The LOFAR Two Meter Sky Survey: Deep Fields

I - Direction-dependent calibration and imaging


1. Introduction

With its low observing frequency, wide fields of view, high sensitivity, large fractional bandwidth and high spatial resolution, the Low Frequency Array (LOFAR) is well suited to conduct deep extragalactic surveys. The LOFAR Surveys Key Science Project is building tiered extragalactic surveys with LOFAR, of different depth and areas, and at frequencies ranging from ~ 30 to ~ 200 MHz. Specifically the LOFAR LBA Sky Survey (LoSS) aims at surveying the northern hemisphere using the LOFAR LBA antennas while the LOFAR Two Meter Sky Survey (LoTSS) uses the High Band Antennas (HBA). Its widest component (LoTSS-wide) has been described by Shimwell et al. (2017a, 2019), and aims at reaching noise levels of ~ 100 μJy·beam⁻¹ over the whole northern hemisphere. While the bright sources identified in LoTSS-wide are largely radio loud Active Galactic Nuclei (AGN), the population of faint sources consists of star forming galaxies and radio quiet AGN.
calibration and imaging algorithms developed by Tasse (2014a), Smirnov & Tasse (2015) and Tasse et al. (2018) we were able to estimate and compensate for the \( \phi \), and thus use LOFAR to produce thermal noise limited maps from 8-hour LOFAR observations in a systematic and automated way, while keeping the computational efficiency high enough to be able to deal with the high LOFAR data rates.

In this first paper of a series we present an improved strategy to reach thermal noise limited images after hundreds of hours of integration on the Boötes and Lockman Hole extragalactic fields, reaching \( \sim 30 \mu \text{Jy beam}^{-1} \) noise levels, while being more robust against absorbing faint unmodeled extended emission and dynamic range issues. In Sec. 2 we introduce the \( \phi \) calibration and imaging problem, together with the existing software that we use to tackle it. In Sec. 3 we use ndf-Pipeline-v2 to synthesize deep images over the Boötes and Lockman Hole extragalactic fields and present these deep low frequency images. The subsequent papers in this series will present the deeper ELAIS-N1 data products (Sabater et al. 2020, in prep.), the multiwavelength cross matching (Kondapally et al. 2020, in prep.) and the photometric redshifts and galaxy characterisation (Duncan et al. 2020 in prep.).

2. LoTSS and the third generation calibration and imaging problem

Calibration and imaging techniques have greatly evolved since the first radio interferometers have become operational. First generation calibration is commonly referred as direction-independent (\( \phi \)) calibration, where calibration solutions are transferred to the target from an amplitude and/or phase calibrator field. Second generation calibration is the innovation, beginning in the mid-1970s, of using the target field to calibrate itself (self-calibration; Pearson & Readhead 1984). Third generation calibration and imaging consists in estimating and compensating for direction-dependent effects (\( \phi \phi \)).

As mentioned above, it is challenging to synthesize high resolution thermal noise limited images with LOFAR (van Haarlem et al. 2013). Specifically, LOFAR (i) operates at very low frequency (\( \nu < 250 \text{ MHz} \)), (ii) has very large fields of view (\( \text{FWHM} \) of \( \sim 2 - 10 \) degrees), and (iii) combines short (\( \sim 100 \) m) and long (\( \sim 2000 \) km) baselines to provide sensitivity to both the compact and extended emission. Because of the presence of the ionosphere and the usage of phased array beams, the combination of (i) and (ii) make the calibration problem direction-dependent by nature. In Sec. 2.1 we introduce the mathematical formalism used throughout this paper, while in Sec. 2.2 and 2.3 we describe the algorithms and software used for \( \phi \phi \) calibration and imaging.

2.1. Measurement equation formalism

The Radio Interferometry Measurement Equation (\( \text{RMME} \), see Hamaker et al. 1996) describes how the underlying electric field coherence (the sky model), and the various direction-independent and direction-dependent Jones matrices (denoted \( \mathbf{G} \) and \( \mathbf{J} \) respectively, map to the measured visibilities. In the following, we consider the electric field in linear notation (along the \( x \) and \( y \) axes), at frequency \( \nu \) in direction \( s = [l,m,n] = \sqrt{1 - P - m^2} \hat{T} \) (where \( T \) is the matrix transpose) and write the \( 4 \times 1 \) sky coherency matrix as \( \mathbf{x}_s = [xx,xy,yy,yy]^T \). If \( \mathbf{G}_b = \mathbf{G}^{gb}_{\nu v} \otimes \mathbf{G}^{prv}_{\nu v} \) and \( \mathbf{J}_b = \mathbf{J}^{gb}_{\nu v} \otimes \mathbf{J}^{prv}_{\nu v} \) are the direction-

![Image of effective 1 GHz RMS (Jy beam\(^{-1}\)) vs. surveyed area (deg\(^2\))](image-url)

Fig. 1: This figure shows the effective noise in the LoTSS-Deep continuum maps as compared to other existing and future surveys. A spectral index of \( \alpha \approx -0.7 \) has been used to convert flux densities to the 1.4 GHz reference frequency.
independent and direction-dependent $4 \times 4$ Mueller matrices on a baseline $b \leftrightarrow (pq) \to [u,v,w]^T$ between antenna $p$ and $q$ at time $t$, then the $4$-visibility $v_b$ is given by

$$v_b = G_b \int J_b^T b_{pq} d_s \times \nu + n_b$$

(1)

where $c$ is the speed of light in the vacuum, and $n_b$ is a $4 \times 1$ random matrix following a normal distribution $\mathcal{N}(0, \sigma_b)$. Depending on the context, in the rest of this paper we will either make use of the antenna-based Jones matrices or the baseline-based Mueller matrices.

The elements of $G_b$ describe the direction-independent effects such as the individual station electronics (the bandpass), or their clock drifts and offsets. The $J_b^T$ models the role including the ionospheric distortion (phase shift, Faraday rotation, scintillative decoherence) and phased array station beam that depend on time, frequency, and antenna. Importantly, in order to estimate the intrinsic flux densities we use a description of the LOFAR station primary beam that is built from semi-analytic models and write it as $B_p^s$ in Eq. 1.

Solving for the third generation calibration and imaging problem consists of estimating the terms on the right-hand side of Eq. 1, namely the Mueller matrices $G_b$ and $J_b^T$ and the sky model $X$ from the set of visibilities $v_b$. Due to the bilinear structure of the RIME, instead of estimating all these parameters at once, inverting Eq. 1 is split into two steps. In the first step, the sky term $x_b$ is assumed to be constant, and the Jones matrices are estimated. The step is referred as “calibration” and as the no-$\nu$-RIME system later in this text (or simply C-RIME depending on the context). In the second step the Jones matrices are assumed to be constant, and the sky term $x_b$ is estimated. This step is commonly called “imaging” and is referred as solving the no-$\nu$-I-RIME system later in the text. The C-RIME and I-RIME problems constitute two sub-steps in inverting the RIME system. We will later describe the idea of alternating between no-$\nu$-C-RIME and no-$\nu$-I-RIME as dd-self-calibration.

While the vast majority of modern developments in the field of algorithmic research for radio interferometry aim at addressing either the no-$\nu$-C-RIME (direction dependent calibration, see Yatawatta et al. 2008, Kazemi et al. 2011, Tasse 2014a, Smirnov & Tasse 2015) or no-$\nu$-I-RIME (direction-dependent imaging, see Bhatnagar et al. 2008, Tasse et al. 2010, 2018), in this article we aim at developing a robust approach using existing no-$\nu$-C-RIME and no-$\nu$-I-RIME algorithms to tackle the complete RIME inversion problem.

2.2. Direction-independent calibration

The standard LoTSS HBA observations consist of a 10-minute scan on a bright calibrator source (in general 3C 196 or 3C 295) before observing the target field for 8 hours. On both calibrator and target fields, the visibilities of the 240 subbands (SB) are regularly distributed across the 120-168 MHz bandpass, with 64 channels per 195.3 kHz subband and 1 sec integration time. The data are first flagged using AОFLAGGER (Offeringa et al. 2012) and averaged to 16ch/sb and 1s.

The interferometric data taken on the calibrator field are used to estimate the direction independent Jones matrices $G$ that are, to first order, the same in the target and calibrator fields. These include (i) the individual LOFAR station electronics and (ii) the clock offsets and drifts. This first pass of calibration is conducted using the Prefactor software package (de Gasperin et al. 2019). Specifically, as the calibrator field essentially consists of a single bright source, the measurement equation is direction independent and the visibilities are modeled as

$$v_b^\text{cal} = G_b^\text{model} v_b$$

(5)

where $v_b^\text{model} = \int X_s \times K_b d_s$ is the sky model of the calibrator. We cannot just use $G_b^\text{cal}$ to calibrate the target field, since the ionosphere is different for the calibrator and target fields. Instead, we want to extract (i) the bandpass and (ii) the clock offsets from the calibrator field, these being valid for the target field. The effective Mueller matrix of a given baseline $G_b^\text{tel}$ can be decomposed as the product of a direction independent and direction dependent term

$$G_b^\text{tel} = G_b^\text{cal} J_b^T$$

(6)

with $G_{\nu \nu} = A_{\nu \nu} \exp \{(i\nu \Delta_{\nu})\}$

(7)

and $J_b^T = \exp \{(i\nu K - \Delta_{\nu})\}$

(8)

where $K = 8.44 \times 10^8$ m$^2$s$^{-2}$, and $A_{\nu \nu}$, $\Delta_{\nu}$ and $\Delta_{\nu}$ are real-valued and represent respectively the bandpass, the clock and ionospheric Total Electron Content offset with respect to a reference antenna. The terms $\Delta_{\nu}$ and $\Delta_{\nu}$ can be disentangled from the frequency dependent phases because the former are linear with $\nu$ while the latter are non-linear.

Assuming the clocks and bandpass are the same for the calibrator and for the target field, the corrected visibilities $v_b^c$ for the target field can be built from the raw data $v_b$ as $v_b^c = G_b^{-1} v_b$. In order to calibrate for the remaining phase errors, the target field is calibrated against the TIFR GMRT Sky Survey (ross) catalogs (Intema et al. 2017). The calibration and imaging terms are averaged to 2 ch/sb and 8s.

2.3. Direction-dependent calibration and imaging

As discussed by Tasse (2014b) there are two families of calibration algorithms. “Physics-based” solvers aim at estimating the underlying Jones matrices whose product gives the effective $G^\text{tel}_{\nu \nu}$ and $J_b^T$. Depending on the observing frequency and instrumental setup, these can be the product of the ionospheric Faraday rotation matrix, the scalar phase shift, and the individual station

1 As described by Hamaker et al. (1996), the Mueller matrices and the Jones matrices can be related to each other using the Vec operator. In the context of the rime, if $J_s$ and $J_e$ are $2 \times 2$ Jones matrices of antenna $p$ and $q$, and $X$ is the $2 \times 2$ source’s coherency matrix then we have $\text{Vec}(J_s X J_e) = \text{Vec}(J_s) \otimes \text{Vec}(X)$, where $\otimes$ is the Kronecker product.

2 https://github.com/lofar-astron/LOFARBeam

3 https://sourceforge.net/p/aoflagger/wiki/Home

4 https://github.com/lofar-astron/prefactor
Though there are some limitations on the shape of the neighboring domains in the \( \{ \text{spatial} \} \) space. These "Jones-based" solvers have the advantage of not requiring any physical modeling of the DDE. However, this method may not be feasible or even possible to estimate the effects of beam smearing. This means practice that the no solvers can only solve for the data, leading to the unmolded sky flux being absorbed by the calibration solutions. This happens differently at different scales, and has a greater effect on the extended emission, which is measured only by the less numerous shorter baselines. Similar to linear problems, the situation depends on the sizes of the parameter space, and also on the size of the neighboring domains in the \( \{ \text{spatial} \} \) spaces. Also, as explained by Shimwell et al.[2019], experience shows that we need to split the sky model into a few tens of directions ("facets") to be able to properly describe the spatial variation of the no Jones matrices. This effect is amplified by the difficulty of properly modeling the extended emission itself. Indeed, even in the absence of calibration errors, the deconvolution problem consisting of inverting Eq. 1 by estimating the \( \hat{\sigma}_b \) for a given \( G \) and \( v \) is ill-posed. Furthermore, the situation is more severe when the Point Spread Function (psf) is less point-like (i.e. when the uv plane is not well sampled). The true measured visibilities are described by Eq. 1 of the "model" visibilities \( \hat{\nu}_b \), which estimate the piecewise constant \( \nu \) over the piecewise constant domains \( p, \Omega_y, \Delta t, \Delta v \). 

\[ \hat{\nu}_b = \hat{G}_b \nu_b \]  

(9)

\[ \nu_b^2 = \sum b \nu_b^2 \]  

(10)

\[ \nu_b^2 = \int_b \nu_b^2 \varphi d \nu \]  

(11)

where \( \Omega_y \) is the set of directions \( s \) for a facet \( \varphi \), \( \hat{x}_y \) is the estimated underlying sky, and \( \hat{G}_b \) and \( \hat{f}_b \) are the \( \nu \) and no Mueller matrices for baseline \( b \), built from the corresponding estimated Jones matrices in \( p, \Delta t, \Delta v \). 

Specifically, in order to solve for the no, the size and shape of the domains are critical. Intuitively, if the domains are too small, not enough data points are used, and the solutions are subject to ill-conditioning. On the other hand if the domains are too large the true Jones matrices can vary within the domain and the piecewise constant approximation cannot account for the physics that underlies the measurement. Due to (i) the non-linear nature of Eq. 1 and (ii) the complexity of the background radio sky, optimising over the shape of these piecewise constant domains is a difficult problem (and is non-differentiable to some extent). 

The faceted Jones-based approach is to find sky \( \hat{x}_y \) as well as the \( \nu \) \( G_{\text{prv}} \) and the no piecewise constant \( f_{\text{prv}} \) for all \( \{ \text{spatial} \} \) such that \( \nu_b \sim v_b \) (we remain intentionally vague here, since the cost function that is minimised depends on the specific no algorithm). In practice, the second facet, \( \nu \) (estimating the Jones matrices and sky terms) is done by (i) solving for the Jones matrices assuming the sky is known (the calibration step), and (ii) solving for the sky-term assuming the Jones matrices are given (imaging step). Using this skymodel and repeating steps (i) and (ii) is known as self-calibration, but in a third-generation approach we must explicitly model the DD aspects. 

Since the computing time evolves as \( \sim n_a^3 \) where \( n_a \) is the number of directions in the no-solvers, the problem of no-calibration has in general been tackled using direction alternating peeling-like techniques. Major breakthroughs have been made in the field of no-CRIME solvers in the past decade by Yatawatta et al.[2008], Kazemi et al.[2011] making this no-calibration computationally affordable. In addition, Tasse[2014a] and Smirnov & Tasse[2015] have described an alternative way to write the Jacobian of the cost function by using the Wirtinger differentiation method. The Jacobian and Hessian harbor a different structure and shortcuts can be taken to invert the calibration problem. The net gain over the classical method is not trivial, but can be as high as \( n_a^3 \) (Smirnov & Tasse[2015]) where \( n_a \) is the number of antennas in the interferometer. This Jones-based approach is therefore fast, but is still subject to the same flaws as any Jones-based solvers. 

Only a very few CJ-CRIME algorithms using a full no self-calibration loop have been described and implemented. They include point self-calibration[Butnagar & Cornwell, 2017], or peeling-based techniques such as mimage (implemented in the obit\footnote{https://www.cv.nrao.edu/~bcotton/Obit.html} package) and factor\footnote{https://github.com/mhardcastle/dd-pipeline} (van Weeren et al.[2016], see also Sec. 4.5). Similar to peeling, and developed for reducing LOFAR data, factor is sequential along the direction axis. Looping over the different facets it consists of (i) subtracting all sources besides calibration sources in that one facet, and (ii) \( \nu \)-self-calibrating in that direction. In addition to the ill conditioning issues discussed above on DD-CRIME and DD-J-CRIME solvers, an expensive computational problem arises when estimating the \( f_{\text{prv}} \).

The approach presented by Shimwell et al.[2019] (also described in detail Sec. 3, and referred to as ddf-pipeline-v1 in the following) is based on the kms DD-CRIME solver (Tasse[2014a]; Smirnov & Tasse[2015]) and ddfacet DD-J-CRIME imager (Tasse et al.[2018]), and is algebraically simultaneous in directions. The direction dependent pipeline ddf-pipeline\footnote{https://github.com/saopicc/DDFacet} is a high level wrapper that mainly calls ddfacet\footnote{https://github.com/saopicc/DDFacet} and kms\footnote{https://github.com/saopicc/KILLMS} for direction dependent self-calibration. This type of algorithm has a number of advantages. Specifically, the interaction terms between the different directions are properly taken into account within the DD-CRIME solver, i.e. the no affected sidelobes leaking from any facet to any other facet are accounted for within the algebraic operations of the algorithm. Another advantage compared to the factor approach is that the data need only to be read rather than modified, making the ddf-pipeline move I/O efficient.

3. Calibration and imaging robustness

3.1. Dynamic range issue

With LOFAR's very large field of view, it is quite common to observe bright sources within the station's primary beam. As explained in Sec. 2, the initial phase calibration is done against rss at 150 MHz (Intema et al., 2017). However, since LoTSS resolution is much higher than rss's \( 6'' \times 6'' \) against

Article number, page 4 of 17
25″×25″ respectively), small spatial uncertainties on how the individual bright sources are modeled lead to large Jones matrix errors.

Solution time and frequency variability is however hard to interpret. Indeed, because the prime formalism is subject to unitary ambiguity (see [Hamaker 2000] for a detailed discussion), the off-diagonal or absolute phase terms found by a solver are not meaningful. Instead, these are given with respect to a reference antenna. When Jones matrices are scalar, this amounts to zeroing the phases ϕ₀ of the reference antenna, by subtracting ϕ₀ from all phases of all antennas. To do this in the general case of non-diagonal Jones matrices, we use a polar decomposition on the Jones matrix J₀ of the reference antenna such that J₀ = Uϕ₀ where U is a unitary matrix. We then apply U to all Jones matrices as Jₚ ← U⁻¹ₚ Jₚ. Intuitively, when the Jones matrices are all scalar, the unitary matrix U is simply exp(iϕ₀), and that step makes the phases of all Jₚ relative to the reference antenna (and specifically zeros the phases of J₀). In the case of non-trivial 2 × 2 Jones matrices, finding and applying U has the

11 The unitary matrix U is found by doing a singular value decomposition J₀ = WΣV and is then built as U = WV⁻¹
Algorithm 1: Overview of the algorithm implemented in ddf-pipeline-v2. The function $I$ represents the imaging step and takes as input the visibility vector $v$ together with the beam model $B_\Omega$ and kms-estimated Jones matrices $J_\Omega$ at locations $\Omega_n$. The function $K$ abstracts the do calibration step, and takes as arguments the visibilities $v$, the skymodel $x$, a solver mode (estimating for either scalar or full Jones matrices), a time-frequency solution interval (in min and MHz), and a set of directions $\Omega_n$ in which to solve for. The extra functions $C$, $B$, and $F$ represent the clustering, bootstrapping and smoothing steps respectively.

**Data:** Visibilities $v$ calibrated from $n$ effects using PreFactor.

```
/* On 60 LOFAR HBA subbands */
/* DI initial deconv and clustering */
1.1 \hat{x}_ \leftarrow I(v_0, J_\Omega = 1, B_\Omega);
1.2 \hat{x}_ \leftarrow C(\hat{x}_);
/* DI calibration and imaging */
1.3 \hat{x}_ \leftarrow I(v_0, B_\Omega);
1.4 v_\nu \leftarrow K(v_0, \hat{x}_, B_\Omega|\text{full}, \delta_0, \delta_v, \Omega_0);
1.5 \hat{x}_ \leftarrow I(v_\nu, B_\Omega);
/* Bootstrapping the flux density scale */
1.6 v_\nu \leftarrow B(v_\nu);
/* Phase only DD calibration and imaging */
1.7 \hat{x}_ \leftarrow \varphi \circ F \circ K(v_0, \hat{x}_, B_\Omega|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_0);
1.8 \hat{x}_ \leftarrow I(v_\nu, \hat{J}_B_\Omega);
/* DD calibration and imaging */
1.9 \hat{x}_ \leftarrow \varphi \circ F \circ K(v_0, \hat{x}_, B_\Omega|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_0);
1.10 \hat{x}_ \leftarrow I(v_\nu, \hat{J}_B_\Omega);
/* DI calibration and imaging */
1.11 v_\nu \leftarrow K(v_\nu, \hat{J}_B_\Omega, \hat{x}_|\text{full}, \delta_0, \delta_v, \Omega_0);
1.12 \hat{x}_ \leftarrow I(v_\nu, \hat{J}_B_\Omega);
/* On 240 LOFAR HBA subbands */
/* DD calibration */
1.13 \hat{x}_ \leftarrow F \circ K(v_\nu, \hat{x}_, B_\Omega|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_0);
/* DI calibration */
1.14 v_\nu \leftarrow K(v_\nu, \hat{J}_B_\Omega, \hat{x}_|\text{full}, \delta_0, \delta_v, \Omega_0);
/* DD imaging */
1.15 \hat{x}_ \leftarrow I(v_\nu, \hat{J}_B_\Omega);
/* DD calibration */
1.16 \hat{x}_ \leftarrow F \circ K(v_\nu, \hat{x}_, B_\Omega|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_0);
/* Slow DD calibration */
1.17 \hat{x}_ \leftarrow F \circ K(v_\nu, \hat{J}_B_\Omega, \hat{x}_|\text{scalar}, 43\text{min}, 2\text{MHz}, \Omega_0);
/* Final imaging steps */
1.18 \hat{x}_ \leftarrow I(v_\nu, \hat{J}_B_\Omega);
1.19 Facet-based astrometric correction (see Shimwell et al. 2019 for details);
```

Since the polar transform has been applied, the variations of the amplitude of the off-diagonal Jones matrices are genuine. These are interpretable in terms of differential Faraday rotation: the rotation of the electric field polarisation changes across the LOFAR array. This demonstrates the need to conduct a full-Jones calibration on the PreFactor-calibrated LoTSS data.

Therefore in Step [1.4] the visibilities are calibrated against modeled visibilities generated by n facet in Step [1.3]. The sky being mostly unpolarised, in this full-Jones calibration step, we assume $\vec{Q} = U = V = 0 \text{Jy}$ (see Sec. [3.4] for a discussion of polarisation related data products). The solution intervals $\delta_0$ and $\delta_v$ along time and frequency are determined such that $n_0 \propto (T/\langle |\Delta v| \rangle)^2 \text{Var}[n]$ where $n_0$ is the number of points in the $\delta_0 \times \delta_v$ time-frequency domain, $T$ is the target solution SNR, and $\text{Var}[n]$ is the variance of the visibilities’ noise (see Mbou Sob et al. in preparation for a justification).

Note that after the initial do calibration solutions have been obtained in Steps [1.7] and [1.9], a more accurate do calibration can be performed. Specifically, in the do calibration Steps [1.11] and [1.14] on any baseline $b$ the model visibilities $\hat{v}_b$ (Eq. [10]) are predicted based on the previously estimated do-Jones matrices $\hat{J}$ (as is done by Smirnov [2011]).

### 3.2. Regularisation

The absorption of unmodeled flux by calibration is a well known issue connected to the calibration of doe. Intuitively speaking, when real flux is missing from the modeled sky $\hat{x}$ at step $i$, and since the rime inversion is often ill-posed, the estimates $\hat{J}_i$ of $J$ can be biased in a systematic way. Experience and simulations show that building a new estimate $\hat{x}_i$ from $\hat{J}_i$ can be biased in that the unmodeled emission is not and will never be recovered (Fig [5]). This effect is especially severe when the extended emission is poorly modeled or unmodeled since this is detected only by the shortest baselines. Effectively, during the inversion of the rime system of equations, the do-self-calibration algorithm has fallen into the wrong (local) minimum.

In order to address this problem, one idea is to reduce the effective number of free parameters used to describe the Jones matrices in the $|pdtv|-$space (see for example Tasse 2014a, Yatawatta 2015, van Weeren et al. 2016, Repetti et al. 2017).
Fig. 4: This figure shows the amplitude and phase (top/bottom respectively) of a scalar Jones matrix for a given station in a given direction in the example observation. The left panel shows the solution as estimated by the kms solver. The right panel shows the regularised solution, as updated by the $F$ function. The amplitude color scale ranges from 0 to 1.5.

Birdi et al. 2020. Forcing the estimated Jones matrices’ shape to look like that of the real underlying ones improves the conditioning of the inverse problem. In Alg. 1 (implemented in ddf-pipeline-v2) we have replaced that normalisation method by a smoothing of the kms-estimated Jones matrices. This function $F$ updates the Jones matrices $\hat{J} \leftarrow F(\hat{J})$ by imposing on them a certain behavior in the time-frequency space, effectively reducing the size of the unknown stochastic process. This can be thought of as a regularisation. This is done independently on the phases and amplitudes on the scalar Jones matrices generated at Steps 1.9, 1.13, 1.16. The updated Jones matrices take the analytical form

$$\tilde{J}_{\text{pd},\nu} = \hat{a}_{\text{pd},\nu} P(t, \hat{\theta}_{\text{pd},\nu}) \exp \left( i K^{-1} \Delta T_{\text{pd},\nu} \right) I$$

(12)

where $\Delta T_{\text{pd},\nu}$ is the differential $\tau$, $P$ is a polynomial parametrised by the coefficients in $\hat{\theta}_{\text{pd},\nu}$ (of size 10), and $\hat{a}_{\text{pd},\nu}$ is a scalar meant to describe the loss of correlation due to ionospheric scintillation as seen in the left panel of Fig. 4. Typically, for the ~8 hours’ integration of LoTSS pointings and solving every 30 sec. and 2 MHz, this parametrisation of the Jones matrices reduces the number of free parameters by a factor $\gtrsim 20$.

To assess the recovery of unmodelled flux in ddf-pipeline-v2 a series of simulations were conducted in which faint simulated sources of various fluxes and extents were injected into real LOFAR data that had been fully processed with the ddf-pipeline-v2 strategy. The properties of the injected sources were chosen to be typical for large extragalactic objects such as radio halos of galaxy clusters. After the injection of the artificial extended sources the steps 1.16 and 1.17 were repeated using the sky model derived at step 1.18 prior to the injection of the sources. These simulations will be discussed further by Shimwell et al. (in preparation) but in each simulation the recovered flux of the completely unmodelled emission exceeded 60%. Examples of the injected and recovered emission are shown in Fig 5.
As suggested by the results of simulations, the effect on real data is in general very satisfactory and allows us to recover the unmodeled extended emission even when it is quite faint and extended. This is shown in Fig. 6 for a typical LoTSS observation. Here the extended emission is about 10′ across, with a mean flux density at the peak of only 0.7 of the local standard deviation.

3.3. Conditioning and solution interval

The additional issue of arcmin-scale negative haloes appearing around bright compact sources (at a level of $\leq 1\%$ or the peak) could be seen however in $\sim 10 - 20\%$ of the LoTSS pointings processed with df/PIPELINE-V1. As shown in Fig. 6 we believe this to be connected to the solution regularisation itself. This issue is hard to understand in detail because of the non-linearity in the C-RIME inversion, but is likely to be due to the pointings showing these issues being more severely affected by the incompleteness of the sky model. Specifically, conducting several experiments, we were able to observe that the situation was improved by deconvolving deeper or taking into account sources outside the synthesized image field of view.

An additional way to improve the conditioning of the problem is to increase the amount of data used to constrain the estimated Jones matrices. For the no calibration steps presented in Alg. 1 we use solution intervals of $0.5 - 1$ minute.

3.4. Data products

3.4.1. Unpolarised flux

Once the estimated no-Jones matrices and skymodel $\hat{S}$ have been obtained at the highest available spatial resolution following the no/no-self-calibration steps presented in Alg. 1 additional data products are formed.

Users can adapt the weighting scheme depending on the scientific exploitation they want to make of the interferometric data. This is very much tied to how the calibration and deconvolution algorithms are working, and concurrent effects take place along the self calibration loop. Extended emission is hard to properly model since the deconvolution problem is more ill-posed in these cases (more pixels are non-zero). To tackle this issue the rsr can be modified to make the convolution matrix more diagonal and the deconvolution problem correspondingly better conditioned.

For all these concurring reasons the faint and extended flux in the highest resolution maps produced by Alg. 1 is either poorly modeled or not deconvolved at all. Since the pixel values of extended sources are not interpretable in the residual maps, the flux density of the radio sources cannot be measured if they are not deconvolved. We therefore intentionally degrade the resolution density at the peak of only 0.1% or the peak (16 rsr have to be computed (as there are 16 terms in the quadratic mean of the Mueller matrices). As most of the sources are unpolarised, the leakage terms are properly taken into account in the no-predict. Instead of deconvolving the polarised flux, we grid the IQUV residual data. The polarised flux is directly postprocessing to be made such as better calibration towards a particular point on the sky (van Weeren et al in prep.), and also reimaging at different resolutions if required.

3.4.2. QUV images

The no-racet no-imager only deals with I-Stokes deconvolution. As discussed by Tasse et al. (2018), estimating the QUV Stokes parameters is complex in the context of no-imaging due to the leakage terms. Indeed, for the problem to be properly addressed, the rsr have to be computed (as there are 16 terms in the quadratic mean of the Mueller matrices). As most of the sources are unpolarised, the leakage terms are properly taken into account in the no-predict. Instead of deconvolving the polarised flux, we grid the IQUV residual data. The polarised flux is directly interpretable when the sources are unresolved. Hence we also generate the following additional products:

4. Low resolution (20′′) spectral Stokes QU cubes (480 planes - Step 2b.3)
5. Very low resolution Stokes QU cubes (480 planes - Step 2b.4), by cutting the baselines > 1.6 km, giving an effective resolution of $\sim 3′$
6. Low resolution (20′′) wide-bandwidth Stokes V image (Step 2b.5)

The output QU cubes are processed using Faraday rotation measure (RM) synthesis (Brentjens & de Bruyn 2005) to find polarised sources and their RM with the sensitivity of the full bandwidth. The wide bandwidth (120 to 168 MHz) combined with the narrow channel width (97.6 kHz) provides a resolution in RM space of $\sim 1.1 \text{rad/m}^2$ and an ability to measure RMs of up to $\sim 450 \text{rad/m}^2$ (e.g. O’Sullivan et al. 2020).

The 3′ QU cubes are sensitive to the large-scale polarised emission from the Milky Way, while the 20′′ QU cubes are excellent for finding compact polarised sources. However, detailed studies of the polarisation and RM structure of resolved extra-galactic sources will require deconvolution of the Q and U data. The df/PIPELINE-V2 output provides significantly better performance in correcting for the effect of the instrumental polarisation (Fig. 8, which is typically at the level of 1% or less for bright total intensity sources (O’Sullivan et al., in prep).

There is no absolute polarisation angle calibration for each LoTSS observation, meaning that while the RM values of sources in overlapping fields are consistent, the polarisation angles are not. Therefore, to avoid unnecessary depolarisation for both mosaicing and the deep fields, the polarisation angles between the observations need to be aligned. The simplest way to do this is by choosing a reference angle of a polarised source in a single observation and applying a polarisation angle correction to all other observations to align with this reference angle, as presented in Herrera Ruiz et al. (2020). An alternative approach is to use the diffuse polarised emission that is present in the $\sim 3′$ QU cubes.

Bright polarised sources are rare in the LoTSS data, with only three sources having a polarisation intensity greater than 50 mJy beam$^{-1}$ in the DR1 HETDEX sky area (Van Eck et al. 2018, O’Sullivan et al. 2018). However, in the fields containing these bright polarised sources the df/PIPELINE-V2 output becomes unreliable for polarised sources. This limitation likely arises from assuming $Q = U = V = 0$ Jy for a field in the no calibration step. While only a few percent of fields are strongly affected, the exact extent of this issue is being investigated further through simulations, where bright polarised sources are inserted into existing.
Fig. 6: Conserving the unmodeled extended emission while keeping high dynamic range is extremely challenging in the context of no calibration and imaging. The left panel shows that a faint and unmodeled extended emission (on the level of Fig. 6) can be totally absorbed. While regularising the no calibration solutions can help in recovering the unmodeled emission (typically after Step 1.16), it can also produce negative imaging artifacts and ‘holes’ around bright sources (middle panel). The right panel shows that solving the residuals on longer time intervals (Step 1.17) corrects for this issue.

LoTSS uv-datasets. Possible solutions will be tested in future pipeline developments.

3.5. Comparison between ddf-pipeline-v1 and ddf-pipeline-v2

3.6. ddf-pipeline-v2 robustness and performance

As explained above ddf-pipeline-v2 is a high level script interfacing kms and nVes. Both of the underlying software packages are efficiently parallelised using a custom version of the Python multiprocess package for process-level parallelism, and using the.SharedArray module. As explained by Tasse et al. (2018), this pythonic approach minimizes the process interconnections for both the kms and nVes software.

This paper considers the application of ddf-pipeline-v2 to the LoTSS-Deep Fields. The pipeline is also being used to process data from the wider and shallower LoTSS survey. The LoTSS project is presently observing at a rate of up to 1,500 hrs every 6 month cycle which corresponds to approximately two 8 hr pointings observed simultaneously) each day. The ddf-pipeline-v2 compute time is roughly split equally between calibration and imaging tasks (see Fig 3). The total run time for an 8 hour pointing is ~ 5 days (on a node equipped with 192 GBytes RAM and 2 Intel Xeon Gold 6130 CPU@2.10GHz, giving 32 physical compute cores), and takes an extra ~ 30% of computing time to completion as compared to ddf-pipeline-v1. Hence 10 compute nodes are sufficient to keep up with the observing rate. However, in practice more compute nodes are used because LoTSS has been observing since 2014 and as of June 1st 2019 over 1,000 pointings exist in the archive. Over ~ 1000 pointings and ~ 12 PB of averaged and compressed LOFAR data (~ 40 PB uncompressed) have now been processed with ddf-pipeline-v2.

4. LoTSS deep fields data and processing

4.1. Observations

LoTSS-Deep Fields observations are being carried out over the four northern fields with high-Galactic latitude and the highest-quality multi-degree-scale ancillary data across the electromagnetic spectrum: the Boötes field, the Lockman Hole, ELAIS-N1 and the North Ecliptic Pole fields. The ultimate aim of the LoTSS Deep Fields project is to reach noise levels of 10-15 \( \mu \text{Jy} \cdot \text{beam}^{-1} \) in each of these fields (requiring ~ 500 hours of integration). The first LoTSS-Deep Fields data release consists of initial observations in three of these fields: Boötes (~ 80 hrs) and Lockman Hole (~ 112 hrs) presented in the current paper, and ELAIS-N1 (presented by Sabater et al. 2020 for an integration time of ~ 170 hrs in paper 2). This first data release also includes an extensive effort of optical/IR cross-matching, which has obtained host galaxy identifications for over 97% of the ~80,000 radio sources detected within the ~ 25 deg\(^2\) overlap with the high-quality multi-wavelength data (Kondapally et al. 2020, Paper 3). This is supplemented by high quality photometric redshifts, and characterisation of host galaxy properties (Duncan et al. 2020, Paper 4), and source classification (e.g. star-forming vs AGN: Best et al. 2020, Paper 5).

In order to put the LoTSS-deep observations in a wider context, in this section we briefly describe the multi-wavelength data available on the Boötes and Lockman Hole fields, focusing on the radio coverage (for a more detailed description see Kondapally et al. 2020, Paper 3).

4.1.1. Boötes field

The Boötes field is one of the NOAO Deep Wide Field Survey (NDWFS Jannuzi & Dev 1999) fields covering ~ 9.2 deg\(^2\). It contains multi-wavelength data including infrared (Spitzer space telescope, see Ashby et al. 2009, Jannuzi et al. 2010), X-rays (Chandra space telescope, see Murray et al. 2005, Kenier et al. 2005), optical data (Jannuzi & Dev 1999, Cool 2007, Brown et al. 2007, 2008). At radio frequencies it has been mapped with the Westerbork Radio Telescope (WSRT, see de Vries et al. 2002), the Very Large Array (VLA, see Croft et al. 2008, Coppejans et al. 2015), the Giant Meterwave Radio Telescope (GMRT,see Intema et al. 2011), and LOFAR (van Weeren et al. 2014, Williams et al. 2016, Retana-Montenegro et al. 2018) at various depths, frequencies, resolutions and cov-
Fig. 7: This figure shows the differences between the maps produced by Alg. 0 and Alg. 1 from a typical 8 hour scans (here the P26H point in the HETDEX field, see Shimwell et al. 2017a). The colorscale is the same on all panels, and displayed using an inverse hyperbolic sine function to render both the low level artifacts and some bright sources.
Cycle 2 and Cycle 4, with a bandwidth of 48 MHz (see Tab. 1). The total integration time of ~ 80 hours is spread over 10 scans of 8 hours.

4.1.2. Lockman hole

The Lockman Hole field is also covered by a large variety of multiwavelength data. Specifically, it has been observed by the Spitzer Wide-area Infrared Extragalactic survey (SWIRE [Lonsdale et al. 2003]) over ~ 11 deg$^2$, and over 16 deg$^2$ by the Herschel Multi-tiered Extragalactic Survey (Oliver et al. 2012). It has also been observed in UV (Martin & GALEX Team 2005), optical (Gonzalez-Solares et al. 2011), near IR (UK Infrared Deep Sky Survey Deep Extragalactic Survey UKIDSS-DXS, see Lawrence et al. 2007), and with the Submillimetre Common-User Bolometer Array (Coppin et al. 2006; Geach et al. 2017). At higher energy, it has been observed with XMM-Newton (Brunner et al. 2008), and Chandra (Polletta et al. 2006).

In the radio domain, the Lockman Hole has been observed over the two deep aforementioned X-ray fields over small sub-deg$^2$ areas (de Ruiter et al. 1997; Ciliegi et al. 2003; Biggs & Ivison 2006; Ibar et al. 2009). Wide surveys of the Lockman Hole have been done with GMRT (Garn et al. 2010), VLA (Owen et al. 2017). At higher energy, it has been observed with XMM-Newton (Brunner et al. 2008), and Chandra (Polletta et al. 2006).

The total integration time of ~ 80 hours is spread over 10 scans of 8 hours.

The Bo"otes pointings data that are presented in this paper consists of 12 pointings of 8 hours centered on (a, $\delta$) = (14h32m00s, +34°30'00") and were observed with the LOFAR-HBA in hba_dual_inner mode during Cycle 2 and Cycle 4, with a bandwidth of 48 MHz (see Tab. [1]).
4.2. Image synthesis

The Lockman Hole and Boötes fields data have been both reduced using Alg. 2. In this approach we first build a wide-band \(\Omega_0\) self-calibrated sky model \(\mathbf{x}_0\), from a single wide band \(\sim 8\) hours observation using Alg. 1. This model is then used to \(\Omega_0\) self-calibrate all the \(n_p\) pointings (with \(n_p = 10\) and \(n_p = 12\) for the Boötes and Lockman Hole datasets respectively) following Alg. 2. This amounts to repeating Steps 1.13 to 1.18 of Alg. 1 on a larger dataset. A comparison between the images synthesised from 8 and 80 hours datasets is presented in Fig. 11. On a single node equipped with \(\sim 500\) GB of 2.4 GHz RAM and 2 Intel Xeon CPU E5-2660 v4@2.00GHz with 14 physical cores each, Alg. 2 took \(\sim 21\) days to process the 80 hours of Boötes data.

Fig. 12 (further discussed in Sec. 4.3) and 13 show the central parts of the of these deep LOFAR Boötes and Lockman Hole observations.

Estimating the noise in radio maps is not straightforward since noise is correlated and non-Gaussian. Also, while the covariance matrix should be entirely described by the rsn, the real covariance matrix is hard to estimate due to the calibration artifacts (see Tasse et al. 2018; Bonnassieux et al. 2018 for a detailed discussion). Here, in order to estimate the local noise we use the statistics of the min(.) estimator (that returns the minimum value of a given sample). Intuitively, while the I-Stokes image max(.) statistics has contributions from both artifacts and real sources, the min(.) only accounts for the artifacts. A min(.) filter with a given box size is therefore run through a restored image, and depending on the box size\(^\text{13}\), the effective standard deviation is derived.

Fig. 14 shows the cumulative distribution of the local noise in the Lockman Hole and Boötes fields maps, reaching \(\lesssim 23\) and \(\lesssim 30\) \(\mu\)Jy\,beam\(^{-1}\) respectively. Taking into account the number of pointings with their respective amount of flagged data, we get total integration times of \(\sim 65\) and \(\sim 88\) hours on the Boötes and Lockman Hole fields respectively, giving a theoretical thermal noise difference of a factor \(\sim 1.16\) compatible with the observed \(13\) The cumulative distribution \(F\) of \(Y = \min [X]\) with \(X \sim N(\mu = 0, \sigma = 1)\) is \(F(y) = 1 - \left[ \frac{1}{2} \left( 1 - \text{erf} \left( \frac{y}{\sqrt{2}} \right) \right) \right]^n\), where \(n\) is the number of pixels in a given box. Finding \(y_{\sigma}\) such that \(F(y_{\sigma}) = 1/2\) given the box size gives us a conversion factor from the minimum estimate to the standard deviation.
4.3. Comparison with deep factor image synthesis

The image of the Boötes field based on 55 hours of LOFAR HBA data and presented by Retana-Montenegro et al. (2018) reaches an unprecedented noise level image of ~ 55 \( \mu \text{Jy beam}^{-1} \) at 150 MHz. To achieve such high sensitivity, Retana-Montenegro et al. (2018) have applied third generation calibration and imaging to correct for the dde using the factor package (developed by van Weeren et al. 2016; see Sec. 2 for more detail). Because the set of LOFAR datasets used by Retana-Montenegro et al. (2018) is different, the comparison can only be approximate. In Fig. 12 we compare the images produced by Retana-Montenegro et al. (2018) and by Alg. 2. While the noise difference should be on the order of 20%, as shown in Fig. 14, the measured one is on the level of ~ 60%. Consistently artifacts around bright sources are also much less severe in the maps generated by Alg. 2 and implemented in DDF-Pipeline-v2.

Algorithm 2: Overview of the algorithm implemented in DDF-Pipeline-v2. The function \( I \) represents the imaging step and takes as input the visibility vector \( \mathbf{v} \) together with the beam model \( \mathbf{B}_{\Omega} \) and km-s-estimated Jones matrices \( \mathbf{J}_{\Omega} \) at locations \( \Omega_\nu \). The function \( K \) abstracts the no calibration step, and computes the sky-model \( \mathbf{\tilde{X}}_\nu \), a solver mode (estimating for either scalar or full Jones matrices), a time-frequency solution interval (in min and MHz), and a set of directions \( \Omega_\nu \) in which to solve for. The extra functions \( C, \mathcal{B}, \) and \( \mathcal{F} \) represent the clustering, bootstrapping and smoothing steps respectively.

**Data:** Visibilities \( \mathbf{v} \) calibrated from dm effects using

PreFactor of \( n_p \times 8 \) hours observations (each with 240 LOFAR-HBA subbands), as well as the high resolution skymodel built in step 1.18

**Result:** Deconvolved image \( \mathbf{x}_\nu \)

\[
\begin{align*}
&/\ast \text{ On } n_p \times 240 \text{ LOFAR HBA subbands } \ast / \\
&/\ast \text{ DD calibration } \ast / \\
&2.1 \quad \mathbf{\tilde{J}} \leftarrow \mathcal{F} \circ K \left( \mathbf{v}_{n_p \times 24}, \mathbf{B}_{\Omega_\nu}, \mathbf{\tilde{X}}_\nu; \text{scalar, 1min, 2MHz, } \Omega_\nu \right) \\
&/\ast \text{ DI calibration } \ast / \\
&2.2 \quad \mathbf{v}_{n_p \times 24} \leftarrow K \left( \mathbf{v}_{n_p \times 24}, \mathbf{J}_{\Omega_\nu}, \mathbf{\tilde{X}}_\nu; \text{full, } \delta_{\Omega_\nu}, \delta_{\nu_0}, \Omega_\nu \right) \\
&/\ast \text{ DD imaging } \ast / \\
&2.3 \quad \mathbf{\tilde{x}}_\nu \leftarrow I \left( \mathbf{v}_{n_p \times 24}, \mathbf{J}_{\Omega_\nu}; \text{scalar} \right) \\
&/\ast \text{ DD calibration } \ast / \\
&2.4 \quad \mathbf{\tilde{J}} \leftarrow \mathcal{F} \circ K \left( \mathbf{v}_{n_p \times 24}, \mathbf{B}_{\Omega_\nu}, \mathbf{\tilde{x}}_\nu; \text{scalar, 1min, 2MHz, } \Omega_\nu \right) \\
&/\ast \text{ Slow DD calibration } \ast / \\
&2.5 \quad \mathbf{\tilde{J}}_\nu \leftarrow K \left( \mathbf{v}_{n_p \times 24}, \mathbf{J}_{\Omega_\nu}, \mathbf{\tilde{x}}_\nu; \text{scalar, 43min, 2MHz, } \Omega_\nu \right) \\
&/\ast \text{ Final imaging steps } \ast / \\
&2.6 \quad \mathbf{\tilde{x}}_\nu \leftarrow I \left( \mathbf{v}_{n_p \times 24}, \mathbf{\tilde{J}}_\nu; \mathbf{J}_{\Omega_\nu}; \text{scalar} \right) \\
&/\ast \text{ Absolute flux density scale correction } \ast / (\text{see Sabater et al. 2020, for details}) \\
&/\ast / \\
&2.7 \quad \mathbf{\tilde{x}}_\nu \leftarrow \lambda \mathbf{x}_\nu; \\
&/\ast \text{ Facet-based astrometric correction (see Shimwell et al. 2019 for details) } \ast / \\
\end{align*}
\]

4.4. Cataloguing

In order to extract astrophysical information we build a catalogue of radio sources from the images produced by Alg. 2 and the data described in Sec. 4. Even in the apparent flux
(a) The central $\geq 2$ deg$^2$ part of the Bootes field as imaged by the direction dependent factor algorithm (Retana-Montenegro et al. 2018).

(b) Zoom in on region (1) of the map synthesised by Retana-Montenegro et al. (2018).

(c) Zoom in on region (1) of the map synthesised by kms-ddfacet (this work).

(d) The same as in 12a but imaged with Alg. 2.

(e) Zoom in on region (2) of the map synthesised by Retana-Montenegro et al. (2018).

(f) Zoom in on region (2) of the map synthesised by kms-ddfacet (this work).

Fig. 12: Comparison between the LOFAR-HBA maps generated at 150 MHz by Retana-Montenegro et al. (2018) and in the current work. The colorscale is the same on all panels, and displayed using an inverse hyperbolic sine function to render both the low level artifacts and some bright sources.

Maps, because of the imperfect calibration and imaging, the LotSS-deep images have spatially variable noise, and to deal
(a) The central $\geq 2 \text{ deg}^2$ part of the Lockman Hole field as imaged by Alg. 2 (Sec. 4.2).

(b) Zoom in on region (1) of the map shown in Fig. 13a.

(c) Zoom in on region (2) of the map shown in Fig. 13a.

Fig. 13: This figure shows the central region of the deep LOFAR-HBA maps of the Lockman Hole field generated at 150 MHz. The colorscale is the same on all panels, and displayed using an inverse hyperbolic sine function to render both the low level artifacts and some bright sources.

Fig. 14: The cumulative distribution of the local noise estimates in the various maps discussed here. As shown here, we have imaged a larger fraction of LOFAR’s HBA primary beam than the image presented in Retana-Montenegro et al. (2018).

with this issue we use PyBDSF\(^\text{15}\) (Python Blob Detector and Source Finder, see Mohan & Rafferty 2015) since it measures noise locally rather than globally. The sources were detected with a 3 and 5$\sigma$ for the island and peak detection threshold respectively. The position-dependent noise was estimated using a sliding box algorithm with a size of $40 \times 40$ synthesised beams, except around bright sources where the box size was decreased to $15 \times 15$ beams to more accurately capture the increased noise in these regions. The columns kept in the final catalogue are the source position, peak and integrated flux density, source size and orientation, the associated uncertainties, the estimated local rms at the source position, as well as a code describing the type of structure fitted by PyBDSF. As described in Sabater et al. (2020), the peak and integrated flux densities of the final catalogs and images are corrected from overall scaling factors of 0.920 and 0.859 for the for Lockman Hole and Boötes fields respectively. The full catalogues cover out to 0.3 of the power primary beam and contain 36,767 entries over 26.5 square degrees and 50,112 over 25.0 square degrees for Boötes and Lockman Hole respectively. These raw PyBDSF catalogues are available online on the LOFAR survey webpage https://www.lofar-surveys.org/ and a thorough analysis of the source catalogues will be presented by Mandal et al. in preparation.

5. Conclusion and future plans

Imaging low-frequency LOFAR data at high resolution and over wide fields of view is extremely challenging. This is mainly due to the rime system being complex in this regime: the background wide-band sky is unknown, as are the time-frequency-antenna $\varpi$-Jones matrices. Due to the high number of free parameters in that system, and to the finite amount of data points

\footnote{https://www.astron.nl/citt/pybdsf}
in the non-linear rime system, the inversion can be subject to ill-conditioning and the do-C-rime solver can absorb unmodeled extended flux.

In order to address this robustness issue we have developed a strategy that aims at conserving the unmodeled emission without affecting the final dynamic range. The method we have developed has similarities with those presented by Yatawatta et al. (2015); van Weeren et al. (2016); Repetti et al. (2017); Birdi et al. (2020), and relies on reducing the effective size of the unknown stochastic process. We show that this allows us to recover most of the faint unmodeled extended emission.

We have applied this third generation calibration and imaging do-algorithm both to the wide-field imaging of the LoTSS survey and to the synthesis of deep 150 MHz resolution images on the Boötes and Lockman Hole fields. The synthesized images are the deepest ever obtained at these frequencies. In the future we plan to continue increasing the depth of these fields: data are already in hand, or scheduled, to double the integration time on each field, with a further aim to increase this to 500 hours in each field.

6. Acknowledgements

This paper is based (in part) on data obtained with the International LOFAR Telescope (ILT). LOFAR (van Haarlem et al. 2013) is the Low Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources), and which are collectively operated by the ILT foundation under a joint scientific policy. The ILT resources have benefitted from the following recent major funding sources: CNRS-INSU, Observatoire de Paris and Université d’Orléans, France; BMBF, MIWF-NRW, MPG, Germany; Science Foundation Ireland (SFI), Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; The Science and Technology Facilities Council, UK; Ministry of Science and Higher Education, Poland.

This work makes use of kern astronomical software package (available at https://kernsuite.info and presented in Molenaar & Smirnov 2018).

MB acknowledges support from INAF under PRIN SKA/CTA FORECaST. MB acknowledges the support from the Ministero degli Affari Esteri della Cooperazione Internazionale - Direzione Generale per la Promozione del Sistema Paese Progetto di Grande Rilevanza ZA18GR02.

MJH acknowledges support from the UK Science and Technology Facilities Council [ST/N000919/1] and the Oxford Hintze Centre for Astrophysical Surveys which is funded through generous support from the Hintze Family Charitable Foundation.

PNB and JS are grateful for support from the UK STFC via grant ST/R000972/1.

MJH acknowledges support from STFC via grant ST/R000905/1.

WLW acknowledges support from the ERC Advanced Investigator programme NewClusters 321271. WLW also acknowledges support from the CAS-NWO programme for radio astronomy with project number 629.001.024, which is financed by the Netherlands Organisation for Scientific Research (NWO).

AB acknowledges support from the VIDI research programme with project number 639.042.729, which is financed by the Netherlands Organisation for Scientific Research (NWO).

IP acknowledges support from INAF under the SKA/CTA PRIN “FORECaST” and the PRIN MAIN STREAM “SAuROS” projects.

MB acknowledges support from INAF under PRIN SKA/CTA FORECaST and from the Ministero degli Affari Esteri della Cooperazione Internazionale - Direzione Generale per la Promozione del Sistema Paese Progetto di Grande Rilevanza ZA18GR02.

RK acknowledges support from the Science and Technology Facilities Council (STFC) through an STFC studentship.

Appendices

A. LoTSS first data release: overview of ddf-pipeline-v1

The data processing strategy of the LoTSS first data release (DR1) has been extensively described by Shimwell et al. (2019). Since addressing the issues described in Sec. 2 involves making improvements relative to this approach, we give here a brief description of the data reduction strategy in ddf-pipeline-v1 (the various steps are outlined in Alg. 0).

As discussed in Sec. 2, the calibration and imaging problem is non-convex and ill-posed. Beyond the computational issues, the great difficulty of the calibration of the ddf is sky incompleteness, because the do-C-rime non-linear system can be subject to ill-conditioning. This is due to the fact that the extended emission (i) is hard to model in the deconvolution step, and (ii) is seen by only the shortest baselines, and therefore sky incompleteness biases the Jones matrices in the calibration step. Experience shows that this leads to some of the unmodeled extended emission being absorbed when running a ddf deconvolution with ddfacet.

To try to compensate for this effect, in ddf-pipeline-v1 (Alg. 0) we introduced an inner w-distance cut during calibration, as well as a normalization of the Jones matrix. With this the ddf-pipeline-v1 was able to recover some of the unmodeled extended emission. The underlying idea was to assume the sky incompleteness was generating some baseline-dependent systematic errors. However, as shown in Fig. 4 and explained by Shimwell et al. (2019) it also produced large scale fake haloes centered on extended sources together with artifacts around bright sources. On fields having a bright ≳ 1 Jy source within the primary beam (such as 3C sources), ddf-pipeline-v1 was not able to converge.

References

Algorithm 0: Overview of the algorithm implemented in ddf-pipeline-v1 to produce the LoTSS-DR1 images. The function $I$ represents the imaging step and takes as input the visibility vector $v$ together with the beam model $B_0$, and ksm-estimated Jones matrices $J_0$ at locations $\Omega_0$. The function $K$ abstracts the no calibration step, and takes as arguments the visibilities $v$, the sky model $\chi$, a solver mode (estimating for either scalar or full Jones matrices), a time-frequency solution interval (in min and MHz), and a set of directions $\Omega$ in which to solve for. The extra functions $C$, $N$, and $B$ represent the clustering, normalisation (see text), and bootstrap-strapping steps respectively.

Data: Visibilities $v$ calibrated from $m$ effects using Prefactor.

Result: Deconvolved image $\hat{\chi}$.

/* On 60 LOFAR HBA subbands */
/* DI initial deconv and clustering */
0.1 $\hat{\chi} \leftarrow I(v_m B_0)$;
0.2 $\Omega_0 \leftarrow C(\hat{\chi})$;
/* Phase only DD calibration */
0.3 $\hat{J} \leftarrow \varphi \cdot K(v_0, \hat{\chi}, B_0)$|scalar, 1min, 2MHz, $\Omega_0$;
/* Absolute flux density scale bootstrapping */
0.4 $\nu \leftarrow B(\nu_0)$;
0.5 $\chi_0 \leftarrow I(v_0 B_0)$;
/* DD calibration and imaging */
0.6 $\hat{J} \leftarrow N \cdot K(v_0, \chi_0, B_0)$|scalar, 1min, 2MHz, $\Omega_0$;
0.7 $\hat{\chi} \leftarrow I(v_0 B_0)$;
/* On 240 LOFAR HBA subbands */
/* Deep DD calibration and imaging */
0.8 $\hat{J} \leftarrow N \cdot K(v_24, \hat{\chi}, B_0)$|scalar, 1min, 2MHz, $\Omega_0$;
0.9 $\chi_0 \leftarrow I(v_24 B_0)$;
0.10 Facet-based astrometric correction (see Shimwell et al. 2019);