




Research Article

Comparison of fast, hybrid imaging architectures for multi-scale, hierarchical aperture arrays

Nithyanandan Thyagarajan 

Commonwealth Scientific and Industrial Research Organisation (CSIRO), Space & Astronomy, P. O. Box 1130, Bentley, WA 6102, Australia

Abstract

Two major areas of modern radio astronomy, namely, explosive astrophysical transient phenomena and observations of cosmological structures, are driving the design of aperture arrays towards large numbers of low-cost elements consisting of multiple spatial scales spanning the dimensions of individual elements, the size of stations (groupings of individual elements), and the spacing between stations. Such multi-scale, hierarchical aperture arrays require a combination of data processing architectures – pre-correlation beamformer, generic version of fast Fourier transform (FFT)-based direct imager, post-correlation beamformer, and post-correlation FFT imager – operating on different ranges of spatial scales to obtain optimal performance in imaging the entire field of view. Adopting a computational cost metric based on the number of floating point operations, its distribution over the dimensions of discovery space, namely, field of view, angular resolution, polarisation, frequency, and time is examined to determine the most efficient hybrid architectures over the parameter space of hierarchical aperture array layouts. Nominal parameters of specific upcoming and planned arrays – the SKA at low frequencies (SKA-low), SKA-low-core, a proposed long baseline extension to SKA-low (LAMBDA-I), Compact All-Sky Phased Array (CASPA), and a lunar array (FarView-core) – are used to determine the most optimal architecture hierarchy for each from a computational standpoint and provide a guide for designing hybrid architectures for multi-scale aperture arrays. For large, dense-packed layouts, a FFT-based direct imager is most efficient for most cadence intervals, and for other layouts that have relatively lesser number of elements or greater sparsity in distribution, the best architecture is more sensitive to the cadence interval, which in turn is determined by the science goals.

Keywords: Astronomical instrumentation; methods and techniques; instrumentation: interferometers; techniques: image processing; techniques: interferometric; telescopes

(Received 2 July 2024; revised 22 November 2024; accepted 25 November 2024)

1. Introduction

The variable and transient radio sky encompasses a vast range of astrophysical phenomena. Notable examples include pulsars (Keane 2013), fast radio bursts (FRBs; Lorimer et al. 2007; Thornton et al. 2013), magnetars and their potential link to FRBs (Bochenek et al. 2020), rotating radio transients (RRAT; McLaughlin et al. 2006), gamma-ray burst (GRB) afterglows, Jovian and solar bursts, flare stars, cataclysmic variables (CVs), exoplanet and stellar outbursts (Zhang et al. 2023), X-ray binaries, novae, supernovae, active galactic nuclei (AGNs), blazars, tidal disruption events, counterparts to gravitational wave (GW) events, and several phenomena yet to be classified (Thyagarajan et al. 2011). Their timescales span orders of magnitude, from nanoseconds to days, across wide frequency ranges and polarisation states (Pietka, Fender, & Keane 2015; Chandra et al. 2016; Nimmo et al. 2022, and references therein). For example, pulsars and their giant radio pulses can operate on timescales from milliseconds to seconds, with millisecond pulsars on milliseconds to tens of milliseconds, while FRBs are active on sub-millisecond to tens of millisecond timescales (Crawford et al. 2022; Gupta

et al. 2022; Snelders et al. 2023). Extremely energetic and ultra-fast phenomena, such as sub-millisecond pulsars and magnetars (Du et al. 2009; Haskell et al. 2018), and nanosecond-scale giant pulses from the Crab pulsar (Hankins et al. 2003; Hankins & Eilek 2007; Eilek & Hankins 2016; Philippov et al. 2019), can reveal the fundamental nature of matter and exotic physics. The effectiveness of discovering these transients depends on the instrument's field of view and the duration of on-sky observation (Cordes 2007), ideally requiring continuous real-time operation over a large instantaneous field of view to capture phenomena ranging from microseconds to days.

Another significant goal of modern radio astronomy is the exploration of structure formation and constraining of cosmological parameters in the early Universe using the redshifted 21 cm line of neutral Hydrogen as a tracer of large-scale structures (Morales & Wyithe 2010; Pritchard & Loeb 2012, and references therein). One prominent example includes probing the Cosmic Dawn and the Epoch of Reionisation signifying the appearance of the first luminous objects in the Universe and their impact on structure formation in the Universe at $z \gtrsim 6$. Another example is to probe the Dark Energy and acceleration of the Universe at $1 \lesssim z \lesssim 4$ using intensity mapping of redshifted 21 cm from HI bound to galaxies (Cosmic Visions 21 cm Collaboration et al. 2018). Such cosmological radio observations probing large-scale structures through spatial power spectrum or imaging require

Email: Nithyanandan.Thyagarajan@csiro.au

Cite this article: Thyagarajan N. (2025) Comparison of fast, hybrid imaging architectures for multi-scale, hierarchical aperture arrays. *Publications of the Astronomical Society of Australia* 42, e001, 1–16. <https://doi.org/10.1017/pasa.2024.121>

interferometer arrays of large fields of view (hundreds to thousands of square degrees), dense sampling of the large-scale spatial modes (tens of kpc to tens of Mpc), and high sensitivity ($\lesssim 1$ mK noise levels).

Based on the requirements for transients and observational cosmology, radio interferometers are witnessing a paradigm shift towards packing a large number of relatively small-sized collecting elements densely in a compact region. Modern radio arrays tend to be aperture arrays that operate interferometrically and hierarchically on multiple spatial scales, involving spatial dimensions of the collecting element, a spatial grouping of elements forming a virtual telescope typically referred to as a station or a tile, and a spatial grouping of stations comprising the full array. Examples of multi-scale, hierarchical aperture arrays include the Murchison Widefield Array (MWA; Tingay et al. 2013), the Hydrogen Epoch of Reionization Array (HERA; DeBoer et al. 2017), the Low Frequency Array (LOFAR; van Haarlem et al. 2013), the swarm of Long Wavelength Arrays (LWA Swarm; Dowell & Taylor 2018), the Hydrogen Intensity and Real-time Analysis Experiment (HIRAX; Crichton et al. 2022), and the SKA-low (Dewdney et al. 2009; Braun et al. 2019).

This new paradigm poses severe computational and data rate challenges necessitated by the processing of large numbers of data streams in short cadence intervals. The choice of data processing architecture is usually sensitive to the array layout, cadence, angular resolution, and field-of-view coverage. For example, a correlator-based architecture is generally more suited for relatively smaller number of interferometer elements sparsely distributed over a large area and processing on a slower cadence, whereas a direct imaging architecture based on the fast Fourier transform (FFT) has far greater efficiency for regularly arranged elements (Daishido et al. 1991; Otobe et al. 1994; Tegmark & Zaldarriaga 2009; Tegmark & Zaldarriaga 2010; Foster et al. 2014; Masui et al. 2019). A generalised version of FFT-based direct imaging like the E-field Parallel Imaging ‘Correlator’ (EPIC; Thyagarajan et al. 2017; Thyagarajan et al. 2019; Krishnan et al. 2023) is expected to be highly efficient for large- N , densely packed arrays with or without regularly placed elements from computational and data rate perspectives. Pre- or post-correlation beamforming is expected to be efficient when the elements and the image locations are few in number. Thus, the choice of an optimal imaging architecture is critical for maximising the discovery potential. If the chosen architecture is inadequate, one or more discovery parameters such as time resolution, on-sky observing time, bandwidth, field of view, or angular resolution must be curtailed to obtain a compromised optimum (for example, Price 2024). For multi-scale arrays that involve hierarchical data processing, a single architecture may not be optimal on all scales. A hybrid architecture that is optimal at various levels of data processing hierarchy is required.

In this paper, the primary motivation is to explore different imaging architectures and their combinations appropriate for aperture arrays spanning a multi-dimensional space of array layout parameters and cadence intervals from a computational viewpoint. Conversely, the paper also provides a guide to choosing the array layout parameters and cadence intervals where a given architecture, or combinations thereof, will remain computationally efficient. Several upcoming and hypothetical multi-scale aperture arrays are used as examples.

The paper is organised as follows. Section 2 describes the parametrisation of multi-scale aperture arrays studied. Section 3 describes the metric of computational cost density in discovery

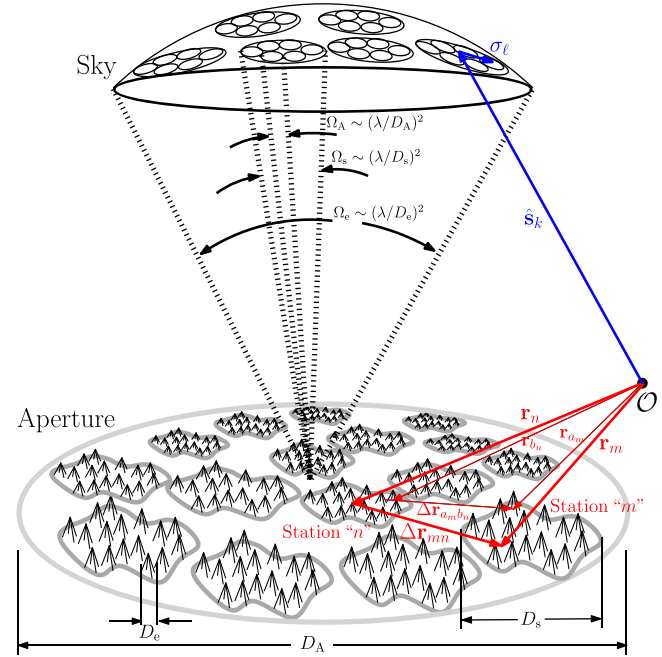


Figure 1. Geometric view of a hierarchical multi-scale aperture array, which consists of N_s stations, each a collection of N_{eps} elements. Stations, m and n , at locations, \mathbf{r}_m and \mathbf{r}_n , respectively, relative to the origin, \mathcal{O} , are separated by Δr_{mn} . Elements, a and b , in stations, m and n , are denoted by a_m and b_n , located at \mathbf{r}_{a_m} and \mathbf{r}_{b_n} , respectively. $\Delta r_{a_m b_n}$ is the separation between these elements. The spatial sizes of the element, station, and the array in the aperture plane are denoted by D_e , D_s , and D_A , respectively. The solid angle, $\Omega_e \sim (\lambda/D_e)^2$, corresponding to the element size determines the overall field of view in the sky plane accessible by the aperture array. Intra-station architectures produce station-sized virtual telescopes that simultaneously produce beams of solid angle, $\Omega_s \sim (\lambda/D_s)^2$, in multiple directions, $\hat{\mathbf{s}}_k$, that can fill Ω_e . Inter-station processing towards any of these beams produces virtual telescopes of size, D_A , that simultaneously produce beams of solid angle, $\Omega_A \sim (\lambda/D_A)^2$, in multiple directions, σ_ℓ , relative to $\hat{\mathbf{s}}_k$, that can fill the solid angle of that beam.

phase space based on which different imaging architectures described in Section 4 are compared. The multi-scale, intra- and inter-station imaging architecture options are described in Sections 5 and 6, respectively. Section 7 contains a discussion and summary of the results, and conclusions are presented in Section 8.

2. Multi-scale, hierarchical aperture arrays

The hierarchical, multi-scale aperture array scenario considered involves spatial scales corresponding to the dimensions of the collecting element, station (a grouping of elements), and the array (a grouping of stations), denoted by D_e , D_s , and D_A , respectively. Fig. 1 illustrates the geometry and notation for hierarchical multi-scale aperture arrays studied in this paper.

The number of stations in the array and the number of elements per station (indexed by m) are denoted by N_s and N_{eps_m} , respectively. The number of elements in all stations are assumed to be equal, $N_{\text{eps}_m} = N_{\text{eps}}$, for convenience. Without loss of generality, this study can be extended to arrays with heterogeneous station densities and layouts such as LOFAR, but the idiosyncrasies of the array will have to be taken into account.

The specific aperture array examples considered are the SKA-low, the core of the SKA-low (SKA-low-core), the first phase of the concept of a long baseline extension to SKA-low called Low-frequency Australian Megametre Baseline

Table 1. Table of hierarchical, multi-scale interferometer array parameters.

Array name	λ (m) ^a	D_A/λ	N_s	D_s/λ	N_{eps}	D_c/λ	f_A^b	f_s^c
LAMBDA-I	2	3.9×10^6	4	17.5	256	1	1.61×10^{-10}	0.836
SKA-low-core	2	5.25×10^2	256	17.5	256	1	0.284	0.836
SKA-low	2	3.5×10^4	512	17.5	256	1	1.28×10^{-4}	0.836
CASPA	0.25	4.6×10^4	3	8.08	65	0.5	9.2×10^{-8}	0.996
FarView-core ^d	10	6.77×10^2	81	38.5	625	1	0.262	0.422

^aNominal operating wavelength of the array. Results may differ at other wavelengths depending on aperture efficiency, filling factor, etc.

^bArray filling factor, $f_A = N_s (D_s/D_A)^2$.

^cStation filling factor, $f_s = N_{\text{eps}} (D_c/D_s)^2$.

^dOnly the 6 km core of the proposed array (Polidan et al. 2024) is considered here.

Demonstrator Array (LAMBDA-I^a), the Compact All-Sky Phased Array (CASPA^b; Luo et al. 2024), and the core of a lunar array (FarView; Polidan et al. 2024) which is hereafter referred to as FarView-core. LAMBDA-I, CASPA, and FarView-core are hypothetical interferometers that are still in concept phase.

The parameters for the chosen examples are summarised in Table 1. Nominal wavelengths, λ , at which the arrays are planned to operate are listed. The computational costs are typically independent of the wavelength and only depend on the spatial frequencies as shown in Table 1. However, the aperture efficiencies can, in general, vary with wavelength. For a fair comparison, the apertures are all assumed to be 100% efficient at the nominal wavelengths chosen, and that the arrays are used for real-time imaging.

Note that, traditionally, cosmological applications such as probing the epoch of reionisation (EoR) using redshifted 21 cm have relied on storing the visibilities and processing them offline. In such cases, although imaging costs like gridding, Discrete Fourier Transform (DFT), or FFT will not apply, the correlations must still be obtained in real-time. With large arrays, the traditional correlator approach may be too expensive. For example, the Packed Ultra-wideband Mapping Array (PUMA; Slosar et al. 2019), a planned radio telescope for cosmology and transients with $\sim 5\,000$ – $32\,000$ elements, is considering an FFT correlator architecture, a specific adaptation of the EPIC architecture for redundant arrays. This study, with a breakdown of costs for different operations, is also relevant for such cosmological arrays.

3. Cost metric

The primary metric for computational cost used in this study is the number of real floating point operations (FLOP) and the density of such operations in the discovery phase space volume whose dimensions span time, frequency, polarisation, and location on the sky. The computational cost will be estimated per independent interval of time (in FLOP per second or FLOPS) for all possible polarisation states per frequency channel ($\delta\nu$) per independent angular resolution element ($\delta\Omega$). In this paper, a discovery phase space volume element comprising of a single frequency channel at a single two-dimensional angular resolution element (independent pixel) will be referred to as a *voxel* ($\delta\nu \delta\Omega$).

The cost estimates rely on the following floating point operations. A complex addition consists of two FLOPs, and a complex multiplication consists of four real multiplications and two

real additions and therefore six FLOPs. A complex multiply-accumulate (CMAC) operation includes one complex multiplication and addition, therefore requiring eight FLOPs.

Besides the computational metric employed in this paper, a practical implementation of any of these architectures requires careful consideration of several other practical factors such as memory bandwidth, input and output data rates, power consumption, and other processing costs such as calibration. These factors, elaborated further in Section 7.4, will vary significantly with real-life implementations and hardware capabilities available during deployment. Therefore, in a complex space of multi-dimensional factors, this work adopts a simple theoretical approach as a first step wherein the budget of floating point operations will form an irreducible requirement regardless of any particular implementation.

4. Fast imaging architectures

Two broad classes of two-stage fast imaging architectures are considered. In both classes, the first stage is intra-station interferometry of the element electric fields using the options of voltage beamforming (BF), E-field Parallel Imaging Correlator (EPIC), beamforming of correlations (XBF), or correlation followed by FFT (XFFT). The second stage involves inter-station combination of the data products from the first stage. The main difference between the two classes is in whether the second stage uses the first stage products coherently or incoherently. Although aperture arrays consisting of only three hierarchical scales are considered here, the analysis can be extended without loss of generality to arrays of even larger levels of spatial hierarchy. A schematic of these hybrid architectures, their computational cost components, and their extension methodology to multiple stages are illustrated in Fig. 2.

Fast imaging involves processing data on multiple time scales. The channelisation of a voltage time stream of length δt with a time resolution of t_s will correspond to spectral channels of width $\delta\nu = 1/\delta t$ and a bandwidth of $B = 1/t_s$. The number of spectral channels is $N_\nu = B/\delta\nu = \delta t/t_s$. Images are expected to be produced at a cadence of t_{acc} , which is typically larger than δt and smaller than the time scale on which calibration needs to be updated. In this paper, the nominal value for δt is $25 \mu\text{s}$, and the possibilities for t_{acc} explored range from $100 \mu\text{s}$ to 10 s.

In the discussions and equations that follow, a real-time calibration is assumed to have been performed on the input voltages. By assuming the system is reasonably stable, a much less frequent ($\ll 1/t_{\text{acc}}$) calibration will be sufficient, which therefore will not significantly affect the total cost budget (Beardsley et al. 2017; Gorthi, Parsons, & Dillon 2021). Although in principle an accurate

^a<https://www.atnf.csiro.au/projects/LAMBDA.html>.

^bPresently, CASPA's stations are intended for localisation and astrometry, rather than for aperture synthesis.

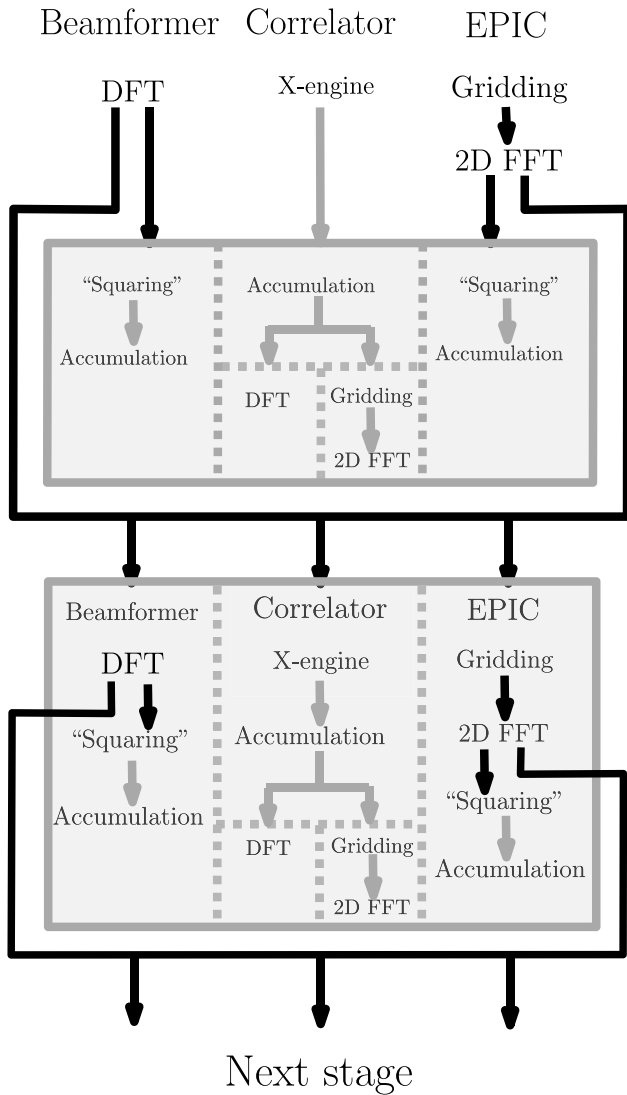


Figure 2. A hybrid, two-stage architecture for hierarchical aperture arrays showing the methodology for extension to multiple stages. Black arrows indicate pathways where the signal consists of electric fields with phase coherence. Grey arrows indicate pathways of a ‘squared’ or a correlated signal with no absolute but only relative phase information. At the top are intra-station architectures, namely, BF, correlator, and EPIC. The top box denotes intra-station imaging options. The bottom box denotes inter-station imaging architectures that process coherent outputs obtained from multiple stations deploying intra-station BF and/or EPIC architectures. The electric fields phase coherently synthesised by BF or EPIC at any stage can form the input for the next stage, thereby allowing the extension of the hierarchy.

calibration is possible in real time, applications such as EoR that require high-accuracy calibration have, for practical reasons, relied on offline calibration.

5. Intra-station imaging architectures

This first stage essentially synthesises station-sized apertures with simultaneous, multiple virtual pointings. These station-sized synthesised apertures from the first stage will act as building blocks for the second stage of inter-station processing. The element, a , in station, m is denoted by a_m . The element and station locations are denoted by \mathbf{r}_{a_m} and \mathbf{r}_m , respectively. Similarly, $\Delta\mathbf{r}_{a_m b_n}$ is the separation between an elements, a and b , in stations, m and n ,

respectively. $\Delta\mathbf{r}_{mn}$ denotes the station separation. This notation is illustrated in Fig. 1. A summary of the intra-station costs are provided in Table 2.

5.1. Voltage Beamforming (BF)

Let $\tilde{E}_{a_m}^p(\nu)$ represent the calibrated and noise-weighted electric field at frequency, ν , in polarisation state, p , measured by a station element, a_m , in station, m . Because each spectral channel is assumed to be processed independently, specifying the frequency can be dropped for convenience to write $\tilde{E}_{a_m}^p(\nu) \equiv \tilde{E}_{a_m}^p$. Then, such electric fields from individual station elements can be phase-coherently superposed to obtain the electric field in polarisation state, α , towards any desired direction as

$$\tilde{\mathcal{E}}_m^\alpha(\hat{\mathbf{s}}_k) = \sum_{a_m} \sum_p \tilde{\mathcal{W}}_{a_m}^{\alpha p*}(\hat{\mathbf{s}}_k) \tilde{E}_{a_m}^p e^{i\frac{2\pi}{\lambda} \hat{\mathbf{s}}_k \cdot \mathbf{r}_{a_m}}, \quad (1)$$

where, $\hat{\mathbf{s}}_k$ denotes the direction of the superposed beam, k , and $\tilde{\mathcal{W}}_{a_m}^{\alpha p*}(\hat{\mathbf{s}}_k)$ denotes a complex-valued directional weighting^c with the superscript indices α, p , representing the contribution of polarisation state, p in the station element, a_m , to polarisation state, α , towards direction, $\hat{\mathbf{s}}_k$. The $*$ denotes complex conjugate. Equation (1) resembles a DFT of the calibrated Electric field measurement after applying a complex-valued weighting.

The polarised intensity in the beamformed pixel is then obtained by

$$\tilde{\mathcal{I}}_m^{\alpha\beta}(\hat{\mathbf{s}}_k) = \langle \tilde{\mathcal{E}}_m^\alpha(\hat{\mathbf{s}}_k) \tilde{\mathcal{E}}_m^{\beta*}(\hat{\mathbf{s}}_k) \rangle, \quad (2)$$

where, angular brackets denote a temporal averaging across an interval of t_{acc} . $\tilde{\mathcal{I}}_m^{\alpha\beta}(\hat{\mathbf{s}}_k)$ results from the outer product of indices α and β . Superscript, $\alpha\beta$, thus denotes the four pairwise combinations of polarisation states of the intensity. Even though it is an outer product over the polarisation indices, it will be referred to as a ‘squaring’ operation, hereafter, for convenience because of what it reduces to if $\alpha = \beta$.

The solid angle of the beamformed pixel using the intra-station data is given by $\Omega_s \simeq (\lambda/D_s)^2$. Beamforming can be applied to all independent beams (n_{bs}) filling the field of view with solid angle, $\Omega_e \simeq (\lambda/D_e)^2$. Thus, $n_{\text{bs}} \simeq \Omega_e/\Omega_s = (D_s/D_e)^2$. This beamforming step has to be executed at a cadence of δt .

Fig. 3 shows a breakdown of the computational cost per station per voxel for voltage beamforming as a function of the different station parameters. In each panel, all parameters except the one on the x -axis are kept fixed at the characteristic values of the LAMBDA-I, SKA-1low, and SKA-1low-core stations as highlighted in cyan. The total cost is dominated by the DFT beamforming in Equation (1) denoted by black dashed lines. The DFT beamforming cost per voxel scales linearly with the number of elements per station. The costs of squaring and temporal averaging in Equation (2) are subdominant. Because the squaring and temporal averaging are performed on a per-pixel basis and not dependent on N_{eps} , they remain constant across N_{eps} . And because the number of independent pixels covering the field-of-view scales as $n_{\text{bs}} \simeq (D_s/D_e)^2$, the cost per independent pixel remains constant across D_s and D_e . All these operations have to be performed at every time step, δt , and hence remain constant regardless of t_{acc} . Although

^cOne can choose $\tilde{\mathcal{W}}_{a_m}^{\alpha\beta}(\hat{\mathbf{s}}_k) = \mathcal{W}_{a_m}^{\alpha\beta}(\hat{\mathbf{s}}_k)$, the directional electric field sensitivity, if the signal-to-noise ratio is to be optimised. But it can also be generically chosen depending on the property desired of the estimator (Morales 2011).

Table 2. Computational budget of intra- and inter-station coherent imaging architectures.

Architecture	Components	Intra-station FLOP count (per station per voxel)	Inter-station FLOP count (per voxel)	Timescale	Equations
BF	Beamforming	$8n_p^2 N_{\text{eps}}$	$8n_p^2 N_s$	δt	1, 8
	Squaring	$6n_p^2$	$6n_p^2$	δt	2, 9
	Accumulation	$2n_p^2$	$2n_p^2$	δt	2, 9
EPIC	Gridding ^a	$\frac{6n_p^2 N_{\text{eps}} N_{\text{ke}}}{(D_s/D_e)^2}$	$\frac{6n_p^2 N_s N_{\text{ks}}}{(D_A/D_s)^2}$	δt	3, 10
	FFT ^{b c}	$8n_p \frac{r}{\log_2 r} \gamma_s^2 \log_2 \left(\gamma_s \frac{D_s}{D_e} \right)^2$	$8n_p \frac{r}{\log_2 r} \gamma_A^2 \log_2 \left(\gamma_A \frac{D_A}{D_s} \right)^2$	δt	3, 10
	Squaring	$6n_p^2 \gamma_s^2$	$6n_p^2 \gamma_A^2$	δt	2, 9
	Accumulation	$2n_p^2 \gamma_s^2$	$2n_p^2 \gamma_A^2$	δt	2, 9
XBF	Correlator	$\frac{6n_p^2 N_{\text{eps}} (N_{\text{eps}} - 1)/2}{(D_s/D_e)^2}$	$\frac{6n_p^2 N_s (N_s - 1)/2}{(D_A/D_s)^2}$	δt	4, 11
	Accumulation	$\frac{2n_p^2 N_{\text{eps}} (N_{\text{eps}} - 1)/2}{(D_s/D_e)^2}$	$\frac{2n_p^2 N_s (N_s - 1)/2}{(D_A/D_s)^2}$	δt	4, 11
	DFT	$8n_p^2 N_{\text{eps}} (N_{\text{eps}} - 1)/2$	$8n_p^2 N_s (N_s - 1)/2$	t_{acc}	5, 12
XFFT	Correlator	$\frac{6n_p^2 N_{\text{eps}} (N_{\text{eps}} - 1)/2}{(D_s/D_e)^2}$	$\frac{6n_p^2 N_s (N_s - 1)/2}{(D_A/D_s)^2}$	δt	4, 11
	Accumulation	$\frac{2n_p^2 N_{\text{eps}} (N_{\text{eps}} - 1)/2}{(D_s/D_e)^2}$	$\frac{2n_p^2 N_s (N_s - 1)/2}{(D_A/D_s)^2}$	δt	4, 11
	Gridding ^a	$\frac{8n_p^4 (2^2 N_{\text{ke}}) N_{\text{eps}} (N_{\text{eps}} - 1)/2}{(D_s/D_e)^2}$	$\frac{8n_p^4 (2^2 N_{\text{ks}}) N_s (N_s - 1)/2}{(D_A/D_s)^2}$	t_{acc}	6, 13
	FFT ^b	$8n_p^2 \frac{r}{\log_2 r} \log_2 \left(\frac{D_s}{D_e} \right)^2$	$8n_p^2 \frac{r}{\log_2 r} \log_2 \left(\frac{D_A}{D_s} \right)^2$	t_{acc}	6, 13

^a N_{ke} and N_{ks} denote the sizes (in number of cells) of the gridding kernels of the antenna elements and the stations, respectively. In this paper, $N_{\text{ke}} = 1^2$ and $N_{\text{ks}} = 1^2$. For XFFT, compared to pre-correlated data used in EPIC, the gridding sizes for intra- and inter-station correlations are quadrupled to $2^2 N_{\text{ke}}$ and $2^2 N_{\text{ks}}$, respectively, to account for the expansion of the correlated kernels and to remove the aliasing effects within the respective fields of view. If the gridding kernels incorporate w -projection, the kernel sizes will be correspondingly larger.

^b r denotes the radix of the FFT algorithm (Cooley & Tukey 1965). Here, $r = 2$.

^c γ_s and γ_A denote padding factors for station- and array-level FFT, respectively, in EPIC. Here, $\gamma_s = 2$ and $\gamma_A = 2$ along each dimension to achieve identical image and pixel sizes as that from XFFT.

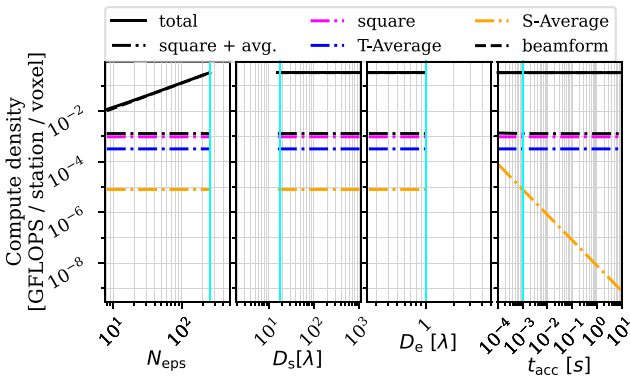


Figure 3. A breakdown of the computational cost density function over the station parameters for the voltage beamforming (BF) architecture at the station level. Each panel shows the variation with the respective parameter keeping the rest fixed at the characteristic values of the LAMBDA-I, SKA-low-core, and SKA-low stations (cyan lines). The beamforming cost in Equation (1) denoted by black dashed lines dominates the squaring (pink dot-dashes) and temporal averaging (blue dot-dashes) costs in Equation (2). Orange dot-dashes denote the incoherent spatial average of the intensities across stations and has negligible contribution to the total cost.

spatial averaging across multiple stations (orange dot-dashed line) is not a station-level operation, it is shown to emphasise that it constitutes only a negligible cost relative to the other components, noting that it needs to be only performed on a slower cadence of t_{acc} on a per-pixel basis, and thus scales inversely with t_{acc} .

Fig. 4 shows a covariant view of the total computational cost density per station per voxel (after summing the DFT, squaring, and averaging components) for the beamforming architecture taken two station parameters at a time while keeping the rest fixed at the nominal values of a LAMBDA-I, SKA-low, or a SKA-low-core station (grey dashed lines). The variation in the cost per voxel is dominated by the beamforming cost and exhibits the same trends observed in Fig. 3 – linear in N_{eps} while remaining constant with D_e and D_s .

5.2. E-field Parallel Imaging Correlator (EPIC)

Instead of DFT beamforming at discrete arbitrary locations as described above, forming independent beams that simultaneously fill the entire field of view can be achieved by a two-dimensional spatial FFT of the calibrated electric field measurements on the aperture plane. Typically, this has been associated with a regularly gridded array layout (Daishido et al. 1991; Otake et al. 1994; Tegmark & Zaldarriaga 2009; Tegmark & Zaldarriaga 2010; Foster et al. 2014; Masui et al. 2019). However, it can be enabled even for an irregular layout of station elements by gridding the calibrated electric fields measured by the station elements. The Optimal Map Making (OMM; Tegmark 1997) formalism adopted by Morales (2011) is the basis of the architecture called E-field Parallel Imaging Correlator (EPIC; Thyagarajan et al. 2017),

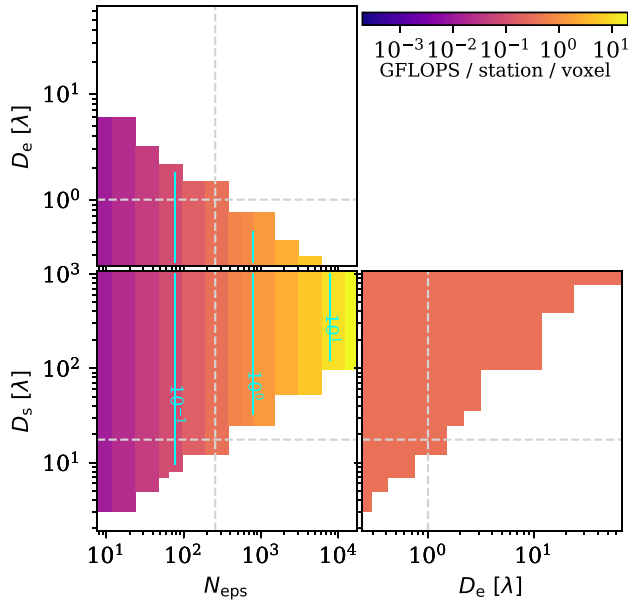


Figure 4. A multi-dimensional covariant view of the net computational cost density function over the station parameters for the voltage beamforming (BF) architecture at the station level by varying two parameters at a time while fixing the rest at the nominal values of the LAMBDA-I, SKA-low-core, and SKA-low stations shown in grey dashed lines. The empty portions in the panels denote parts of parameter space physically impossible to sample given the constraints of the station parameters. Contours corresponding to the colour scale are shown in cyan.

$$\tilde{\mathcal{E}}_m^\alpha(\hat{\mathbf{s}}_k) = \sum_j \delta^2 \mathbf{r}_j e^{i \frac{2\pi}{\lambda} \hat{\mathbf{s}}_k \cdot \mathbf{r}_j} \left(\sum_{a_m} \sum_p \tilde{W}_{a_m}^{\alpha p*}(\mathbf{r}_j - \mathbf{r}_{a_m}) \tilde{E}_{a_m}^p \right). \quad (3)$$

The expression within the parenthesis represents the gridding operation of each calibrated electric field measurement at an arbitrary location, \mathbf{r}_{a_m} , of element a_m onto a common grid at locations, \mathbf{r}_j , using a gridding kernel, $\tilde{W}_{a_m}^{\alpha p*}(\mathbf{r})$ corresponding to that element. The outer summation denotes the Fourier transform of the weighted and gridded electric fields implemented through FFT.

Application of the FFT will have the effect of simultaneously beamforming over the entire field of view, Ω_e . The convolution with $\tilde{W}_{a_m}^{\alpha p*}(\mathbf{r})$ in the aperture, which is the Fourier dual of $\tilde{\mathcal{V}}_{a_m}^{\alpha p*}(\hat{\mathbf{s}}_k)$ in Equation (1), has several purposes. Firstly, it can be chosen to optimise specific properties in the synthesised image. For example, choosing $\tilde{W}_{a_m}^{\alpha p*}(\mathbf{r}) = W_{a_m}^{\alpha p*}(\mathbf{r})$, the complex conjugate of the holographic illumination pattern of the antenna element will optimise the signal-to-noise ratio in the image (Morales 2011). Secondly, it can incorporate w -projection for non-coplanar element spacings (Cornwell, Golap, & Bhatnagar 2008), direction-dependent ionospheric effects, wide-field distortions from refractive and scintillating atmospheric distortions, etc. (Morales & Matejek 2009; Morales 2011). Finally, through gridding convolution it transforms data from discrete and arbitrary locations onto a regular grid, thereby enabling the application of a two-dimensional spatial FFT. Finally, the polarised intensities are obtained by application of Equation (2), that is pixelwise ‘squaring’ (outer product of polarisation states) and temporal averaging.

Fig. 5 shows a breakdown of the costs per station per voxel of the EPIC architecture, with nominal values of a LAMBDA-I, SKA-low, or a SKA-low-core station marked in cyan. The dominant cost is from the two-dimensional FFT (black dashed lines)

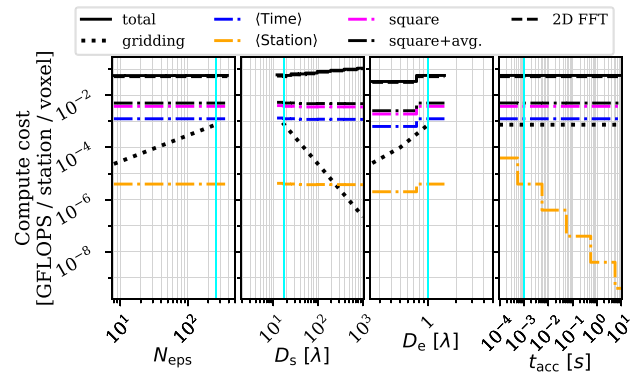


Figure 5. A breakdown of the computational cost density function over the station parameters for the EPIC architecture at the station level. Each panel shows the variation with the respective parameter keeping the rest fixed at the characteristic values of the LAMBDA-I, SKA-low-core, and SKA-low stations (cyan lines). The spatial FFT cost in Equation (3) denoted by black dashed lines dominates the squaring (pink dot-dashes) and temporal averaging (blue dot-dashes) costs in Equation (2). The gridding operation (black dotted lines) in Equation (3) contributes even lesser to the cost budget. The incoherent spatial averaging of the intensities across stations (orange dot-dashed line) has negligible contribution to the total cost.

in Equation (3). The spatial FFT cost depends only on the grid dimensions. So, it is independent of N_{eps} as the grid can hold as many elements as physically possible. The computational cost of the spatial FFT scales roughly as $(\gamma_s D_s / D_e)^2 \log_2 (\gamma_s D_s / D_e)^2$, where, γ_s is a constant padding factor. γ_s is used to control the pixel scale in the image. For example, $\gamma_s = 2$ (adopted in this paper) along each dimension will provide identical pixel scale and image size as that formed from gridded visibilities. Notably, because the spatial FFT produces an image over the entire field of view at every independent pixel instantly, it is not a pixelwise operation, and therefore, its cost per voxel scales as $\sim \log_2 (D_s / D_e)$, which exhibits a weak dependence on D_s and D_e . The squaring (pink dot-dashed lines) and temporal averaging (blue dot-dashed lines) in Equation (2) operate on a per-pixel basis. Thus, those costs per voxel remain constant with N_{eps} , D_s and D_e similar to the voltage beamforming architecture. And they are subdominant. The gridding (black dotted lines) with weights in Equation (3) is a per-element operation and depends only on N_{eps} and the size of the gridding kernel, N_{ke} . Here, $N_{\text{ke}} = 1^2$ is chosen to provide unaliased full field-of-view images. The incorporation of w -projection in the gridding kernel will correspondingly increase N_{ke} . Nevertheless, the gridding cost, scales linearly with N_{eps} and inverse squared with D_s per voxel. Overall, the gridding cost is insignificant. Again, the spatial averaging of pixels across multiple stations (orange dot-dashed line) is not a station-level operation, but is only shown to emphasise that it operates on a much slower cadence of t_{acc} , scaling inversely with t_{acc} , and adds only negligible cost to the overall cost budget.

Fig. 6a shows a covariant view of the total computational cost density per station per voxel (sum of gridding, FFT, squaring and temporal averaging components) for the EPIC architecture taken two station parameters at a time while keeping the rest fixed at the nominal values of a LAMBDA-I, SKA-low, or a SKA-low-core station denoted by grey dashed lines. The variation in the cost per voxel is dominated by the spatial FFT cost which is relatively constant in all these parameters and displays the same trends observed in Fig. 5, namely, nearly constant in N_{eps} , D_e , and D_s . Fig. 6b shows the ratio of the computational cost of the EPIC architecture relative to the voltage beamforming architecture in Fig. 4. It is clearly

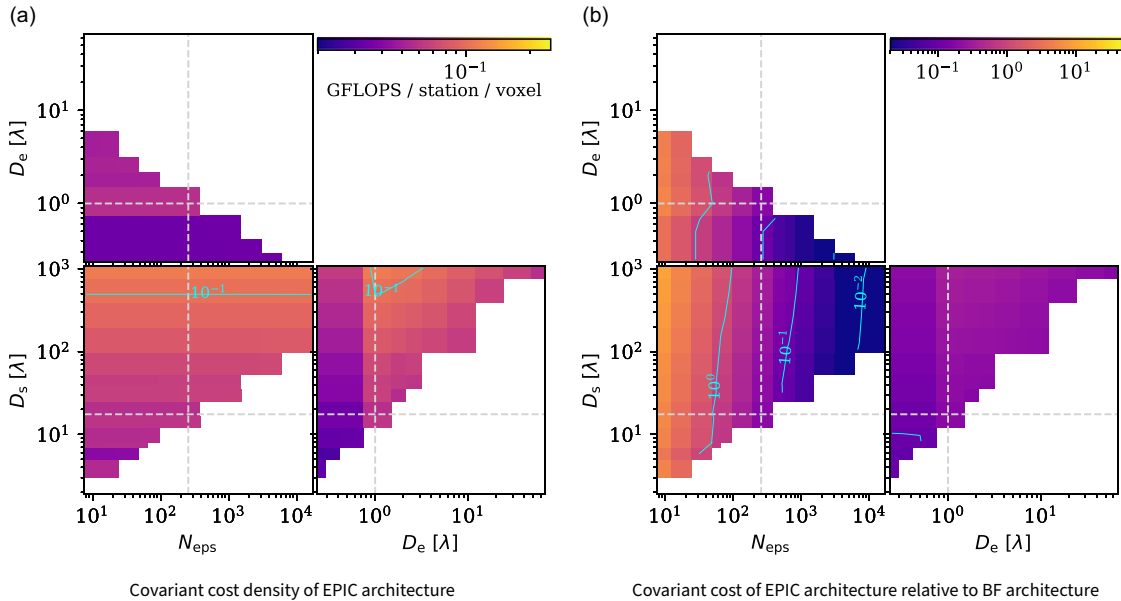


Figure 6. (Left): Two-dimensional slices of the computational cost per station per voxel for imaging using station-level EPIC. Because the spatial FFT dominates the overall cost budget, the computational cost density is relatively insensitive to N_{eps} , D_s , and D_e as seen in Fig. 5, thereby not showing much variation in the colour scale. (Right): Ratio of computational cost of EPIC to voltage beamforming. For $N_{\text{eps}} \gtrsim 100$, EPIC has a significant advantage over voltage beamforming (BF). Grey dashed lines indicate nominal values for LAMBDA-I, SKA-1ow-core, and SKA-1ow stations. Logarithmic contour levels corresponding to the colour scale are shown in cyan.

noted that the cost of the EPIC architecture is significantly less than that of voltage beamforming for $N_{\text{eps}} \gtrsim 100$, and continues to decrease with increasing N_{eps} , implying that for full field-of-view imaging, EPIC has a clear and significant computational advantage over voltage DFT beamforming as N_{eps} and packing density of elements in a station increase.

The EPIC architecture has been implemented on the Long Wavelength Array (LWA) station in Sevilleta (NM, USA) (Kent et al. 2019, 2020; Krishnan et al. 2023) using a GPU framework and operates on a commensal mode. With recent optimisations, the EPIC pipeline on LWA-Sevilleta computes 128×128 all-sky (visible hemisphere) images at 25 000 frames per second and outputs at a cadence of 40–80 ms accumulations (Reddy et al. 2024). For reference, the software correlator and visibility-based imager can produce images at 5 s cadence (Taylor et al. 2012).

The EPIC architecture will be relevant for large cosmological arrays like PUMA that are considering a FFT-based correlator architecture. Obtaining visibilities from an EPIC architecture will involve an additional step of an inverse-FFT or an inverse-DFT back to the aperture plane from the accumulated images (not included in Fig. 5). However, the inverse-FFT or DFT operation to obtain visibilities needs to be performed only once every accumulation interval and is thus not expected to add significantly to the total cost budget.

5.3. Correlator Beamforming (XBF)

In the traditional correlator approach, a pair of electric field measurements, $E_{a_m}^p$ and $E_{b_m}^q$, in polarisation states, p and q , at station elements, a_m and b_m , respectively, in station, m , can be cross-correlated and temporally averaged to obtain calibrated visibilities, $\tilde{V}_{a_m b_m}^{pq}$. It can be written as

$$\tilde{V}_{a_m b_m}^{pq} = \langle E_{a_m}^p E_{b_m}^{q*} \rangle. \quad (4)$$

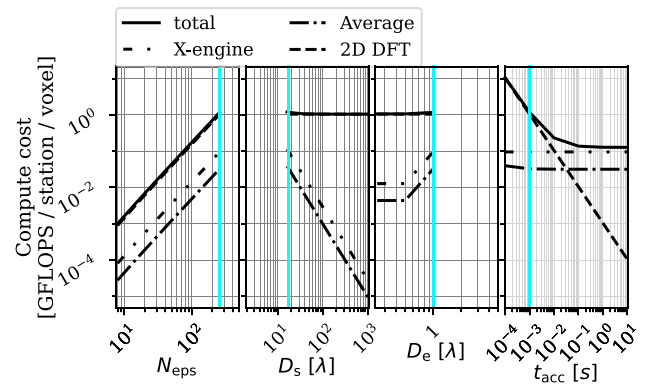


Figure 7. A breakdown of the computational cost density function over the station parameters for the correlator beamforming architecture (XBF) at the station level. Each panel shows the variation with the respective parameter keeping the rest fixed at the characteristic values of the LAMBDA-I, SKA-1ow-core, and SKA-1ow stations (cyan lines). The two-dimensional DFT (dashed lines) dominates the cost for the chosen parameters at an imaging cadence of $t_{\text{acc}} = 1$ ms over correlation/X-engine (double dot-dashed) and temporal averaging of the correlations (dot-dashed).

The visibilities in the station can be beamformed using a DFT to obtain polarised intensities towards \hat{s}_k as:

$$\tilde{\mathcal{I}}_m^{\alpha\beta}(\hat{s}_k) = \sum_{a_m, b_m} \sum_{p, q} \tilde{\mathcal{B}}_{a_m b_m}^{\alpha\beta; pq*}(\hat{s}_k) \tilde{V}_{a_m b_m}^{pq} e^{i \frac{2\pi}{\lambda} \hat{s}_k \cdot \Delta \mathbf{r}_{a_m b_m}}. \quad (5)$$

$\tilde{\mathcal{B}}_{a_m b_m}^{\alpha\beta; pq*}(\hat{s}_k)$, is the counterpart of $\tilde{\mathcal{W}}_{a_m}^{p\alpha*}(\hat{s}_k)$ in Equation (1) for a correlated quantity, and it denotes a complex-valued directional weighting applied to the calibrated visibility representing the contribution of polarisation state, pq , to a state, $\alpha\beta$, in the measurement and sky planes, respectively. It can be chosen to produce images with desired characteristics (Masui et al. 2019).

Fig. 7 shows the computational cost density breakdown for correlator beamforming at the station level. The correlation and

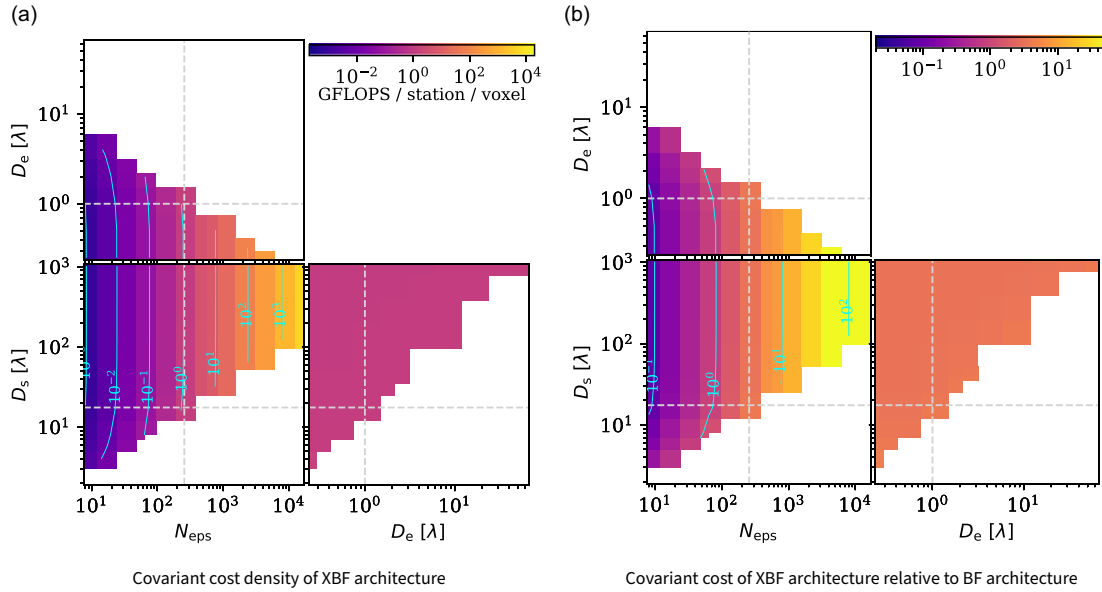


Figure 8. (Left): Two-dimensional slices of the computational cost for imaging using a DFT beamforming of station-level cross-correlations (XBF) for LAMBDA-I, SKA-low-core, and SKA-low station parameters (dashed grey lines). Cyan lines denote contours of the colour scale in logarithmic increments. (Right): Same as the left but relative to a voltage beamformer (BF) architecture. The computational cost density is dominated by the two-dimensional DFT cost for $t_{\text{acc}} = 1$ ms and is lower than that of voltage beamforming (BF) for $N_{\text{eps}} \lesssim 100$ while getting more expensive with increasing N_{eps} .

temporal averaging in Equation (4) are performed on every pair of station elements and scale as $N_{\text{eps}}(N_{\text{eps}} - 1)/2$, with no dependence on D_e or D_s . Hence, their costs per voxel scale as $(D_s/D_e)^{-2}$. The two-dimensional DFT in Equation (5) also scales as $N_{\text{eps}}(N_{\text{eps}} - 1)/2$, but being a per-pixel operation, stays constant with D_e and D_s . The correlation and temporal averaging have to be performed at every interval of δt and are therefore, independent of t_{acc} , whereas the DFT needs to be performed at a cadence of t_{acc} , and thus scales as t_{acc}^{-1} . For the chosen parameters of LAMBDA-I, SKA-low, and SKA-low-core, at $t_{\text{acc}} = 1$ ms, the computational cost density is dominated by the cost of the DFT. However, for imaging on a slower cadence, $t_{\text{acc}} \gtrsim 10$ ms, the cost of correlation begins to dominate.

Fig. 8a is a covariant view of the total computational cost density per station per voxel (sum of correlator, temporal averaging, and DFT components) for the correlator beamformer architecture taken two station parameters at a time while keeping the rest fixed at the nominal values of a SKA-low, SKA-low-core, or a LAMBDA-I station (grey dashed lines). The variation in the cost per voxel for $t_{\text{acc}} = 1$ ms is dominated by the spatial DFT cost which is relatively constant in D_e and D_s , and scales as $\sim N_{\text{eps}}^2$, displaying the same trends observed in Fig. 7. Fig. 8b shows the ratio of the computational cost of the XBF architecture relative to the voltage beamforming architecture in Fig. 4. It is clear that the cost of the XBF architecture is less expensive only for $N_{\text{eps}} \lesssim 100$ and gets worse with increasing N_{eps} .

5.4. Correlation and FFT (XFFT)

In this approach, $\tilde{V}_{a_m b_m}^{pq}$, the calibrated and inverse noise covariance weighted visibilities in station, m , obtained after the correlator in Equation (4) are not weighted and beamformed using a DFT. Instead, they are gridded with weights, $\tilde{B}_{a_m b_m}^{\alpha\beta; pq*}(\Delta\mathbf{r})$, which is the Fourier dual of $\tilde{B}_{a_m b_m}^{\alpha\beta; pq*}(\hat{\mathbf{s}}_k)$ in Equation (5), onto a grid in the

aperture plane. The gridded visibilities are Fourier transformed using an FFT rather than a DFT used in correlator beamforming,

$$\tilde{\mathcal{I}}_m^{\alpha\beta}(\hat{\mathbf{s}}_k) = \sum_j \delta^2 \Delta\mathbf{r}_j e^{i\frac{2\pi}{\lambda} \hat{\mathbf{s}}_k \cdot \Delta\mathbf{r}_j} \times \left(\sum_{a_m, b_m} \sum_p \tilde{B}_{a_m b_m}^{\alpha\beta; pq*}(\Delta\mathbf{r}_j - \Delta\mathbf{r}_{a_m b_m}) \tilde{V}_{a_m b_m}^{pq} \right). \quad (6)$$

The gridding and Fourier transform (using FFT) are represented by the parenthesis and outer summation, respectively. $\tilde{B}_{a_m b_m}^{\alpha\beta; pq*}(\Delta\mathbf{r})$ denotes the polarimetric gridding kernel that is used to place the visibility data in polarisation state, pq , centred at location, $\Delta\mathbf{r}_{a_m b_m}$, onto a grid of locations, $\Delta\mathbf{r}_j$ in polarisation state, $\alpha\beta$. The purpose of $\tilde{B}_{a_m b_m}^{\alpha\beta; pq*}(\Delta\mathbf{r})$ is similar (Morales & Matejek 2009; Bhatnagar et al. 2008) to its counterpart gridding operator in the EPIC architecture (Morales 2011) and can also incorporate the w -projection kernel (Cornwell et al. 2008).

This is the mathematical counterpart to the operations performed on calibrated electric fields from the elements in EPIC, with few key differences – (1) the calibrated visibilities can be accumulated and allowed for Earth rotation and bandwidth synthesis, (2) the gridding and spatial Fourier transform operate not on electric fields but visibilities, and (3) the spatial Fourier transform can occur on a slower cadence of t_{acc} . In the radio interferometric context, $\tilde{\mathcal{I}}_m^{\alpha\beta}(\hat{\mathbf{s}}_k)$ is referred to as the ‘dirty image’ (Thompson, Moran, & Swenson 2017; Taylor, Carilli, & Perley 1999).

Fig. 9 shows the computational cost density breakdown for XFFT at the station level. The correlation and temporal averaging in Equation (4) are the same as the correlator beamformer, scaling as $N_{\text{eps}}(N_{\text{eps}} - 1)/2$ and have no dependence on D_e or D_s . Hence, their costs per voxel scale as $(D_s/D_e)^{-2}$. The gridding in Equation (9) also scales as $N_{\text{eps}}(N_{\text{eps}} - 1)/2$ and the size

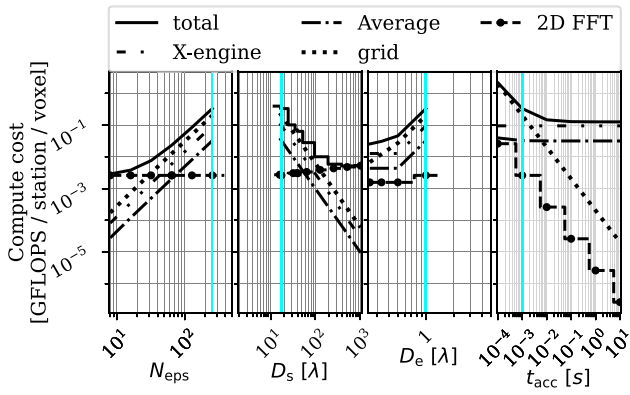


Figure 9. A breakdown of the computational cost density function over the station parameters for the correlator FFT architecture (XFFT) at the station level. Each panel shows the variation with the respective parameter keeping the rest fixed at the characteristic values of the LAMBDA-I, SKA-low-core, and SKA-low stations (cyan lines). The correlator/X-engine (double dot-dashed lines) and gridding (dotted lines) costs are comparable and dominate over the temporal averaging (dot dashed lines) and two-dimensional FFT (dashed lines) costs for the chosen LAMBDA-I parameters at an imaging cadence of $t_{acc} = 1$ ms.

of the gridding kernel, $2^2 N_{ke}$, where, N_{ke} was the size of the kernel used in EPIC. The quadrupling is because the extent of the visibility gridding kernel is expected to be twice as that of the element gridding kernel along each dimension. Thus, $2^2 N_{ke} = 4$ is chosen to provide unaliased full field-of-view images for the calculations here. The incorporation of w -projection in the gridding kernel will correspondingly increase the kernel size, but being independent of grid dimensions, the gridding cost per voxel scales as $(D_s/D_e)^{-2}$. The two-dimensional FFT in Equation (9) depends only on the grid size and is independent of N_{eps} . The FFT scales as $(D_s/D_e)^2 \log_2 (D_s/D_e)^2$, and thus its cost per voxel scales only weakly as $\log_2 (D_s/D_e)$. The correlation and temporal averaging have to be performed at every interval of δt and are therefore, independent of t_{acc} , whereas the gridding and FFT need to be performed at a cadence of t_{acc} , and thus scale as t_{acc}^{-1} . For the chosen parameters of LAMBDA-I, SKA-low, and SKA-low-core, and $t_{acc} = 1$ ms, gridding dominates the overall cost, followed by the correlator, temporal averaging of correlations, and FFT. However, for imaging on a slower cadence, $t_{acc} \gtrsim 5-10$ ms, the cost of correlation begins to dominate.

Fig. 10a shows a covariant view of the total computational cost density per station per voxel (sum of correlator, temporal averaging, gridding, and FFT components) for the correlator FFT architecture taken two station parameters at a time while keeping the rest fixed at the nominal values of a LAMBDA-I, SKA-low, or a SKA-low-core station denoted by grey dashed lines. The variation in the cost per voxel for $t_{acc} = 1$ ms is dominated by the gridding cost, displaying the same trends observed in Fig. 9. Fig. 10b shows the ratio of the computational cost of the XFFT architecture relative to the voltage beamforming architecture in Fig. 4. Clearly, the cost of the XFFT architecture is less expensive than the BF in most of the parameter space, particularly for larger D_s/D_e . For the chosen parameters (grey dashed lines), the XFFT cost roughly matches that of BF.

For applications where visibilities are processed offline and real-time imaging is not required, like traditional cosmology experiments including the EoR, only the computational cost of correlations in forming the visibilities will apply, which is still one

of the dominant components of the overall cost in Figs. 7 and 9. The DFT, gridding, and FFT costs can be ignored for real-time processing.

6. Inter-station imaging architectures

The products from the intra-station architectures form the inputs for inter-station imaging. Here, the inter-station processing is also assumed to occur in real time.

6.1. Incoherent architectures

Incoherent inter-station imaging simply involves accumulating the images from each station at a cadence of t_{acc} . The angular resolution remains the same as the intra-station images (assuming all stations have equal angular resolution) obtained through any of the four approaches described in Section 5. The weighted averaging of intra-station intensities gives

$$\tilde{\mathcal{I}}^{\alpha\beta}(\hat{s}_k) = \frac{\sum_m w_m^{\alpha\beta}(\hat{s}_k) \tilde{\mathcal{I}}_m^{\alpha\beta}(\hat{s}_k)}{\sum_m w_m^{\alpha\beta}(\hat{s}_k)}, \quad (7)$$

where, $w_m^{\alpha\beta}(\hat{s}_k)$ represents station weights for a beam pointed towards \hat{s}_k . The incoherent nature of this inter-station intensity accumulation implies that it does not depend on inter-station spacing or the array diameter, D_A .

6.2. Coherent architectures

A coherent combination of data between stations requires electric field data products with phase information from the first stage of processing on individual stations. Thus, intra-station visibilities in which absolute phases have been removed and only phase differences remain, as well as the image intensity outputs of XBF and XFFT, are unusable for coherent processing at the inter-station level. Hence, the two remaining inputs are considered, namely, the electric fields, $\tilde{\mathcal{E}}_m^{\alpha}(\hat{s}_k)$, from BF and EPIC from Equations (1) and (3), respectively.

The intra-station coherent combinations, $\tilde{\mathcal{E}}_m^{\alpha}(\hat{s}_k)$, produced through BF and EPIC, respectively, in Equations (1) and (3), effectively convert the station, m , to act as virtual station-sized telescopes measuring the electric fields but simultaneously pointed towards multiple locations, denoted by \hat{s}_k . Each of these virtual station-sized telescopes has a field of view determined by inverse of the station size as $\sim (\lambda/D_s)^2$. Together, with all the different virtual pointings, they can cover a field of view given by the inverse of the element size as $\sim (\lambda/D_e)^2$. Using these electric fields as the input measurements from the virtually synthesised telescopes pointed towards any given direction, \hat{s}_k , the inter-station imaging architecture can now employ any of the four architectures that were previously discussed in Section 5, for all possible directions filling the entire field of view.

For inter-station processing, all element- and station-related quantities in intra-station processing will be replaced with station- and array-related quantities, respectively. Similarly, the element-based measurements of electric fields, $\tilde{E}_{a_m}^p$, will be replaced with the synthesised electric fields, $\tilde{\mathcal{E}}_m^{\alpha}(\hat{s}_k)$, measured by the virtual station-sized telescopes obtained either through BF (Equation 1) or EPIC (Equation 3). Since the computational costs of the inter-station processing now depend on the array layout and that of intra-station processing on the station layout, it necessitates hybrid architectures appropriate for multi-scale aperture arrays. Table 2 summarises the computational costs of inter-station processing.

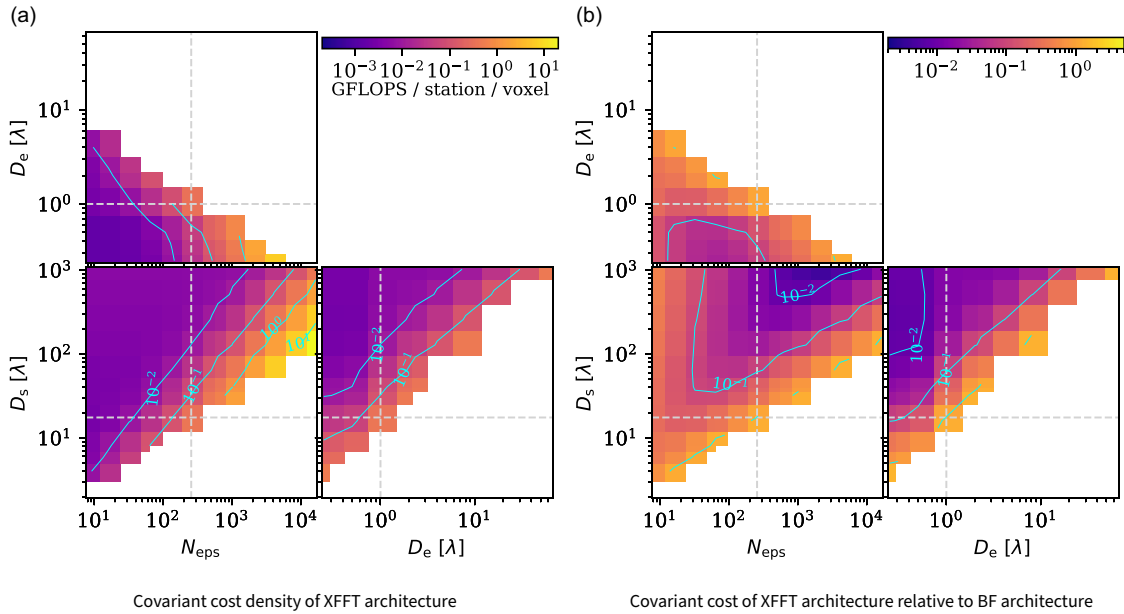


Figure 10. (Left): Two-dimensional slices of the computational cost for imaging using an FFT of station-level cross-correlations (XFFT) for LAMBDA-I, SKA-low-core, and SKA-low station parameters (dashed grey lines). Cyan contours denote logarithmic levels in the colour scale.

6.2.1. Beamforming

Using the electric fields, $\tilde{\mathcal{E}}_m^p(\hat{\mathbf{s}}_k)$, synthesised from each intra-station processing as input, the inter-station DFT beamforming can be written very similar to Equation (1), as

$$\tilde{\mathcal{I}}_{\hat{\mathbf{s}}_k}^\alpha(\boldsymbol{\sigma}_\ell) = \sum_m \sum_p \tilde{\mathcal{W}}_m^{\alpha p*}(\boldsymbol{\sigma}_\ell, \hat{\mathbf{s}}_k) \tilde{\mathcal{E}}_m^p(\hat{\mathbf{s}}_k) e^{i\frac{2\pi}{\lambda} \boldsymbol{\sigma}_\ell \cdot \mathbf{r}_m}. \quad (8)$$

$\boldsymbol{\sigma}_\ell$ denotes locations within the beam solid angle synthesised by the virtual station-sized telescopes pointed towards $\hat{\mathbf{s}}_k$. The directional weighting applied to the synthesised electric fields from station m towards directions, $\boldsymbol{\sigma}_\ell$, centred on the beam directed towards $\hat{\mathbf{s}}_k$ is denoted by $\tilde{\mathcal{W}}_m^{\alpha p*}(\boldsymbol{\sigma}_\ell, \hat{\mathbf{s}}_k)$. Superscripts α and p denote the contribution of polarisation state p from the measured input field to polarisation state α in the synthesised field.

The polarised intensity distribution within the beam solid angle is given by the pixel-wise outer product over polarisation states ('squaring') and averaging, similar to Equation (2),

$$\tilde{\mathcal{I}}_{\hat{\mathbf{s}}_k}^{\alpha\beta}(\boldsymbol{\sigma}_\ell) = \left\langle \tilde{\mathcal{E}}_{\hat{\mathbf{s}}_k}^\alpha(\boldsymbol{\sigma}_\ell) \tilde{\mathcal{E}}_{\hat{\mathbf{s}}_k}^{\beta*}(\boldsymbol{\sigma}_\ell) \right\rangle, \quad (9)$$

where, the angular brackets denote a temporal averaging across an interval of t_{acc} . The angular resolution of the inter-station beamforming will correspond to the dimensions of the station layout within the array as $\sim (\lambda/D_A)^2$.

6.2.2. EPIC

Instead of DFT beamforming with electric fields synthesised from the stations, inter-station EPIC can be used to simultaneously form beams in all independent directions filling the available field of view using FFT. It takes the form

$$\tilde{\mathcal{E}}_{\hat{\mathbf{s}}_k}^\alpha(\boldsymbol{\sigma}_\ell) = \sum_j \delta^2 \mathbf{r}_j e^{i\frac{2\pi}{\lambda} \boldsymbol{\sigma}_\ell \cdot \mathbf{r}_j} \times \left(\sum_m \sum_p \tilde{\mathcal{W}}_m^{\alpha p*}(\mathbf{r}_j - \mathbf{r}_m, \hat{\mathbf{s}}_k) \tilde{\mathcal{E}}_m^p(\hat{\mathbf{s}}_k) \right). \quad (10)$$

The parenthesis represents gridding of the synthesised input electric fields, $\tilde{\mathcal{E}}_m^p(\hat{\mathbf{s}}_k)$ at locations, \mathbf{r}_m , onto the grid locations at \mathbf{r}_j with the gridding kernels, $\tilde{\mathcal{W}}_m^{\alpha p*}(\mathbf{r}, \hat{\mathbf{s}}_k)$, that are applicable for intra-station syntheses towards $\hat{\mathbf{s}}_k$ from stations indexed by m . $\tilde{\mathcal{W}}_m^{\alpha p*}(\mathbf{r}, \hat{\mathbf{s}}_k)$ is the Fourier dual of $\tilde{\mathcal{W}}_m^{\alpha p*}(\boldsymbol{\sigma}_\ell, \hat{\mathbf{s}}_k)$ used in DFT beamforming in Equation (8). The outer summation denotes the two-dimensional FFT of the weighted and gridded electric fields. A gridding kernel size of $N_{\text{ks}} = 1^2$ is assumed. Incorporating any w -terms using w -projection into the gridding will increase N_{ks} . A padding of the grid can be done to control the pixel scale in the image. A constant padding factor of $\gamma_A = 2$ (adopted in this paper) along each dimension will provide identical pixel scale and image size as that formed from gridded visibilities. The polarised intensities are produced by 'squaring' and temporal averaging as in Equation (9). The image within the solid angle of each intra-station beam towards $\hat{\mathbf{s}}_k$ have a solid angle corresponding to the full array layout including all stations as $\sim (\lambda/D_A)^2$.

For arrays that are considering FFT correlators to obtain visibilities, there will be an additional inverse-FFT (not included here) on the polarised intensities. However, the computational cost of this additional step is not expected to be significant compared to the overall cost due to the slower cadence of the operation.

6.2.3. Correlator beamforming

In this correlator-based approach, the station-synthesised electric fields, $\tilde{\mathcal{E}}_m^p(\hat{\mathbf{s}}_k)$, virtually pointed towards $\hat{\mathbf{s}}_k$, are cross-correlated and temporally averaged to produce the inter-station visibilities,

$$\tilde{\mathcal{V}}_{mn}^{pq}(\hat{\mathbf{s}}_k) = \left\langle \tilde{\mathcal{E}}_m^p(\hat{\mathbf{s}}_k) \tilde{\mathcal{E}}_n^{q*}(\hat{\mathbf{s}}_k) \right\rangle. \quad (11)$$

The visibilities phase-centred on \hat{s}_k are then beamformed using DFT similar to Equation (5) to produce polarised intensities at locations, σ_ℓ , within the solid angle of the beam centred on \hat{s}_k as

$$\tilde{I}_{\hat{s}_k}^{\alpha\beta}(\sigma_\ell) = \sum_{m,n} \sum_{p,q} \tilde{B}_{mn}^{\alpha\beta;pq*}(\sigma_\ell, \hat{s}_k) \tilde{V}_{mn}^{pq}(\hat{s}_k) e^{i\frac{2\pi}{\lambda} \hat{\sigma}_\ell \cdot \Delta \mathbf{r}_{mn}}. \quad (12)$$

$\tilde{B}_{mn}^{\alpha\beta;pq*}(\sigma_\ell, \hat{s}_k)$ is a polarimetric directional weighting towards locations, σ_ℓ , within the solid angle of the station-synthesised beam centred on \hat{s}_k , and is applied to the corresponding visibility phase-centred towards \hat{s}_k . Each of the images within the beam solid angle centred on \hat{s}_k will have an angular size that scales with the array size as $\sim (\lambda/D_A)^2$.

6.2.4. Correlation FFT

An alternative to creating images from the visibilities using DFT beamforming is through FFT after gridding. The visibilities obtained in Equation (1) are transformed using an FFT following the same formalism in Equation (6), to yield the polarised intensities simultaneously in all independent pixels filling the field of view,

$$\tilde{I}_{\hat{s}_k}^{\alpha\beta}(\sigma_\ell) = \sum_j \delta^2 \Delta \mathbf{r}_j e^{i\frac{2\pi}{\lambda} \sigma_\ell \cdot \Delta \mathbf{r}_j} \left(\sum_{m,n} \sum_{p,q} \tilde{B}_{mn}^{\alpha\beta;pq*}(\Delta \mathbf{r}_j - \Delta \mathbf{r}_{mn}, \hat{s}_k) \tilde{V}_{mn}^{pq}(\hat{s}_k) \right). \quad (13)$$

The parenthesis represents the gridding operation. The term, $\tilde{B}_{mn}^{\alpha\beta;pq*}(\Delta \mathbf{r}, \hat{s}_k)$, is the Fourier dual of $\tilde{B}_{mn}^{\alpha\beta;pq*}(\sigma_\ell, \hat{s}_k)$ in Equation (12). It acts as the polarimetric gridding kernel that weights and grids the inverse noise covariance weighted inter-station visibilities, $\tilde{V}_{mn}^{pq}(\hat{s}_k)$, at station separations, $\Delta \mathbf{r}_{mn}$, phased towards \hat{s}_k , onto an array-scale grid at locations, $\Delta \mathbf{r}_j$. Any visibilities overlapping on the grid points are thus weighted and averaged. A gridding kernel size of $2^2 N_{ks}$ is assumed, where the quadrupling (doubling in each direction) is relative to the kernel size used in EPIC. In this paper, $2^2 N_{ks} = 4$ is assumed to provide unaliased imaging within the field of view. The kernel size would increase if w -projection is incorporated to account for any w -terms across the array.

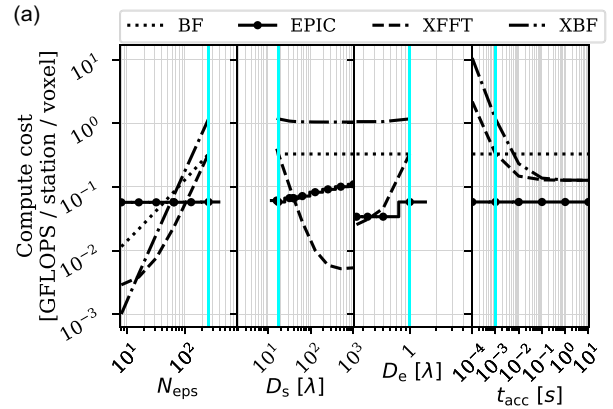
If only visibilities, not images, are required such as in traditional cosmology experiments, then the costs associated with the DFT in Equation (12), and gridding and FFT in Equation (13) can be ignored for real-time processing. However, the cost of correlations in forming the visibilities will continue to dominate the total cost budget for large arrays.

7. Results and discussion

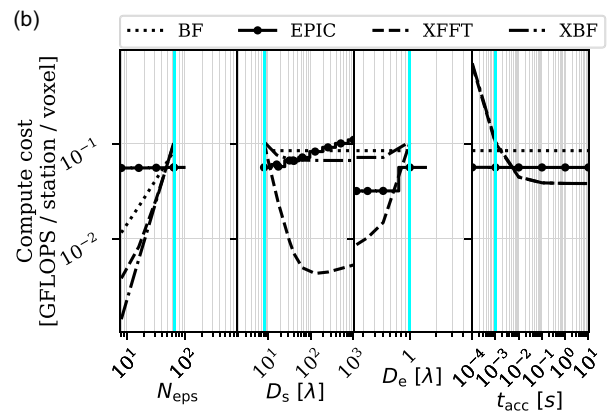
Below is a comparative view of the various architectures from the viewpoint of computational cost density for the range of array layouts and imaging cadence considered in this paper.

7.1. Station-scale imaging architectures

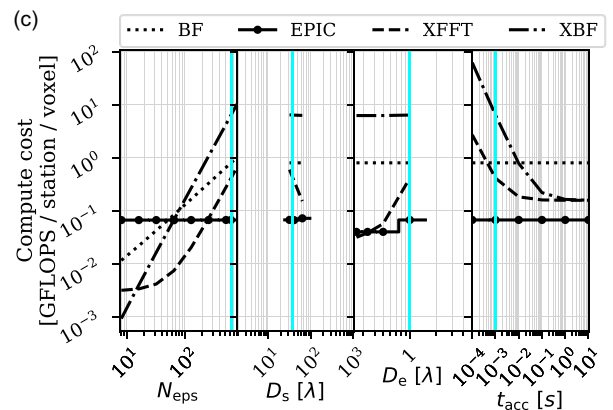
Fig. 11 compares the computational cost densities of different imaging architectures discussed in Section 5 for varying parameters of the arrays in Table 1. The station layout parameters for LAMBDA-I, SKA-low-core, and SKA-low are identical and so are their station-level computing costs. For station parameters of these arrays and that of FarView-core, EPIC is computationally



Station imaging compute cost densities for LAMBDA-I, SKA-low, and SKA-low-core stations.



Station imaging compute cost densities for a CASPA station.



Station imaging compute cost densities for a FarView-core station.

Figure 11. Computational cost densities of different station-level imaging architectures for (a) SKA-low-core, SKA-low, and LAMBDA-I, (b) CASPA, and (c) FarView-core stations. In each subpanel, all parameters are fixed at the array's nominal value (cyan line) except the parameter shown on the x-axis. For SKA-low-core, SKA-low, LAMBDA-I, and FarView-core, the EPIC architecture is found to be most efficient computationally. For CASPA, EPIC is more efficient than other architectures when imaging cadence is $t_{acc} \lesssim 1 - 5$ ms, while XFFT is more efficient for $t_{acc} \gtrsim 5 - 10$ ms.

advantageous clearly. For CASPA, EPIC and XFFT hold the advantage for imaging cadence faster and slower than $t_{acc} \simeq 5$ ms, respectively.

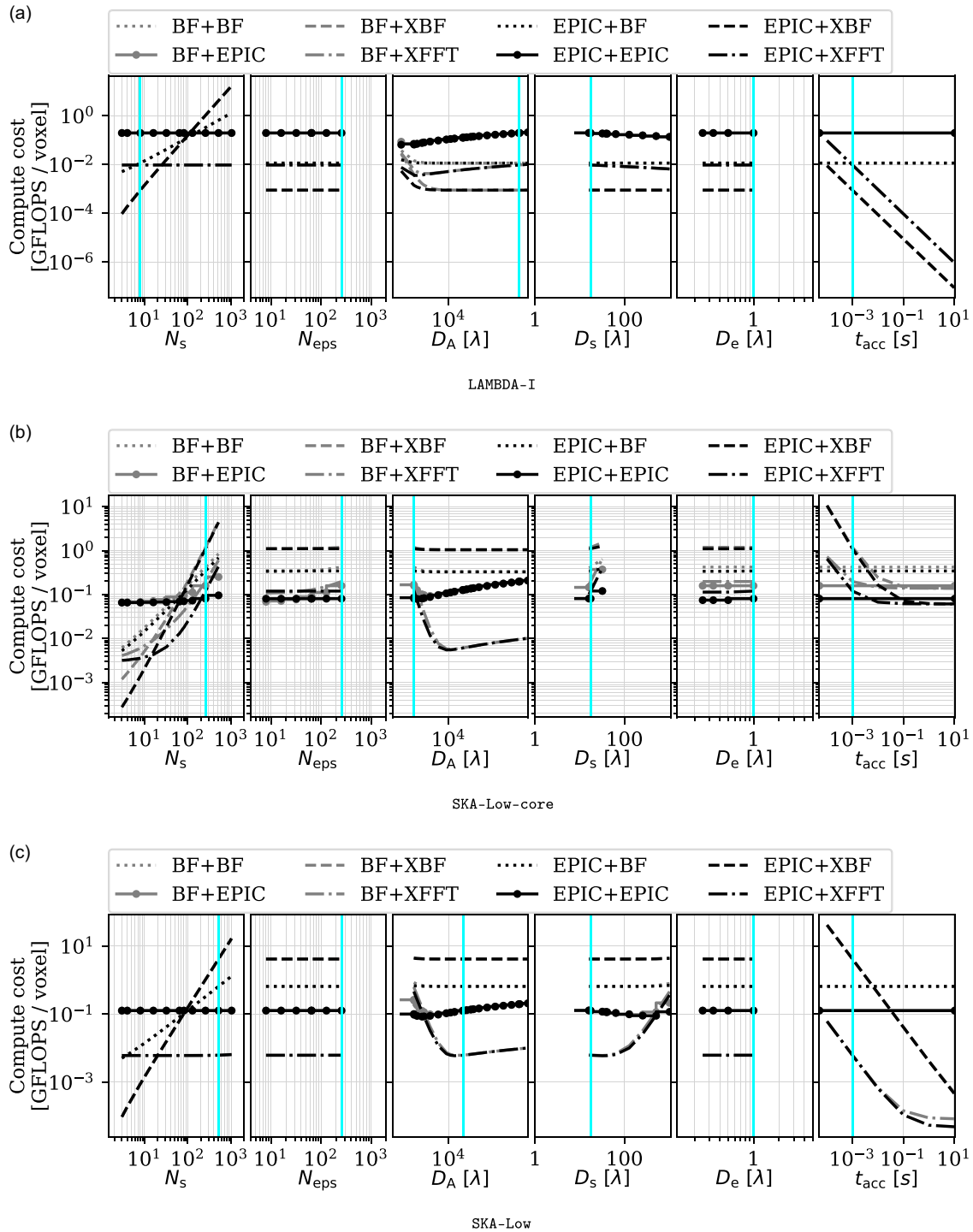


Figure 12. One-dimensional slices of the computational cost density for (a) LAMBDA-I, (b) SKA-low-core, and (c) SKA-low using two-stage coherent imaging at intra- and inter-station levels. Station-synthesised electric fields are obtained through BF (grey) and EPIC (black) in intra-station processing. Inter-station processing of data is obtained by BF (dotted), EPIC (solid), XBF (dashed), and XFFT (dot-dashed). The text in the legend lists the first (intra-station) and second (inter-station) stage architectures delimited by ‘+’.

7.2. Multi-scale imaging architectures

Fig. 12 shows a comparison of the compute cost densities in the discovery phase space of different multi-scale imaging architectures combining four options – BF, EPIC, XBF, and XFFT – on the inter-station data with two options – BF and EPIC – on

the intra-station data for various array parameters and imaging cadences.

In all cases, EPIC (black lines) is more efficient than BF (grey lines) for intra-station full field-of-view data processing (first stage). This is because of the relatively large number of elements densely packed in stations and due to the capability of

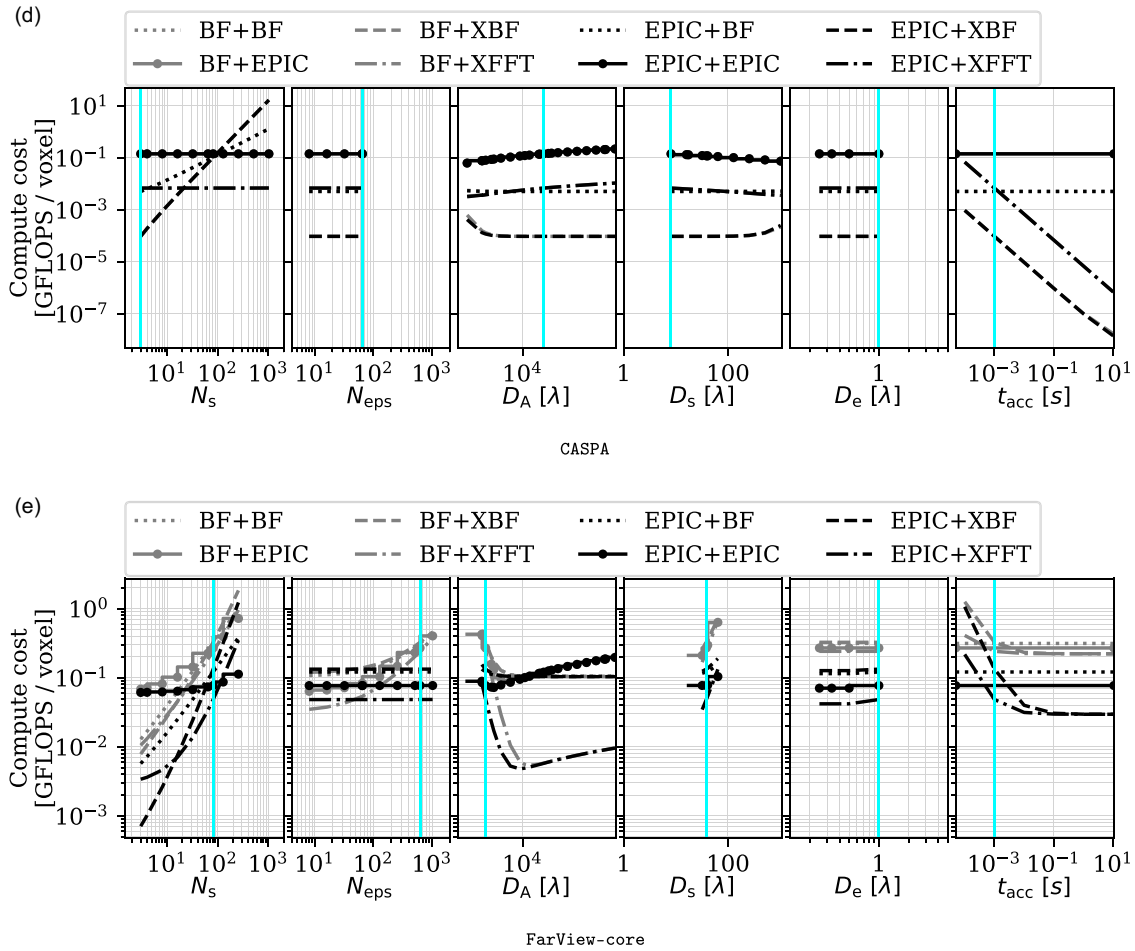


Figure 12. Continued. Same as above but for (d) CASPA, and (e) FarView-core.

simultaneous full field of view beamforming by the EPIC architecture powered by the FFT’s efficiency.

The inter-station processing, however, varies significantly due to the diversity in the various inter-station array parameters and imaging cadence considered. For example,

- LAMBDA-I will benefit from deploying the correlator beamforming (XBF) architecture for inter-station processing by a factor of ≈ 10 over XFFT or BF.
- SKA-low-core will have a marginal advantage from EPIC for the chosen parameters. However, if the number of stations, N_s , decreases, the array diameter, D_A , increases or if the imaging cadence interval is slowed to $t_{acc} \gtrsim 5$ ms, the XFFT architecture starts to have a significant advantage over EPIC on the inter-station scale of data processing.
- SKA-low will have an advantage from the XFFT architecture for inter-station processing by a factor ≈ 20 .
- CASPA will have an advantage by a factor $\gtrsim 20$ from using the XBF architecture for inter-station processing due to significantly smaller number of stations.
- FarView-core appears to have a marginal advantage in using XFFT for inter-station processing for the chosen parameters. If N_s increases or if the cadence interval decreases to $t_{acc} \lesssim 1$ ms, then EPIC becomes more efficient.

7.3. Effect of imaging cadence

The scientific goals enabled by access to the wide range of timescales exhibited by transient phenomena (microseconds to days) are key determinants of whether the architecture deployed can support the processing, as they are intricately tied to different timescales (see Table 2). Here, the impact of the cadence interval on the choice of imaging architecture for multi-scale arrays for intra- and inter-station processing is examined. Table 3 summarises the findings from Figs. 11 and 12 based on the dependence of the computational cost densities of the different architectures and arrays on t_{acc} .

For intra-station processing, EPIC emerges as the most efficient among various architectures for all arrays considered on all timescales by a factor of $\lesssim 10$, except for CASPA when $t_{acc} \gtrsim 100$ ms where the XFFT architecture emerges to be more efficient than EPIC by a factor of few. This is because the stations have a sufficiently large number of elements packed densely (from Table 1). For arrays like CASPA that have lesser number of elements and/or are not as densely packed, the cadence interval becomes an important factor. For cadence intervals, $t_{acc} \gtrsim 100$ ms, the XFFT architecture becomes advantageous. Even for cosmological applications that only require intra-station visibilities at cadence slower than 10 s, an FFT-correlator architecture implemented via EPIC will have the computational advantage over a

Table 3. Impact of imaging cadence on choice of efficient imaging architecture.

Array	Intra-station architecture				Inter-station architecture			
	$\lesssim 0.1$ ms	$\simeq 1$ ms	$\gtrsim 100$ ms	$\gtrsim 10$ s	$\lesssim 0.1$ ms	$\simeq 1$ ms	$\gtrsim 100$ ms	$\gtrsim 10$ s
LAMBDA-I	EPIC	EPIC	EPIC	EPIC	BF/XBF	XBF	XBF	XBF
SKA-low-core	EPIC	EPIC	EPIC	EPIC	EPIC	EPIC	XFFT/EPIC/XBF	XBF/XFFT
SKA-low	EPIC	EPIC	EPIC	EPIC	XFFT/EPIC	XFFT	XFFT	XFFT
CASPA	EPIC	EPIC	XFFT	XFFT	XBF	XBF	XBF	XBF
FarView-core	EPIC	EPIC	EPIC	EPIC	EPIC	XFFT	XBF/XFFT	XBF/XFFT

traditional correlator architecture. However, for CASPA, the latter will have an advantage over the former.

The same reasoning can be extended for inter-station processing, where there is significant diversity. There is a large range in the number of stations and in the array filling factor. As a result, both the layout and cadence interval play significant roles in determining the efficiency of architectures. For arrays with small numbers of stations such as LAMBDA-I and CASPA, the XBF architecture holds an advantage by a factor of $\simeq 10$ on $t_{\text{acc}} \gtrsim 1$ ms and BF for $t_{\text{acc}} \lesssim 0.1$ ms. For arrays with large numbers of stations such as SKA-low-core with a reasonably high packing density, EPIC holds the advantage on all timescales by a factor of few, but with slower cadence intervals, $t_{\text{acc}} \gtrsim 100$ ms, XBF and XFFT also have comparable efficiency. Although the SKA-low has a large number of stations, they are not as densely packed as in the core, and therefore, the role of t_{acc} is more significant. XFFT and EPIC are comparable (within a factor of few) at $t_{\text{acc}} \lesssim 1$ ms, but for $t_{\text{acc}} \gtrsim 1$ ms, XFFT is clearly efficient by a factor $\gtrsim 10$. Similarly, FarView-core has a comparable array filling factor as SKA-low-core, but has far fewer stations. Hence, EPIC is only marginally efficient for $t_{\text{acc}} \lesssim 0.1$ ms, whereas XFFT becomes more efficient by a factor of few for $t_{\text{acc}} \gtrsim 1$ ms. If cosmological applications require inter-station visibilities at a cadence slower than $\gtrsim 100$ ms, a correlator-based architecture holds the computational edge for all the arrays considered here.

7.4. Other considerations

In transient search and cosmological applications, there will be additional processing required. For example, a de-dispersion trial process is required to discover new FRBs and pulsars with optimum signal-to-noise ratio. Computationally efficient algorithms for de-dispersion like the Fast Dispersion Measure Transform (FDMT; Zackay & Ofek 2017) need to be performed either on intensities in the image plane pixels or on visibilities in the aperture plane, the cost of which will scale with the number of independent pixels in the image or the visibilities, respectively. For compact, dense, large- N arrays, the number of independent pixels will be typically lesser than the number of measured visibilities, and thus it will be advantageous to perform de-dispersion in the image plane. Conversely, it will be computationally preferable to de-disperse visibilities for sparse arrays. A full accounting of such costs for most general cases is beyond the scope of this paper.

In intensity mapping applications for cosmology, the images need to be post-processed for deconvolution, foreground, removal, etc. As these post-processing computations are additional and common to all architectures, including them in adopted cost metric will not affect the results. Calibration is another aspect

that is not included in this computational cost analysis. Because calibration typically operates on a significantly slower cadence compared to t_{acc} , and the computational cost is not tedious (Beardsley et al. 2017; Gorthi et al. 2021), its contribution is negligible to the overall cost budget, and therefore will not affect the results.

Besides computational costs, there are other practical considerations in the deployment of an architecture such as the hardware's processing bandwidth, internal memory bandwidth management, input and output data rates, etc. For example, the GPU-based deployment of EPIC on the LWA is severely limited by the memory bandwidth constraints on chip.^d The architecture of current generation of GPUs is found to be inefficient at handling the transpose operation required in a two-dimensional FFT, whereas they may be more efficient at handling correlations. On the contrary, the current generation of FPGAs seem more suited for low bit width two-dimensional FFT from computational and on-chip memory bandwidth perspectives. It is extremely difficult to forecast the continually evolving landscape of hardware capabilities accessible to instruments at the instant of deployment. Thus, this analysis has been restricted to the metric of computational cost density, which is independent of evolving hardware capabilities. For example, even if the computations are parallelised over a large number of GPU cores as noted in Sokolowski et al. (2024), the total computational cost remains unaffected. So, it is intended to be viewed as a first step among many towards an actual implementation.

8. Conclusions

Explosive transient phenomena and observations of cosmological structures are two high impact science drivers in radio astronomy. The paradigm of aperture arrays used in radio astronomy is undergoing a transformation towards large numbers of low-cost elements grouped together in a hierarchy of spatial scales to achieve the desired science outcomes. Such a paradigm faces enormous challenges in computing and throughput because of the large numbers of elements spanning a multiplicity of spatial scales, a range of cadence intervals, real-time processing, and a continuous operation model. Different processing architectures are optimal for specific layouts and processing cadences. However, a single architectural may not be optimal for a hierarchical spatial layout.

This paper explores hybrid architecture solutions for existing and planned multi-scale aperture arrays (SKA-low-core, SKA-low, LAMBDA-I, CASPA, and FarView-core) on a wide

^dSee https://github.com/epic-astronomy/Memos/blob/master/PDFs/009_EPIC_Code_Optimizations.md for a detailed analysis.

range of cadence intervals, $0.1 \text{ ms} \leq t_{\text{acc}} \lesssim 10 \text{ s}$, over the full field of view, using the metric of computational cost density evaluated over the discovery parameter space. The interferometric data processing architectures considered here are pre-correlation voltage beamformer (BF), a generic FFT-based direct imager (EPIC), beamforming from correlations (XBF), and FFT-based imaging of correlations (XFFT). The compute budget for various architecture combinations for these different arrays are broken down into their detailed components.

For densely packed layouts with large numbers of elements, EPIC is computationally quite efficient regardless of processing cadence. This was the case for the station layouts of all the arrays considered, and hence EPIC emerges as the computationally most efficient architecture for full field of view, station-level data processing on all cadence intervals except for CASPA on $t_{\text{acc}} \gtrsim 100 \text{ ms}$, when XFFT becomes more efficient, although marginally. For cosmological applications with large stations where only intra-station visibilities are required at a slower cadence $\gtrsim 10 \text{ s}$ with the arrays studied here, an EPIC-based FFT correlator holds the advantage over a traditional correlator architecture, except in the case of CASPA where the latter will be preferred.

The inter-station array layout has a wider range in the number of stations and density of their distribution. In such cases, the cadence interval plays a significant role in determining the computationally most efficient architecture. For inter-station data processing in LAMBDA-I and CASPA, XBF is found to be the most suitable option for all cadence intervals, while for SKA-low-core it is EPIC that holds the advantage on all cadence intervals. For SKA-low, XFFT appears most efficient for $t_{\text{acc}} \gtrsim 1 \text{ ms}$, while EPIC and XFFT become comparable for $t_{\text{acc}} \lesssim 0.1 \text{ ms}$. If only slower cadence ($\gtrsim 10 \text{ s}$) inter-station visibilities are required from these arrays such as in cosmological applications, a traditional correlator architecture holds a computational advantage. Thus, a multi-scale aperture array requires an optimal hybrid architecture strategy depending on the processing cadence, which in turn depends on the scientific objectives. This study provides a guide for designing computationally strategic data processing architecture for hierarchical, multi-scale aperture arrays, and conversely, can be used to design array parameters for a specified computational capability.

This work considers only computational cost density over the discovery parameter space as a metric for finding the most efficient combination of processing architectures for multi-scale aperture arrays. Additional computing will be required for post-processing steps such as de-dispersion, deconvolution, etc. depending on the application. Actual deployment will have to take other significant practical factors into account such as power consumption, memory bandwidth, data rate, cost, and capabilities of current hardware.

Acknowledgement. Inputs from Ron Ekers, Vivek Gupta, David Humphrey, Nivedita Mahesh, and Rajaram Nityananda are gratefully acknowledged.

Data availability. Not applicable.

References

- Beardsley, A. P., Thyagarajan, N., Bowman, J. D., & Morales, M. F. 2017, *MNRAS*, **470**, 4720
 Bhatnagar, S., Cornwell, T. J., Golap, K., & Uson, J. M. 2008, *A&A*, **487**, 419
 Bochenek, C. D., et al. 2020, *Natur*, **587**, 59

- Braun, R., Bonaldi, A., Bourke, T., Keane, E., & Wagg, J. 2019, arXiv e-prints, arXiv:1912.12699
 Chandra, P., et al. 2016, *JApA*, **37**, 30
 Cooley, J. W., & Tukey, J. W. 1965, *MC*, **19**, 297
 Cordes, J. 2007, in *From Planets to Dark Energy: the Modern Radio Universe*, **35**
 Cornwell, T. J., Golap, K., & Bhatnagar, S. 2008, *IEEE JSTSP*, **2**, 647
 Cosmic Visions 21 cm Collaboration, et al. 2018, arXiv e-prints, arXiv:1810.09572
 Crawford, F., et al. 2022, *MNRAS*, **515**, 3698
 Crichton, D., et al. 2022, *JATIS*, **8**, 011019
 Daishido, T., et al. 1991, in *Astronomical Society of the Pacific Conference Series*, Vol. 19, IAU Colloq. 131: Radio Interferometry. Theory, Techniques, and Applications, ed. T. J. Cornwell, & R. A. Perley, **86**
 DeBoer, D. R., et al. 2017, *PASP*, **129**, 045001
 Dewdney, P. E., Hall, P. J., Schilizzi, R. T., & Lazio, T. J. L. W. 2009, *IEEE Proc.*, **97**, 1482
 Dowell, J., & Taylor, G. B. 2018, *JAI*, **7**, 1850006
 Du, Y. J., Xu, R. X., Qiao, G. J., & Han, J. L. 2009, *MNRAS*, **399**, 1587
 Eilek, J. A., & Hankins, T. H. 2016, *JPP*, **82**, 635820302
 Foster, G., Hickish, J., Magro, A., Price, D., & Zarb Adami, K. 2014, *MNRAS*, **439**, 3180
 Gorthi, D. B., Parsons, A. R., & Dillon, J. S. 2021, *MNRAS*, **500**, 66
 Gupta, V., et al. 2022, *MNRAS*, **514**, 5866
 Hankins, T. H., & Eilek, J. A. 2007, *ApJ*, **670**, 693
 Hankins, T. H., Kern, J. S., Weatherall, J. C., & Eilek, J. A. 2003, *Nature*, **422**, 141
 Haskell, B., et al. 2018, *A&A*, **620**, A69
 Keane, E. F. 2013, in *Neutron Stars and Pulsars: Challenges and Opportunities after 80 years*, ed. J. van Leeuwen, *Vol. 291*, 295
 Kent, J., et al. 2019, *MNRAS*, **486**, 5052
 Kent, J., et al. 2020, *MNRAS*, **491**, 254
 Krishnan, H., et al. 2023, *MNRAS*, **520**, 1928
 Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, *Sci*, **318**, 777
 Luo, R., et al. 2024, arXiv e-prints, arXiv:2405.07439
 Masui, K. W., et al. 2019, *ApJ*, **879**, 16
 McLaughlin, M. A., et al. 2006, *Nature*, **439**, 817
 Morales, M. F. 2011, *PASP*, **123**, 1265
 Morales, M. F., & Matejek, M. 2009, *MNRAS*, **400**, 1814
 Morales, M. F., & Wyithe, J. S. B. 2010, *ARA&A*, **48**, 127
 Nimmo, K., et al. 2022, *NatAs*, **6**, 393
 Otobe, E., et al. 1994, *PASJ*, **46**, 503
 Philippov, A., Uzdensky, D. A., Spitkovsky, A., & Cerutti, B. 2019, *ApJ*, **876**, L6
 Pietka, M., Fender, R. P., & Keane, E. F. 2015, *MNRAS*, **446**, 3687
 Polidan, R. S., et al. 2024, *ASR*, **74**, 528
 Price, D. C. 2024, *PASA*, **41**, e037
 Pritchard, J. R., & Loeb, A. 2012, *RPPH*, **75**, 086901
 Reddy, K., et al. 2024, in *Software and Cyberinfrastructure for Astronomy VIII*, ed. J. Ibsen, & G. Chiozzi, Vol. 13101, International Society for Optics and Photonics (SPIE), 131011C
 Slosar, A., et al. 2019, *BAAS*, **51**, <https://baas.aas.org/pub/2020n7i053>
 Snelders, M. P., et al. 2023, *NatAs*, **7**, 1486
 Sokolowski, M., et al. 2024, arXiv e-prints, arXiv:2405.13478
 Taylor, G. B., Carilli, C. L., & Perley, R. A., eds. 1999, *Astronomical Society of the Pacific Conference Series*, Vol. 180, *Synthesis Imaging in Radio Astronomy II* (San Francisco, CA: Astronomical Society of the Pacific)
 Taylor, G. B., et al. 2012, *JAI*, **01**, 1250004
 Tegmark, M. 1997, *ApJ*, **480**, L87
 Tegmark, M., & Zaldarriaga, M. 2009, *PhRvD*, **79**, 083530
 Tegmark, M., & Zaldarriaga, M. 2010, *PhRvD*, **82**, 103501
 Thompson, A. R., Moran, J. M., & Swenson Jr., G. W. 2017, *Interferometry and Synthesis in Radio Astronomy* (3rd edn.; Cham: Springer), doi: 10.1007/978-3-319-44431-4
 Thornton, D., et al. 2013, *Sci*, **341**, 53
 Thyagarajan, N., et al. 2019, *BAAS*, **Vol. 51**, 263

- Thyagarajan, N., Beardsley, A. P., Bowman, J. D., & Morales, M. F. 2017, *MNRAS*, **467**, 715
- Thyagarajan, N., Helfand, D. J., White, R. L., & Becker, R. H. 2011, *ApJ*, **742**, 49
- Tingay, S. J., et al. 2013, *PASA*, **30**, e007
- van Haarlem, M. P., et al. 2013, *A&A*, **556**, A2
- Zackay, B., & Ofek, E. O. 2017, *ApJ*, **835**, 11
- Zhang, J., et al. 2023, *ApJ*, **953**, 65