mathbf{N}^{-1}(\mathbf{v},\mathbf{v})\,\mathbf{m}(\mathbf{v}) = \left\langle \left[\widetilde{\mathbf{W}}_a^{\mathrm{T}}(\mathbf{r},\mathbf{a})\,\widetilde{\mathbf{N}}^{-1}(\mathbf{a},\mathbf{a})\,\widetilde{\mathbf{E}}(\mathbf{a})\right] * \left[\widetilde{\mathbf{W}}_a(\mathbf{r},\mathbf{a})\,\mathbf{N}^{-1}(\mathbf{a},\mathbf{a})\,\mathbf{E}^{\star}(\mathbf{a})\right]\right\rangle_t \tag{13}$$

We can then use the multiplication-convolution theorem to move the convolution in Equation 13 to a square after the Fourier transform in Equation 12.

$$\mathbf{I}'(\hat{\mathbf{s}}) = \left\langle \left| \mathbf{F}^{\mathrm{T}}(\hat{\mathbf{s}},\mathbf{r})\,\widetilde{\mathbf{W}}^{\mathrm{T}}(\mathbf{r},\mathbf{a})\,\widetilde{\mathbf{N}}^{-1}(\mathbf{a},\mathbf{a})\,\widetilde{\mathbf{E}}(\mathbf{a}) \right|^2 \right\rangle_t . \tag{14}$$

The term inside the angular brackets before squaring has a very similar form as that in equation 12. It signifies that the measured antenna electric fields are weighted by the antenna noise, weighted and gridded by the antenna aperture kernel, Fourier transformed and finally squared to obtain the same image estimated that would have been obtained using equation 12.

Equation 14 is the optimal imaging equation used by the MOFF algorithm. While mathematically equivalent to Equation 12, squaring in image space rather than convolving in uv space potentially saves orders of magnitude in computation.

There are some important differences between the two techniques:

(i) The time-averaging cannot be performed on a stochastic measurement but only on its statistical properties. In FX imaging, the visibilities measured between antenna pairs represent spatial correlations which can be time-averaged followed by gridding and imaging. However, in MOFF imaging both antenna and gridded electric fields are stochastic and therefore must be imaged and squared before time-averaging.

(ii) In FX imaging, electric fields measured by antennas are not correlated with themselves and hence lack zero spacing measurements. In contrast, in MOFF imaging, since the gridded electric fields are imaged and squared, they retain information from auto-correlated electric fields at zero spacing and thus yield the true total power of the imaged field.

3 SOFTWARE IMPLEMENTATION

We have implemented the MOFF imaging technique in our "E-field Parallel Imaging Correlator" – a highly parallelized Object Oriented Python package, [2] now publicly available. Besides implementing the MOFF imaging algorithm it also includes FX imaging using the software holography technique and a simulator for generating electric fields from a sky model.

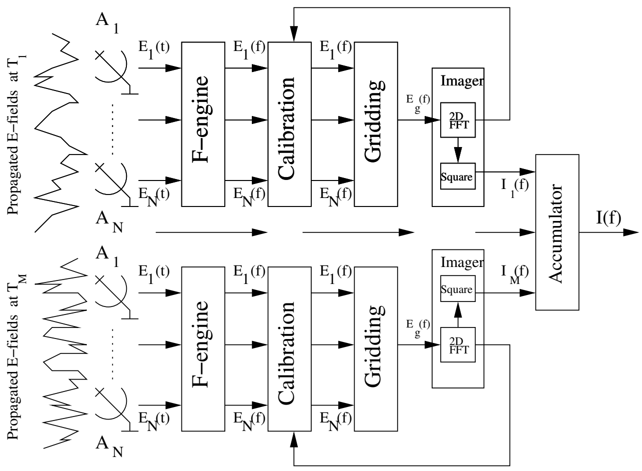

Figure 1. A flowchart of MOFF imaging in EPIC. The propagated electric fields shown on the left are measured as time-series $E_1(t)\ldots E_{\mathrm{N}}(t)$ by the antennas which are then Fourier transformed by the F-engine to produce electric field spectra $E_1(f)\ldots E_{\mathrm{N}}(f)$. They are calibrated and gridded. The gridded electric fields $E_{\mathrm{g}}(f)$ from each time series are imaged to produce an images $I_1(f)\ldots I_{\mathrm{N}}(f)$. These images are time-averaged to obtain the final image $I(f)$.

Fig. 1 shows the flowchart for MOFF imaging. The propagated electric fields are shown on the left at different time stamps, $t_1 \ldots t_{\mathrm{M}}$. At each time stamp, the electric fields measured by antennas are denoted by $E_1(t)\ldots E_{\mathrm{N}}(t)$. The F-engine performs a temporal Fourier transform on the electric field time-series to obtain electric field spectra $E_1(f)\ldots E_{\mathrm{N}}(f)$ ($\widetilde{\mathbf{E}}(\mathbf{a})$ in matrix notation) for each of the antennas. Each of the complex antenna gains are calibrated to correct the corresponding electric field spectra. These calibrated electric fields are gridded using an antenna-based gridding convolution function after which it is spatially Fourier transformed and squared to obtain images for every time stamp. These images are then time-averaged to obtain the accumulated image $I(f)$ ($\mathbf{I}(\hat{\mathbf{s}})$ in matrix notation).

Fig. 2 shows the flowchart for software holographic imaging from a FX correlator. The antenna-based F-engine is identical to that in the MOFF processing. The electric field spectra from each antenna are then cross- multiplied in the X-engine with those from all other antennas to obtain the visibilities $V_{\mathrm{ij}}(f)$ ($\mathbf{m}(\mathbf{v})$ in matrix notation). They are calibrated and time-averaged to obtain $\langle V_{\mathrm{ij}}(f)\rangle$ which are then gridded and imaged to obtain the image $I(f)$. The $I(f)$ obtained from both techniques are identical as explained in §2.

Here we discuss the components of these architectures in detail.

3.1 Temporal Fourier transform

This module is common to the MOFF and FX imaging techniques. Time samples of electric fields measured by the antenna and digitized by the A/D converter is Fourier transformed to generate electric field spectra. This step can be parallelized by antennas as shown in

[2] EPIC package can be accessed at https://github.com/nithyanandan/EPIC

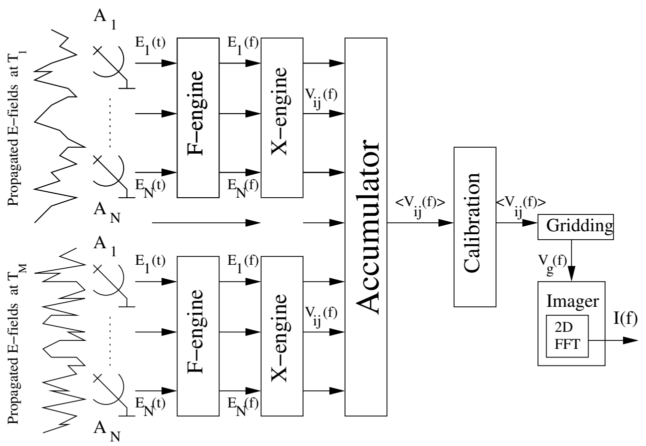

Figure 2. A flowchart of FX imaging in EPIC. The FX process flow shares the F-engine with the MOFF process. Following the F-engine, the electric fields pass through the X-engine to obtain visibilities $V_{\rm ij}(f)$ which are calibrated and time-averaged. Then they are gridded to obtain the gridded visibilities $V_{\rm g}(f)$ which are then imaged to obtain the image $I(f)$.

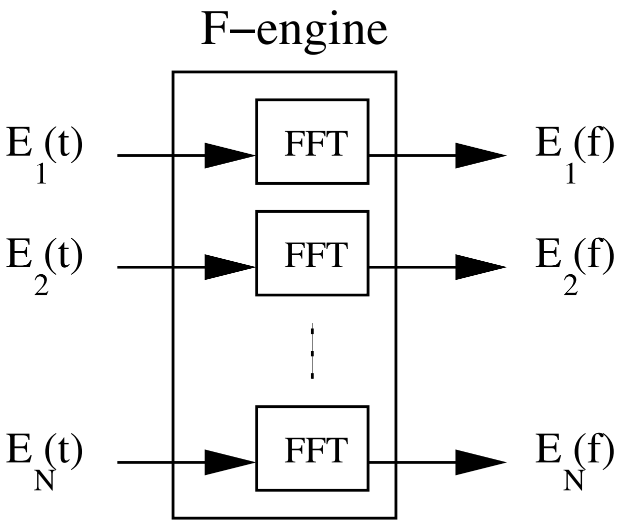

Figure 3. Block diagram of a F-engine. The electric field data streams from antennas are Fourier transformed in parallel to generate electric field spectra.

Fig. 3. The output is then fed to either MOFF and FX imaging pipelines.

3.2 Antenna-to-Grid Mapping

A grid is generated on the coordinate system in which antenna locations are specified with a grid spacing that is at most $\lambda_{\rm min}/2$ even at the highest frequency to ensure there is no aliasing even from regions of the sky far away from the field of view. The number of locations on the grid is restricted to be a power of 2.

The gridding kernel in the simplest case is given by the antenna aperture illumination function, $\widetilde{B}(\mathbf{r}-\mathbf{r}_a)$, which can be specified either by a functional form or as a table of values against locations around the antennas. A nearest neighbor mapping from all antenna footprints to grid locations is created using an efficient k-d tree algorithm (Maneewongvatana & Mount 1999). There is no restriction here that the aperture illumination function has to be identical across antennas.

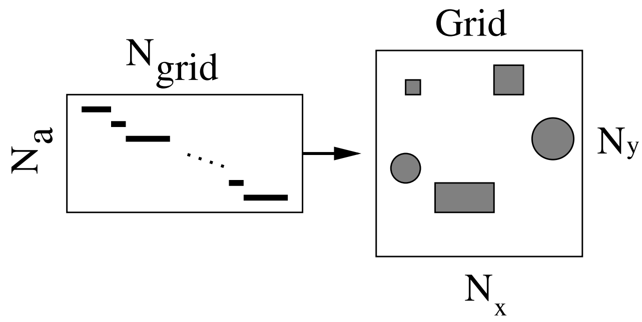

Figure 4. Block diagram of an antenna-to-grid mapping. A sparse block-diagonal matrix of total size $N_{\rm grid} \times N_{\rm a}$ is created where each block contains roughly the number of pixels covered by the respective kernel. The antenna aperture illumination kernels do not have to be identical to each other. A discrete set of arbitrarily placed antennas are now placed onto a regular grid.

In the most general case, this gridding kernel could contain information on the w-projection effect, and even other time-dependent ionospheric effects. For a stationary antenna array in the absence of any time-dependent effects, this mapping must only be determined once in the antenna array coordinate frame. The antenna-to-grid mapping matrix, $\mathbf{M}(\mathbf{r},\mathbf{a})$ is described as a transformation matrix from the space of measured electric fields by the antennas ($\mathbf{a}$) to the antenna array grid denoted by the coordinate $\mathbf{r}$. Since each antenna occupies a footprint typically the size of its aperture, $\mathbf{M}(\mathbf{r},\mathbf{a})$, which is generally of size $N_{\rm grid} \times N_{\rm a}$, reduces to a sparse block-diagonal matrix with only $N_{\rm a}$ blocks and roughly $N_{\rm k}$ non-zero entries per block. This sparse matrix is stored in a Compressed Sparse Row (CSR) format. Fig. 4 illustrates the antenna-to-grid mapping matrix and the grid containing the mapped aperture footprints of the antennas.

3.3 Calibration

Calibration of direct imaging correlators remains a challenge. Contrary to the FX data flow, direct imagers mix the signals from all antennas before averaging and writing to disk. It is therefore essential to apply gain solutions before the gridding step. Previous efforts have resorted to applying FX-generated calibration solutions (Zheng et al. 2014; Foster et al. 2014), or integrating a dedicated FX correlator which periodically forms the full visibility matrix (Wijnholds & van der Veen 2009; de Vos et al. 2009).

In a companion paper to appear soon, we demonstrate a novel calibration technique (EPICal) which leverages the data products formed by direct imaging correlators to estimate antenna complex gains. This method correlates the antenna electric field signals with an image pixel form the output of the correlator in the feedback calibration fashion outlined in Morales 2011 (illustrated in Fig. 1 by the arrow leading from the imager to the calibration block). Furthermore it allows for arbitrarily complex sky models, and following the MOFF algorithm places no restriction on array layout, and accounts for non-identical antenna beam patterns. Because only a single correlation is needed for each antenna, the computation complexity scales only as $N_{\rm ant}$.

The calibration module included in the EPIC repository allows for application of pre- determined calibration solutions, or can solve for the complex gains using the EPICal algorithm.

3.4 Gridding Convolution

The antenna array aperture illumination over the entire grid, $\widetilde{\mathbf{W}}(\mathbf{r})$, is obtained by a projection of the individual antenna aperture illuminations:

$$\widetilde{\mathbf{W}}(\mathbf{r}) = \sum_a \widetilde{\mathbf{W}}_a(\mathbf{r}-\mathbf{r}_a) \tag{15}$$

$$= \mathbf{M}(\mathbf{r},\mathbf{a})\,\mathcal{I}(\mathbf{a}), \tag{16}$$

where, $\mathcal{I}(\mathbf{a})$ is a row of ones. This is achieved by efficient multiplication with the sparse matrix created in the antenna-to-grid mapping process using the sparse matrices module in Python's SciPy package. Unless $\widetilde{\mathbf{W}}(\mathbf{r})$ includes time-dependent effects of the ionosphere or the instrument, it needs to be computed just once for the entire observation. However, the gridding of electric fields must be computed at every readout of the electric field spectra,

$$\widetilde{\mathbf{E}}(\mathbf{r}) = \mathbf{M}(\mathbf{r},\mathbf{a})\,\widetilde{\mathbf{E}}(\mathbf{a}). \tag{17}$$

3.5 Spatial Fourier Transform

Before the spatial Fourier transform, the gridded electric fields are padded with zeros in order to match the grid size and angular size of each image pixel that would have been obtained with software holography of output from an FX correlator.

In MOFF imaging, these are spatially Fourier transformed followed by a squaring operation at every timestamp for every frequency channel. In FX imaging, the spatial Fourier transform is performed only once per integration timescale and does not include a squaring operation.

3.6 Time-averaging

In MOFF imaging, the measured antenna electric fields and the corresponding holographic electric field images are zero-mean stochastic quantities. Hence, they cannot be time-averaged to reduce noise. The statistical quantity stable with time in this case are the square of the holographic electric field images. Thus, squared images have to be formed at every instant of time before averaging as indicated in equation 14.

In contrast, visibilities measured by an antenna are statistically stable within an integration time interval. Hence, they are averaged after calibration as shown in equation 5. It is advantageous to average them in visibilities before imaging because the repeated cost of spatial FFT can be avoided. Since this averaging has been performed already on the visibilities over an integration timescale, the imaging step has to be performed only once per integration cycle.

3.7 Dealing with antenna auto-correlations

The squaring operation under MOFF imaging in the image plane introduces antenna auto-correlations around the zero spacing in the *uv*-plane which are absent in traditional visibility-based imaging. In order to facilitate a robust comparison, these auto-correlations are removed from both the *uv* and image planes, which is otherwise not an essential part of the core algorithm. We describe below how they can be removed from the *uv* plane in a straightforward manner.

The shape and extent of these auto-correlations can be estimated from the antenna aperture illumination pattern. The aperture illumination patterns are already available from the gridding step. Fig. 5 shows the estimated weights from antenna auto-correlations in the *uv*-plane (left) and the corresponding response in the image plane (right). The latter is simply the directional antenna power response.

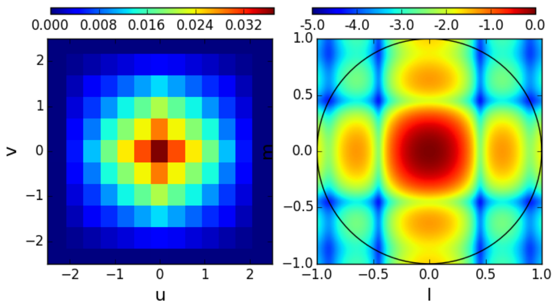

Figure 5. The auto-correlation of weights of a square shaped antenna aperture in the uv plane (left) and the corresponding directional antenna power response on the sky (right) in coordinates specified by direction cosines. The antenna auto-correlation weights are normalized to a sum of unity yielding a peak response of unity in the antenna's directional power pattern on the sky. The color scale for the directional power pattern is logarithmic. The black circle indicates the sky horizon and values beyond it are not physical and hence ignored.

We inverse Fourier transform the squared images and beams back to the *uv* plane and subtract the estimated auto-correlation kernel scaled to the peak value centered at the zero spacing pixel. The final averaged image is obtained by Fourier transforming the *uv* plane data and weights with the auto-correlations subtracted to the image plane. These images are now comparable to those obtained from visibility-based imaging. This step of removing auto-correlations is required to be performed only once per integration timescale and does not add significant cost to the full operation.

4 VERIFICATION

In order to prove the accuracy of the EPIC code, we verify the images produced through simulations. We simulate electric field streams from a model sky and process the data through both the MOFF and visibility based imaging algorithms. We then compare the output images to demonstrate their equivalence.

4.1 Simulations

We use the EPIC simulator to generate electric field samples from a sky model. In our simulations, we use 64 frequency channels each of width δf = 40 kHz, 10 point sources of flux densities 10 Jy at random locations. The number of timestamps integrated in one integration cycle was kept at eight where each A/D timeseries is $1/\delta f$ = 25 μs long. We use the MWA array layout (Beardsley et al. 2012) for demonstration. Only the inner 51 tiles within a square bounding box of 150 m on each side were used. We assumed all tiles are identical and have a square shaped electric field illumination footprint 4.4 m on each side.

Fig. 6 shows the dirty images (top) and synthesized beams (bottom) obtained with antenna-based MOFF and FX visibility-based imaging algorithms packaged in EPIC. The antenna auto-correlations that correspond to zero spacing have been removed from the MOFF image and the corresponding synthesized beam. The sky positions of the simulated sources are indicated by solid black circles. The reconstructed sky image has the simulated sources

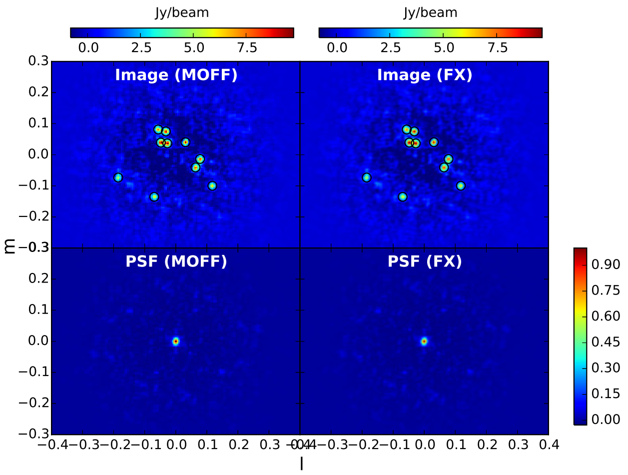

Figure 6. Dirty images (top) and synthesized beams (bottom) obtained from simulated data using antenna-based MOFF algorithm (left) and FX visibility-based software holography (right). The solid black circles in the top panels indicate the simulated source positions. The antenna auto-correlations at zero-spacing have been removed from the MOFF images. The images in either case reconstruct the sources at the right locations with the fluxes expected after multiplication by the antenna power pattern. The synthesized beams from the two algorithms are well matches in size and shape. The overall modulation by the power pattern is seen clearly in both images.

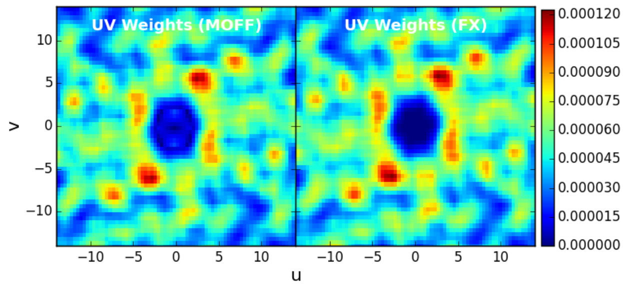

Figure 7. Dirty images (top) and synthesized beams (bottom) obtained from simulated data using antenna-based MOFF algorithm (left) and FX visibility-based software holography (right). The solid black circles in the top panels indicate the simulated source positions. The antenna auto-correlations at zero-spacing have been removed from the MOFF images. The images in either case reconstruct the sources at the right locations with the fluxes expected after multiplication by the antenna power pattern. The synthesized beams from the two algorithms are well matched in size and shape. The overall modulation by the power pattern is seen clearly in both images.

at the expected sky positions in either case. Both algorithms result in images and synthesized beams that are well matched with each other. Their fluxes are modulated by a multiplicative power pattern corresponding to that of a uniform square aperture.

4.2 Comparison of outputs

We investigate the two imaging algorithms for differences from the point of view of the quality of their outputs. We begin by comparing the gridded cross-correlation weights in the *uv*-plane. In MOFF imaging, weights from antenna auto-correlations have been removed as described in §3.7.

Fig. 7 shows the cross-correlation weights obtained with MOFF imaging (left) and visibility-based imaging (right). The first notable difference is in the weights around zero spacing. Though both show a dominant void around zero-spacing, the void obtained with MOFF algorithm shows many pixels with non-zero weights. In contrast, the zero-spacing void from traditional imaging consists of predominantly zero-valued pixels. The gridding process in the former involves rounding the antenna footprint to the nearest grid pixel. Depending on the exact location of the center of the antenna relative to the grid, the grid pixels that receive contribution from an antenna may be a pixel narrower along one or both axes relative to that from another identical antenna but with a different center location relative to the grid. The resulting auto-correlation footprint will also not necessarily be identical. Hence, using a single expected footprint for subtraction will typically leave some residuals behind as seen in the void region in the weights obtained with MOFF imaging.

These residuals can be mitigated by:

(i) making the grid spacing finer which makes the rounding error less susceptible to the location of the antenna center relative to the grid, and

(ii) subtracting each auto-correlation of antenna weights separately by using the shape and extent of the footprint appropriate for that specific antenna aperture.

The latter is the most general solution applicable especially in the case of heterogeneous antenna arrays and is under active development for EPIC.

The other notable difference is that weights outside the zero-spacing void in some regions are different from each other at the few percent level though the sum of weights in these regions are identical. For instance, note the difference in weights at $(u, v) \approx (2, -10)$. This is found to arise because with MOFF imaging the antenna locations were rounded to the nearest grid pixel whereas in visibility-based imaging the baseline locations (difference between antenna locations) were rounded to the nearest grid pixel. Since rounding is not a commutative operation (i.e. rounding of the difference value is not necessarily equal to the difference of the rounded values), the gridding operation in the two cases introduces rounding errors in the placement of antenna aperture and cross-correlated aperture weights. Each antenna aperture weight projected to the nearest grid location can be displaced by $\sim$ 1–2 pixels. This effect can be mitigated only by making the grid spacing finer at the expense of increased computational cost.

We study the effect of the differences in gridded weights on the image plane. Fig. 8 shows the difference between the synthesized beams obtained with the two methods. A difference map between the two synthesized beams is shown on top. The amplitude of the difference appears to be modulated by the directional power response of the antenna. At the bottom, in radial bins, the rms of the synthesized beam (gray) and the rms of the difference map (black) are plotted in percentage units relative to the peak (to be read using the axis on the left side of the plot). The antenna power pattern (red; to be read using the scale on the right) is plotted for reference.

The synthesized beam rms is proportional to the antenna power pattern as expected from a point spread function uncorrected for the antenna power pattern. The rms of differenced synthesized beams is also modulated by the antenna power pattern. The rms of the difference is definitely lesser than the rms of the synthesized beam in the central regions up to $(l^2+m^2)^{1/2} \lesssim 0.3$. This implies that the beams are well matched in the central regions. In the outer regions,

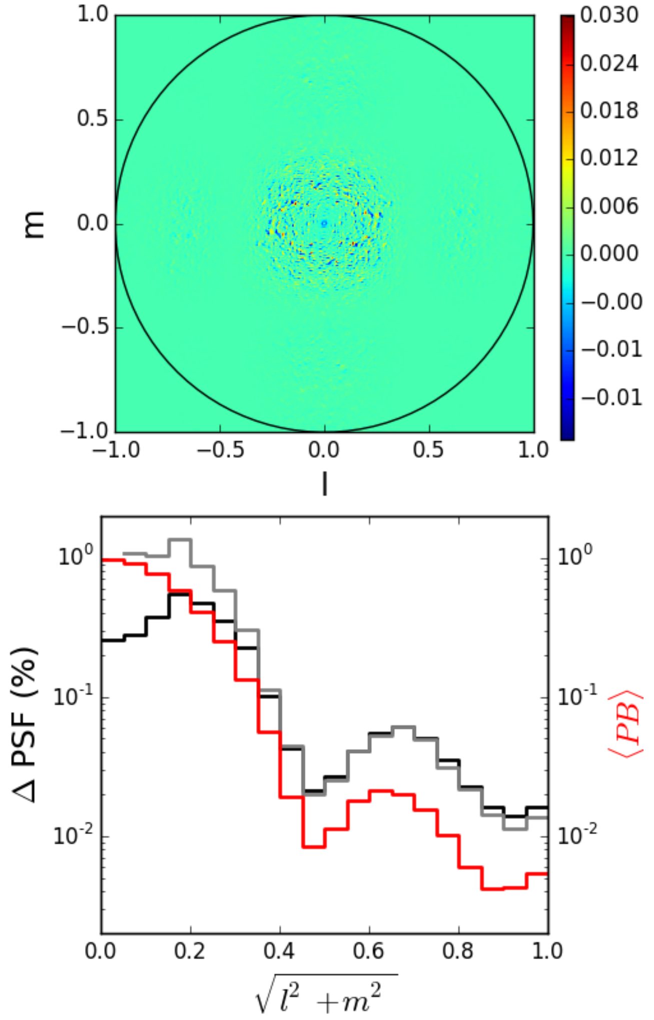

Figure 8. Map of difference between the synthesized beams obtained with the two methods (top) and radial statistics of the synthesized beams and their differences (bottom). The maximum difference is of the order of a few percent. The difference appears to be modulated in amplitude by the power pattern of the antenna.

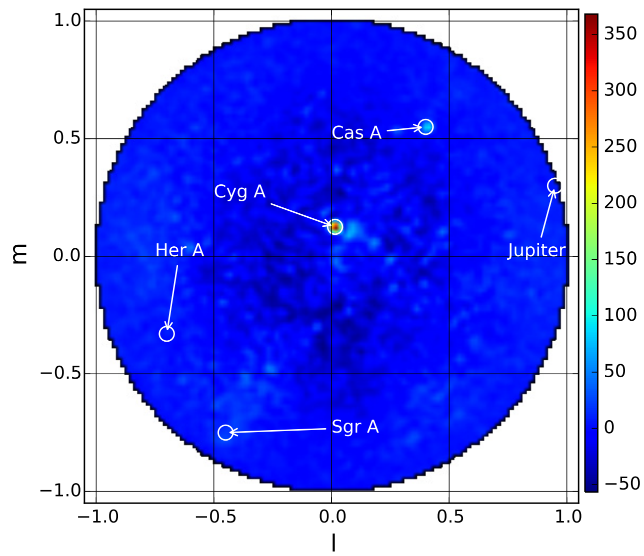

Figure 9. Image from LWA TBN data obtained with MOFF imaging using EPIC package after averaging over 2 s and ≈ 80 kHz. The x- and y-axes denote direction cosines l and m respectively. The antenna voltages are compensated for their respective delays. The flux scale is arbitrary and not calibrated. Locations of Cyg A, Cas A, Sgr A, Her A and Jupiter are annotated.

their mismatch is comparable to the rms of synthesized beams. This indicates the two synthesized beams are not completely randomly different from each other in which case the rms of the difference would have been $\approx \sqrt{2}$ higher than the rms of the each of the synthesized beams. This indicates that while differences exist, large fractions of them are still well matched to each other even out to the horizon. Thus the rounding errors in gridding do not affect the statistics of the images or the synthesized beams.

4.3 Application to LWA data

Here we demonstrate our software using narrow band data from the LWA in New Mexico. This data is in LWA narrow-band transient buffer (TBN) format with 512 voltage time samples from 255 antennas within roughly a diameter of 100 m. The data is centered at a frequency of 74.03 MHz, with a sample rate (equal to the bandwidth) of 100 kHz with 512 complex time samples in a A/D writeout timescale of 5.12 ms, a frequency resolution of 195.3125 Hz and dual polarization. There are 391 such timestamps yielding a total duration of 2 s.

We corrected the cable delays, but otherwise assume the data is sufficiently calibrated to image directly. A test of EPICal on this data will be conducted in the future.

Fig. 9 shows the image produced with MOFF imaging packaged in EPIC after averaging over the entire 2 s of data and the inner ≈ 80% of bandwidth (roughly 80 kHz). The image is shown in direction cosine coordinates – l along x-axis and m along y-axis. The flux scale is arbitrary. Even in this proof-of-concept demonstration, we see Cyg A and Cas A prominently and Sgr A faintly as annotated, thus validating the functionality of the EPIC package.

5 ANALYSIS AND FEASIBILITY

We now investigate the feasibility of implementing the EPIC imager on current and future radio telescopes.

We have profiled the core routines of EPIC line-by-line for various ranges of parameters such as antenna filling fraction, maximum baseline length, bandwidth and frequency resolution, integration timescale, etc. for HERA antenna layouts which are highly compact. However, we note that in general, the hardware and optimization of routines in place will determine the relative speeds of the different stages in the pipeline.

Of all steps in the MOFF pipeline that are repeated for every writeout from the F-engine, the slowest step even for dense HERA layouts is found to be the spatial two-dimensional FFT in the imaging stage relative to applying the sparse matrix gridding convolution, squaring or time-averaging. For instance, even in the dense array layout scenario that makes these stages perform the slowest, the gridding convolution, squaring and time-averaging take up only $\simeq 40\%$, $\simeq 40\%$ and $\simeq 15\%$ respectively of the time taken by the spatial Fourier transform. With sparser arrays the gridding process will be be even faster.

In the visibility-based imaging, the predominant computational cost is at the X-engine requiring N_a^2 complex multiplications per channel per correlator writeout timescale.

In the following discussions, we will assume that the computational cost for the MOFF imaging is determined by the spatial Fourier transform while that for visibility-based imaging comes from N_a^2 cross-correlations. However, if non-linearities such as non-coplanarity of baselines (Cornwell et al. 2008) and wide-field phenomena like the *pitchfork* effect (Thyagarajan et al. 2015a,b) are to be corrected for, the antenna based illumination footprint can start becoming less compact in the measurement plane and can result in a costlier gridding process.

The number of complex multiplications and additions in the spatial Fourier transform implemented via Fast Fourier Transform (FFT; Cooley & Tukey 1965) is $\approx \beta N_\mathrm{g} \log N_\mathrm{g}$ where N_g is the number of pixels on the grid and β is a constant that depends on the implementation of twiddle FFT algorithms (Brigham 1974). In our study, we set $\beta = 5$, a value much more conservative than was indicated in Morales (2011). We set the number of complex multiplications in the X-engine in visibility based imaging to N_a^2.

We consider a variety of current and planned radio telescopes. Their antenna layouts are summarized in Table 1. The size of the layout gives the maximum baseline b_max. The grid spacing is determined by the science goals of the experiment in general. For our purpose, we assume a typical requirement that only the field of view of the antenna is to be imaged. This sets the grid spacing to be equal to the size of the antenna, A_a. Hence, $N_\mathrm{g} \simeq b_\mathrm{max}^2/A_\mathrm{a}$.

Fig. 10 shows the number of complex operations per frequency channel per integration timescale. Telescopes that fall to the left of this line indicate MOFF imaging is computationally more efficient than visibility based imaging. All HERA layouts except possibly HERA-19 are in a parameter space where MOFF imaging holds the advantage. The solid line showing future trajectory of HERA like systems will be clearly favoured by MOFF imaging. The gray shaded area is for a projected LWA expansion and is also predominantly in the region favouring MOFF imaging. It is bounded by the LWA1 and LWA-OV on the left and right respectively. The current (see Table 1) and an expanded layout with a four-fold increase in number of elements over a two-fold increase in b_max provide the bounds at the bottom and top respectively. Current instruments such as MWA and LOFAR lie in parameter space favouring visibility based imaging.

We now consider antenna array layouts described by three quantities essential to radio interferometry, namely, maximum baseline length, number of antennas, and the size of each antenna.

Fig. 11 shows the ratio of the number of computations required with visibility based imaging relative to MOFF imaging and the boundaries where this ratio is unity. The color scale refers to the ratio while using a 1 m antenna element. Pairs of black and white lines of the same line style correspond to a specific antenna size. For a given antenna size (annotated on the plot), the white line denotes the minimum baseline length that could achieve the closest packing as a function of the number of antennas. Regions to the right of the white line for a given antenna size imply a physically impossible scenario where antennas will have to be packed overlapping with

Table 1. Radio telescopes and array layouts.

Telescope	Core size b_max (in m)	Number of Antennas N_a	Antenna size A_a (in m^2)	Frequency f_0 (in MHz)
MWA-112[a]	1400	112	16	150
MWA-240[a]	1400	240	16	150
MWA-496[a]	1400	496	16	150
MWA-112[a]	1400	1008	16	150
LOFAR-LC[b]	3500	24	5809	50
LOFAR-HC[b]	3500	48	745	150
LWA1	100	256	10	50
LWA-OV[c]	200	256	10	50
HERA-19	70	19	154	150
HERA-37	98	37	154	150
HERA-331	294	331	154	150
HERA-6769[d]	1330	6769	154	150
SKA1-LC[e]	1000	750	962	150
SKA1-LCD[f]	1000	192,000	962	150
CHIME	100	1280	8	600
HIRAX	200	1024	6	600

[a] MWA-N denotes N tiles in the specified core diameter
[b] LC and HC denotes low band and high band stations inside the specified core diameter
[c] Owens Valley LWA
[d] Hypothetically chosen to have a total collecting area of 1 km^2
[e] This is the number of beamformed stations expected to be in the core, roughly three-fourths of the total number
[f] All dipoles inside the core are used as independent elements without station beamforming

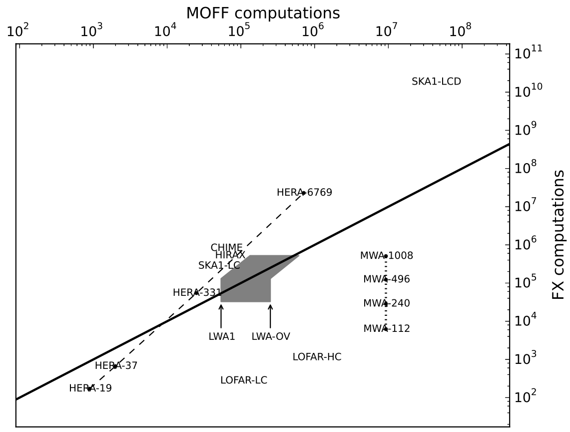

Figure 10. Current and planned instruments in parameter space of number of complex multiplies and adds with MOFF and FX. The dashed line is the boundary at which the number of operations with MOFF and visibility based imaging are equal. MOFF imaging is more efficient for telescopes occupying the left of this line and vice versa. All the HERA layouts except possibly HERA-19 lie in the parameter space favoured by MOFF imaging. And so are SKA1-LC and SKA1-LCD. The solid black curve shows the projected trajectory of bigger close-packed hexagonal layouts similar to HERA. The gray shaded area denotes the projected trajectory of the LWA bounded by LWA1 (left edge), LWA-OV (right edge), current layout (bottom) and a four-fold increase in the number of elements in a two-fold increase in the core size (top). Current instruments such as MWA and LOFAR fall in a region favoured by visibility based imaging.

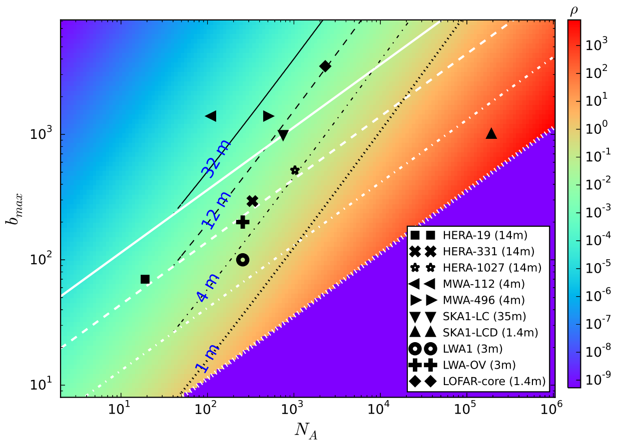

Figure 11. Current and instruments planned for future in parameter space of baseline length and number of antennas with MOFF and FX. Color scale shown is for the ratio of number of computations required by visibility based imaging to that by the MOFF algorithm. Different line styles denote different antenna sizes annotated. For a given antenna size, the corresponding white line denotes the maximum number of antennas that can be packed inside various baseline lengths. The region to the right of the white lines for corresponding antenna size is physically disallowed. Regions to the left of the black lines of the same corresponding antenna size (or line style) favours usage of visibility based imaging. Region inside the wedge enclosed by the black and white lines favours using the MOFF algorithm for any given antenna size. These wedges shift leftward with increasing antenna size. HERA-331, HERA-1027, LWA1, SKA1-LC and SKA1-LCD are inside their corresponding wedges that favour the usage of MOFF imaging.

each other. For the same antenna size, the black line denotes the boundary where the ratio of the number of computations with either algorithm becomes unity. Regions to the right of the black lines favour usage of the MOFF algorithm. Thus the wedge between the white and black lines denotes the region in parameter space where MOFF algorithm holds a definite advantage while visibility based algorithms will be favoured to the left of the black lines. As antenna size increases the maximum number of antennas for a dense packing as a function of baseline length decreases. Hence the white lines shift leftward as antenna size increases. Similarly, with increase in antenna size, N_g also decreases when field of view imaging is achieved with an increasing grid spacing equal to antenna size and hence lowers the amount of computations required with the MOFF algorithm. This shifts the black curves leftward. HERA-331, HERA-1027, LWA1, SKA1-LC and SKA1-LCD are inside their respective wedges indicating they are in regions that favour the usage of MOFF imaging. For fixed baseline length, regions favouring the MOFF algorithm tend to be towards large N_a indicating large-N dense array layouts with smaller antenna elements are best suited for deploying EPIC.

6 CONCLUSIONS

As radio astronomy is entering a new era, advances in instrumentation have to be accompanied by equal advances in processing techniques to manage computational resources. Many future radio telescopes such as the SKA, HERA and LWA are headed towards the large-N dense array layout model for which computational cost from traditional FX/XF correlator based architecture and visibility based imaging starts rising steeply. We have provided the first software demonstration of a general purpose imaging algorithm using our generic and efficient EPIC software that is designed to bring this cost down from $O(N^2)$ to $O(N \log N)$. Under the class of direct imaging techniques, ours is one of the most generic – neither does it place any constraint on the array layout to be on a regular grid nor does it require the antenna array to be homogeneous.

Our package, now publicly available, written in object oriented Python is highly modularized and parallelizable. It includes an implementation of the MOFF algorithm in addition to visibility based software holography imaging and a data simulator for sky models. It has been successfully tested on simulated as well as real LWA observations.

The MOFF algorithm packaged with EPIC is already found to be most suitable for many present and planned radio telescopes such as the LWA, HERA and SKA. In general, MOFF is most suited to operate in the region of parameter space characterized by dense packing of a large number of antennas especially when consisting of a large number of small antenna elements.

A unique and significant advantage is the instantaneous availability of calibrated time-domain images bundled together with the hardware frontend such as the F-engine as an integrated module. Hence, it is a compelling candidate for time-domain radio astronomy, e.g. search for and monitoring of transients. Transient detection pipeline at the backend of EPIC can be fine-tuned to target fast transients such as the Fast Radio Bursts (FRB; Thornton et al. 2013) on millisecond timescales at GHz frequencies or slow transients from planetary and exoplanetary origins at frequencies around 100 MHz.

Thus, EPIC with the MOFF algorithm packaged is uniquely poised to offer a substantial advantage to imaging with large-N dense arrays typical of next-generation radio telescopes as well as push the frontiers of time-domain astronomy to fill gaps in understanding the science behind phenomena responsible for extreme transient events in the Universe.

In the near future, we plan to upgrade our current Python implementation of EPIC to a GPU based pipeline in order to operate on real-time data and develop a transient trigger and monitor backend. In the meanwhile, we plan to demonstrate the capability of EPIC to calibrate and image from heterogeneous arrays and incorporate corrections for non-coplanarity of baselines and direction-dependence of calibration.

ACKNOWLEDGEMENTS

We thank Greg Taylor for providing us the LWA data, and Larry D'Addario, Gregg Hallinan and Harish Vedantham for their valuable inputs. Construction of the LWA has been supported by the Office of Naval Research under Contract N00014-07-C-0147. Support for operations and continuing development of the LWA1 is provided by the National Science Foundation under grant AST-1139974 of the University Radio Observatory program.

REFERENCES

Bandura K., et al., 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. p. 22 (arXiv:1406.2288), doi:10.1117/12.2054950

Beardsley A. P., et al., 2012, MNRAS, 425, 1781

Bhatnagar, S. Cornwell, T. J. Golap, K. Uson, J. M. 2008, A&A, 487, 419

Bowman J. D., et al., 2013, Publ. Astron. Soc. Australia, 30, 31

Brigham E. O., 1974, The fast Fourier Transform

Bunton J. D., 2004, Experimental Astronomy, 17, 251
Cooley J. W., Tukey J. W., 1965, Math. Comput., 19, 297
Cornwell T. J., Golap K., Bhatnagar S., 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 647
Daishido T., et al., 2000, Proc. SPIE, 4015, 73
Ellingson S. W., et al., 2013, IEEE Transactions on Antennas and Propagation, 61, 2540
Foster G., Hickish J., Magro A., Price D., Zarb Adami K., 2014, Monthly Notices of the Royal Astronomical Society, 439, 3180
Lonsdale C. J., Doeleman S. S., Cappallo R. J., Hewitt J. N., Whitney A. R., 2000, in Butcher H. R., ed., Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4015, Radio Telescopes. pp 126–134
Maneewongvatana S., Mount D. M., 1999, eprint arXiv:cs/9901013,
Mellema G., et al., 2013, Experimental Astronomy, 36, 235
Morales M. F., 2011, PASP, 123, 1265
Morales M. F., Matejek M., 2009, MNRAS, 400, 1814
Otobe E., et al., 1994, PASJ, 46, 503
Parsons A. R., et al., 2010, The Astronomical Journal, 139, 1468
Tegmark M., 1997a, Phys. Rev. D, 55, 5895
Tegmark M., 1997b, ApJ, 480, L87
Tegmark M., Zaldarriaga M., 2009, Phys. Rev. D, 79, 083530
Tegmark M., Zaldarriaga M., 2010, Phys. Rev. D, 82, 103501
Thompson A. R., Moran J. M., Swenson Jr. G. W., 2001, Interferometry and Synthesis in Radio Astronomy, 2nd Edition. Wiley
Thornton D., et al., 2013, Science, 341, 53
Thyagarajan N., et al., 2015a, ApJ, 804, 14
Thyagarajan N., et al., 2015b, ApJ, 807, L28
Tingay S. J., et al., 2013, PASA - Publications of the Astronomical Society of Australia, 30
Wijnholds S., van der Veen A.-J., 2009, Signal Processing, IEEE Transactions on, 57, 3512
Zernike F., 1938, Physica, 5, 785
Zheng H., et al., 2014, MNRAS, 445, 1084
de Vos M., Gunst A., Nijboer R., 2009, Proceedings of the IEEE, 97, 1431
van Cittert P. H., 1934, Physica, 1, 201
van Haarlem M. P., et al., 2013, A&A, 556, A2

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