A&A 648, A1 (2021)
https://doi.org/10.1051/0004-6361/202038804
© C. Tasse et al. 2021

Astronomy & Astrophysics

The LOFAR Two Meter Sky Survey

Special issue

The LOFAR Two-meter Sky Survey: Deep Fields Data Release 1

I. Direction-dependent calibration and imaging

C. Tasse[1,2,3], T. Shimwell[4,5], M. J. Hardcastle[6], S. P. O'Sullivan[7], R. van Weeren[5], P. N. Best[8], L. Bester[9,2], B. Hugo[9,2], O. Smirnov[2,9], J. Sabater[8], G. Calistro-Rivera[10], F. de Gasperin[11], L. K. Morabito[12], H. Röttgering[5], W. L. Williams[5], M. Bonato[13,14,15], M. Bondi[13], A. Botteon[5,13], M. Brüggen[11], G. Brunetti[13], K. T. Chyży[16], M. A. Garrett[17,5], G. Gürkan[18], M. J. Jarvis[19,20], R. Kondapally[8], S. Mandal[5], I. Prandoni[13], A. Repetti[21], E. Retana-Montenegro[22], D. J. Schwarz[23], A. Shulevski[24], and Y. Wiaux[21]

[1] GEPI, Observatoire de Paris, CNRS, Université Paris Diderot, 5 place Jules Janssen, 92190 Meudon, France
e-mail: cyril.tasse@obspm.fr
[2] Centre for Radio Astronomy Techniques and Technologies, Department of Physics and Electronics, Rhodes University, Grahamstown 6140, South Africa
[3] USN, Observatoire de Paris, CNRS, PSL, UO, Nançay, France
[4] ASTRON, Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, 7991 PD, Dwingeloo, The Netherlands
[5] Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands
[6] Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK
[7] School of Physical Sciences and Centre for Astrophysics & Relativity, Dublin City University, Glasnevin, D09 W6Y4, Ireland
[8] SUPA, Institute for Astronomy, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK
[9] South African Radio Astronomy Observatory, Observatory 7925, Cape Town, South Africa
[10] European Southern Observatory, Karl-Schwarzchild-Strasse 2, 85748 Garching bei München, Germany
[11] Hamburger Sternwarte, University of Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany
[12] Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham, DH1 3LE, UK
[13] INAF – Istituto di Radioastronomia, Via P. Gobetti 101, 40129 Bologna, Italy
[14] INAF – Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, 35122, Padova, Italy
[15] Italian ALMA Regional Centre, Via Gobetti 101, 40129 Bologna, Italy
[16] Astronomical Observatory, Jagiellonian University, ul. Orla 171, 30-244 Kraków, Poland
[17] Jodrell Bank Centre for Astrophysics, University of Manchester, Alan Turing Building, Oxford Road, M13 9PL, UK
[18] CSIRO Astronomy and Space Science, PO Box 1130, Bentley WA 6102, Australia
[19] Astrophysics, Department of Physics, Keble Road, Oxford, OX1 3RH, UK
[20] Department of Physics & Astronomy, University of the Western Cape, Private Bag X17, Bellville, Cape Town 7535, South Africa
[21] Institute of Sensors, Signals and Systems, Heriot-Watt University, Edinburgh EH14 4AS, UK
[22] Astrophysics & Cosmology Research Unit, School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Durban 3690, South Africa
[23] Fakultät für Physik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany
[24] Anton Pannekoek Institute for Astronomy, University of Amsterdam, Postbus 94249, 1090 GE Amsterdam, The Netherlands

Received 1 July 2020 / Accepted 10 November 2020

ABSTRACT

The Low Frequency Array (LOFAR) is an ideal instrument to conduct deep extragalactic surveys. It has a large field of view and is sensitive to large-scale and compact emission. It is, however, very challenging to synthesize thermal noise limited maps at full resolution, mainly because of the complexity of the low-frequency sky and the direction dependent effects (phased array beams and ionosphere). In this first paper of a series, we present a new calibration and imaging pipeline that aims at producing high fidelity, high dynamic range images with LOFAR High Band Antenna data, while being computationally efficient and robust against the absorption of unmodeled radio emission. We apply this calibration and imaging strategy to synthesize deep images of the Boötes and Lockman Hole fields at ∼150 MHz, totaling ∼80 and ∼100 h of integration, respectively, and reaching unprecedented noise levels at these low frequencies of $\lesssim$30 and $\lesssim$23 μJy beam^{-1} in the inner ∼3 deg^2. This approach is also being used to reduce the LoTSS-wide data for the second data release.

Key words. techniques: interferometric – techniques: image processing – surveys – galaxies: active – galaxies: starburst – radio continuum: galaxies

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, see van Haarlem et al. 2013) is well suited to conduct deep extragalactic surveys. The LOFAR Surveys Key Science Project is using LOFAR to build tiered extragalactic surveys of different depths and areas and at frequencies ranging from ∼30 to ∼200 MHz. Specifically, the LOFAR LBA Sky Survey (LoLSS) aims at surveying the

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

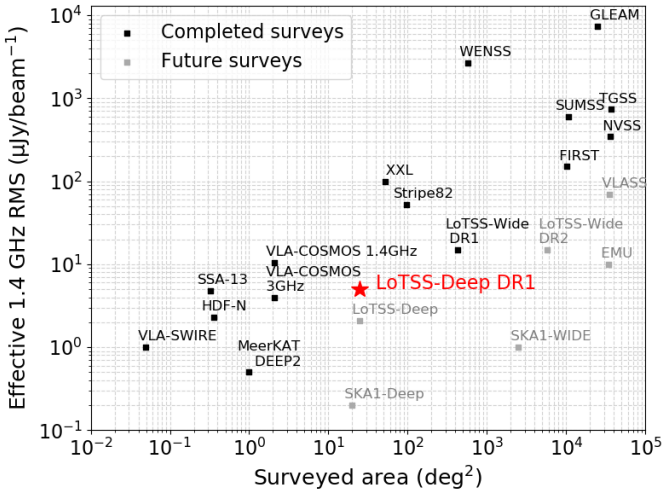

Fig. 1. Effective noise in the LoTSS-Deep Fields continuum maps as compared to other existing and future surveys. A spectral index of $\alpha = -0.7$ has been used to convert flux densities to the 1.4 GHz reference frequency.

northern hemisphere using the LOFAR Low Band Antenna (LBA) antennas, while the LOFAR Two-metre Sky Survey (LoTSS) uses the High Band Antennas (HBAs). Its widest component (LoTSS-wide) has been described by Shimwell et al. (2017, 2019) and aims at reaching noise levels of $\lesssim$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 (see Padovani 2016, and references therein). The LoTSS-Deep Fields targets noise levels of ultimately $\lesssim$10 μJy beam^{-1}, thereby entering a new fainter, higher redshift regime where star forming galaxies and radio quiet AGN will outnumber the population of radio loud AGN, and thereby probing the evolution of those populations over cosmic time. Figure 1 is inspired by that of Smolčić et al. (2017a) and shows a sensitivity and surveyed area comparison between various existing and future surveys. These include the TIFR Giant Metrewave Radio Telescope Sky Survey (TGSS, Intema et al. 2017), the Faint Images of the Radio Sky at Twenty Centimeters (FIRST, Becker et al. 1995), the NRAO Very Large Array Sky Survey (NVSS, Condon et al. 1998), Very Large Array-COSMOS (VLA-COSMOS, Schinnerer et al. 2004; Smolčić et al. 2017b), the Very Large Array Sky Survey (VLASS, Lacy et al. 2020), Evolutionary Map of the Universe (EMU, Norris 2010), the Very Large Array-*Spitzer* Wide-Area Infrared Extragalactic Survey (VLA-SWIRE, Owen & Morrison 2008), SSA13 (Fomalont et al. 2006), Stripe 82 (Heywood et al. 2016), XXL (Butler et al. 2018, and references therein), DEEP2 (Mauch et al. 2020), the LoTSS first data release (DR1) (LoTSS-DR1, Shimwell et al. 2017), *Hubble* Deep Field - North (HDF-N, Richards 2000), the Westerbork Northern Sky Survey (WENSS, Rengelink et al. 1997), the GaLactic and Extra-Galactic All-Sky MWA (GLEAM, Wayth et al. 2015), and the Square Kilometer Array (SKA, Prandoni & Seymour 2015).

The depth of the LoTSS-Deep Fields is unlikely to be routinely surpassed at these low frequencies even into the era of the first phase of the SKA (Dewdney et al. 2009) because, although the SKA will have the raw sensitivity to easily reach such depths, the confusion noise of the SKA-low will likely increase the image rms to values exceeding the target depth of the LoTSS-deep images (see e.g., Zwart et al. 2015; Prandoni & Seymour 2015). In order to construct the LoTSS-Deep Fields, we have selected the Boötes, Lockman Hole, and European Large Area Infrared Space Observatory Survey-North 1 (ELAIS-N1) fields, together with the North Ecliptic Pole (NEP). Each of them is covered by a wealth of multiwavelength data, which are necessary to derive photometric redshifts and low frequency radio luminosities, thereby providing an efficient way to estimate star formation rates (SFRs) in galaxies, for example. Together, these four multiwavelength fields allow us to probe a total sky area of $\gtrsim$30 deg^2 in order to probe all galaxy environments at $z > 1$.

It is, however, quite challenging to make thermal noise limited images at low frequencies because of the presence of direction-dependent effects (DDEs), such as the ionospheric distortions, and because of the complex primary beam shapes of phased arrays. We have shown (Shimwell et al. 2019) that, using a novel set of calibration and imaging algorithms developed by Tasse (2014a), Smirnov & Tasse (2015), and Tasse et al. (2018), we are able to estimate and compensate for the DDEs and thus use LOFAR to produce thermal noise limited maps from eight-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 ~30 μJy beam^{-1} noise levels, while being more robust against absorbing faint unmodeled extended emission and dynamic range issues. In Sect. 2, we introduce the direction-dependent (DD) calibration and imaging problem together with the existing software that we use to tackle it. We describe our DDE calibration and imaging strategy (DDF-PIPELINE-V2) in Sect. 3 (to expand on this, in Appendix A we describe DDF-PIPELINE-V1, which was presented in detail in Shimwell et al. 2019). In Sect. 4, we use DDF-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. 2021), the multiwavelength cross matching (Kondapally et al. 2021), and the photometric redshifts and galaxy characterization (Duncan et al. 2021).

2. LoTSS and the third generation calibration and imaging problem

Calibration and imaging techniques have greatly evolved since the first radio interferometers became operational. First generation calibration is commonly referred to as direction-independent (DI) 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 DDEs.

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 frequencies ($\nu < 250$ MHz), (ii) has very large fields of view (full width at half maximum of 2–10 degrees), and (iii) combines short (~100 m) and long (~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) makes the calibration problem direction-dependent by nature. In Sect. 2.1, we introduce

the mathematical formalism used throughout this paper, while in Sects. 2.2 and 2.3 we describe the algorithms and software used for DI and DD calibration and imaging.

2.1. Measurement equation formalism

The radio interferometry measurement equation (RIME, see Hamaker et al. 1996) describes how the underlying electric field coherence (the sky model), as well as 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 ν in direction $\boldsymbol{s} = [l, m, n = \sqrt{1 - l^2 - m^2}]^T$ (where T is the matrix transpose) and write the 4×1 sky coherency matrix as $\mathbf{x}_{s\nu} = [\mathrm{XX}, \mathrm{XY}, \mathrm{YX}, \mathrm{YY}]^T_{s\nu}$. If $\mathcal{G}_{\boldsymbol{b}} = \mathbf{G}^*_{qt\nu} \otimes \mathbf{G}_{pt\nu}$ and $\mathcal{J}^s_{\boldsymbol{b}} = \mathbf{J}^{s*}_{qt\nu} \otimes \mathbf{J}^s_{pt\nu}$ are the direction-independent and direction-dependent 4×4 Mueller matrices[1] on a baseline $\boldsymbol{b} \leftrightarrow \{pqt\} \rightarrow [u, v, w]^T$ between antenna p and q at time t, then the four-visibility $\mathbf{v}_{\boldsymbol{b}}$ is given by

$$\mathbf{v}_{\boldsymbol{b}} = \mathcal{G}_{\boldsymbol{b}} \int_s \mathcal{J}^s_{\boldsymbol{b}} \mathcal{B}^s_{\boldsymbol{b}} \mathbf{x}_{s\nu} k^s_{\boldsymbol{b}} \mathrm{d}\boldsymbol{s} + \boldsymbol{n}_{\boldsymbol{b}}, \tag{1}$$

$$\text{with } k^s_{\boldsymbol{b}} = \exp\left(-2i\pi \frac{\nu}{c} \left(\boldsymbol{b}^T_{pq,t}(\boldsymbol{s} - \boldsymbol{s}_0)\right)\right) \tag{2}$$

$$\text{and } \boldsymbol{b}_{pq,t} = \begin{bmatrix} u_{pq,t} \\ v_{pq,t} \\ w_{pq,t} \end{bmatrix} = \begin{bmatrix} u_{p,t} \\ v_{p,t} \\ w_{p,t} \end{bmatrix} - \begin{bmatrix} u_{q,t} \\ v_{q,t} \\ w_{q,t} \end{bmatrix} \tag{3}$$

$$\text{and } \boldsymbol{s} = \begin{bmatrix} l \\ m \\ n \end{bmatrix} \text{ and } \boldsymbol{s}_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \tag{4}$$

where c is the speed of light in the vacuum and $\boldsymbol{n}_{\boldsymbol{b}}$ is a 4×1 random matrix following a normal distribution $\mathcal{N}(0, \sigma_{\boldsymbol{b}})$. 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 $\mathcal{G}_{\boldsymbol{b}}$ describe the direction-independent effects, such as the individual station electronics (the bandpass) or their clock drifts and offsets. The $\mathcal{J}^s_{\boldsymbol{b}}$ models the DDEs, including the ionospheric distortion (phase shift, Faraday rotation, and scintillative decoherence) and phased array station beam that depend on time, frequency, and antenna. Importantly, in order to estimate the intrinsic flux densities, we used a description of the LOFAR station primary beam that is built from semi-analytic models[2] and write it as $\mathcal{B}^s_{\boldsymbol{b}}$ 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 $\mathcal{G}_{\boldsymbol{b}}$ and $\mathcal{J}^s_{\boldsymbol{b}}$ and the sky model $\mathbf{x}_\nu$ from the set of visibilities $\mathbf{v}_{\boldsymbol{b}}$. Due to the bilinear structure of the RIME, instead of estimating all these parameters at once, the inversion of Eq. (1) is split into two steps. In the first step, the sky term $\mathbf{x}_\nu$ is assumed to be constant and the Jones matrices are estimated. This step is referred to as "calibration" or as the DD-C-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 $\mathbf{x}_\nu$ is estimated. This step is commonly called "imaging" and is referred to as solving the DD-$\mathcal{I}$-RIME system later in the text. The C-RIME and $\mathcal{I}$-RIME problems constitute two sub-steps in inverting the RIME system. We will later describe the idea of alternating between DD-C-RIME and DD-$\mathcal{I}$-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 DD-C-RIME (DD calibration, see Yatawatta et al. 2008; Kazemi et al. 2011; Tasse 2014a; Smirnov & Tasse 2015) or DD-$\mathcal{I}$-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 DD-C-RIME and DD-$\mathcal{I}$-RIME algorithms to tackle the complete RIME inversion problem.

2.2. Direction-independent calibration

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

The interferometric data taken on the calibrator field are used to estimate the direction-independent Jones matrices $\mathbf{G}$ that are, to the 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[4] (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

$$\widehat{\mathbf{v}^{\mathrm{cal}}_{\boldsymbol{b}}} = \widehat{\mathcal{G}^0_{\boldsymbol{b}}} \mathbf{v}^{\mathrm{model}}_{\boldsymbol{b}}, \tag{5}$$

where $\mathbf{v}^{\mathrm{model}}_{\boldsymbol{b}} = \int_s \mathbf{x}_{s\nu}\, k^s_{\boldsymbol{b}}\, \mathrm{d}\boldsymbol{s}$ is the sky model of the calibrator. We cannot just use $\mathcal{G}^0_{\boldsymbol{b}}$ 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 $\mathcal{G}^0_{\boldsymbol{b}}$ can be decomposed as the product of a direction-independent and direction-dependent term

$$\widehat{\mathcal{G}^0_{\boldsymbol{b}}} = \widehat{\mathcal{G}_{\boldsymbol{b}}} \widehat{\mathcal{J}^0_{\boldsymbol{b}}}, \tag{6}$$

$$\text{with } \widehat{\mathbf{G}_{p\nu}} = \widehat{A_{p\nu}} \exp\left(i\nu \widehat{\Delta^t_p}\right) \tag{7}$$

$$\text{and } \widehat{\mathbf{J}^0_{p\nu}} = \exp\left(iK\nu^{-1} \widehat{\Delta^T_p}\right), \tag{8}$$

where $K = 8.44 \times 10^9$ m^3 s^{-2} and $A_{p\nu}$, Δ^t_p, and Δ^T_p are real-valued and represent, respectively, the bandpass, the clock, and the ionospheric total electron content (TEC) offset with respect to a reference antenna. The terms Δ^t_p and Δ^T_p can be disentangled from the frequency-dependent phases because the former is linear with ν while the latter is nonlinear.

Assuming the clocks and bandpass are the same for the calibrator and for the target field, the corrected visibilities $\mathbf{v}^{\mathrm{c}}_{\boldsymbol{b}}$ for the target field can be built from the raw data $\mathbf{v}_{\boldsymbol{b}}$ as $\mathbf{v}^{\mathrm{c}}_{\boldsymbol{b}} = \widehat{\mathcal{G}_{\boldsymbol{b}}}^{-1} \mathbf{v}_{\boldsymbol{b}}$. In order to calibrate for the remaining phase errors, the target field

[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 $\mathbf{J}_p$ and $\mathbf{J}_q$ are 2×2 Jones matrices of antenna p and q, and $\mathbf{X}$ is the 2×2 source's coherency matrix, then we have $\mathrm{Vec}\left\{\mathbf{J}_p \mathbf{X} \mathbf{J}_q\right\} = \left(\mathbf{J}^*_q \otimes \mathbf{J}_p\right) \mathrm{Vec}\{\mathbf{X}\}$, 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

is DI calibrated against the TGSS catalogs (Intema et al. 2017) and visibilities 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 $\mathbf{G}^s_{ptv}$ and $\mathbf{J}^s_{ptv}$. Depending on the observing frequency and instrumental setup, these can be the product of the ionospheric Faraday rotation matrix, the scalar phase shift, or the individual station primary beams. This approach has the great advantage of constraining the free parameters to a low number, but it requires one to analytically model the physics of the various effects over the $\{sptv\}$ space to be able to disentangle them. The second family of algorithms directly estimates the effective Jones matrices over piecewise constant domains in $\{sptv\}$ space. These "Jones-based" solvers have the advantage of not requiring any physical modeling of the DDEs. However, this makes the number of degrees of freedom increase dramatically, typically by a few orders of magnitude. These additional degrees of freedom can often make the inverse problem ill-posed[5]. This means, in practice, that the DD solvers can overfit the data, leading to the unmodeled 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, as well as on the shape of the neighboring domains in the $\{sptv\}$ spaces. Furthermore, 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 DD 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 $\mathbf{x}_{sv}$ for given $\mathcal{G}$, $\mathcal{J}$, and $\mathbf{v}$ 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). While the true measured visibilities are described by Eq. (1), the ("model") visibilities $\widehat{\mathbf{v}_b}$ that are estimated[6] over the piecewise constant domains $p, \Omega_\varphi, \Delta t$, and $\Delta\nu$ can be written as

$$\widehat{\mathbf{v}_b} = \widehat{\mathcal{G}_b}\widehat{\mathbf{v}^\Sigma_b}, \tag{9}$$

$$\text{with } \widehat{\mathbf{v}^\Sigma_b} = \sum_\varphi \widehat{\mathbf{v}^\varphi_b} \tag{10}$$

$$\text{and } \widehat{\mathbf{v}^\varphi_b} = \widehat{\mathcal{J}^\varphi_b}\mathcal{B}^\varphi_b \int_{s\in\Omega_\varphi} \widehat{\mathbf{x}_{sv}}\, k^s_b \,\mathrm{d}\boldsymbol{s}, \tag{11}$$

where Ω_φ is the set of directions $\boldsymbol{s}$ for a facet φ, $\widehat{\mathbf{x}_{sv}}$ is the estimated underlying sky, and $\widehat{\mathcal{G}_b}$ and $\widehat{\mathcal{J}^\varphi_b}$ are the DI and DD Mueller matrices for baseline $\boldsymbol{b}$, built from the corresponding estimated Jones matrices in $p, \Delta t$, and $\Delta\nu$.

Specifically, in order to solve for the DDEs, 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 nonlinear nature of Eq. (1) and (ii) the complexity of the background radio sky, optimizing 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 $\widehat{\mathbf{x}_{sv}}$ as well as the DI $\widehat{\mathbf{G}_{ptv}}$ and the DD piecewise constant $\widehat{\mathbf{J}^\varphi_{ptv}}$ for all $\{s\varphi ptv\}$ such that $\widehat{\mathbf{v}_b} \sim \mathbf{v}_b$ (we remain intentionally vague here since the cost function that is minimized depends on the specific DD algorithm). In practice, inverting Eq. (1) (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 sky model and repeating steps (i) and (ii) is known as self-calibration, but in a third-generation approach we must explicitly model the direction-dependent aspects.

Since the computing time evolves as $\sim n_\mathrm{d}^3$, where n_d is the number of directions in the DD solvers, the problem of DD calibration has in general been tackled using direction alternating peeling-like techniques. Major breakthroughs have been made in the field of DD-$\mathcal{C}$-RIME solvers over the past decade by Yatawatta et al. (2008); Kazemi et al. (2011), making this DD 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 different structures, and shortcuts can be taken to invert the calibration problem. The net gain over the classical method is not trivial: It can be as high as n_a^2 (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 other Jones-based solver.

Only a very few $\mathcal{CI}$-RIME algorithms using a full DD self-calibration loop have been described and implemented. They include pointing self-calibration (Bhatnagar & Cornwell 2017) and peeling-based techniques such as MFIMAGE (implemented in the OBIT[7] package) and FACTOR (van Weeren et al. 2016, see also Sect. 4.3). 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, with the exception of calibration sources, in that one facet and (ii) DI self-calibrating in that direction. In addition to the ill-conditioning issues discussed above regarding DD-$\mathcal{C}$-RIME and DD-$\mathcal{I}$-RIME solvers, an expensive computational problem arises when estimating the $\widehat{\mathbf{J}^\varphi_{ptv}}$.

The approach presented by Shimwell et al. (2019; also described in detail in Appendix A and referred to as DDF-PIPELINE-V1 in the following) is based on the KMS DD-$\mathcal{C}$-RIME solver (Tasse 2014a; Smirnov & Tasse 2015) and DDFacet DD-$\mathcal{I}$-RIME imager (Tasse et al. 2018) and is algebraically simultaneous in directions. The direction-dependent pipeline DDF-PIPELINE[8] is a high level wrapper that mainly calls DDFacet[9] and KMS[10] for DD 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-$\mathcal{C}$-RIME solver, that is to say, the DD affected sidelobes leaking from any facet to any other facet are accounted for within the algebraic operations of the algorithm. Another advantage over the FACTOR approach is that the data only need to be read, as

[5] Linear and nonlinear problems can be ill-posed, meaning that nonunique solutions can be found.

[6] Throughout this paper, the notation $\widehat{x}$ should be read as "the estimate of x."

[7] https://www.cv.nrao.edu/~bcotton/Obit.html

[8] https://github.com/mhardcastle/ddf-pipeline

[9] https://github.com/saopicc/DDFacet

[10] https://github.com/saopicc/killMS

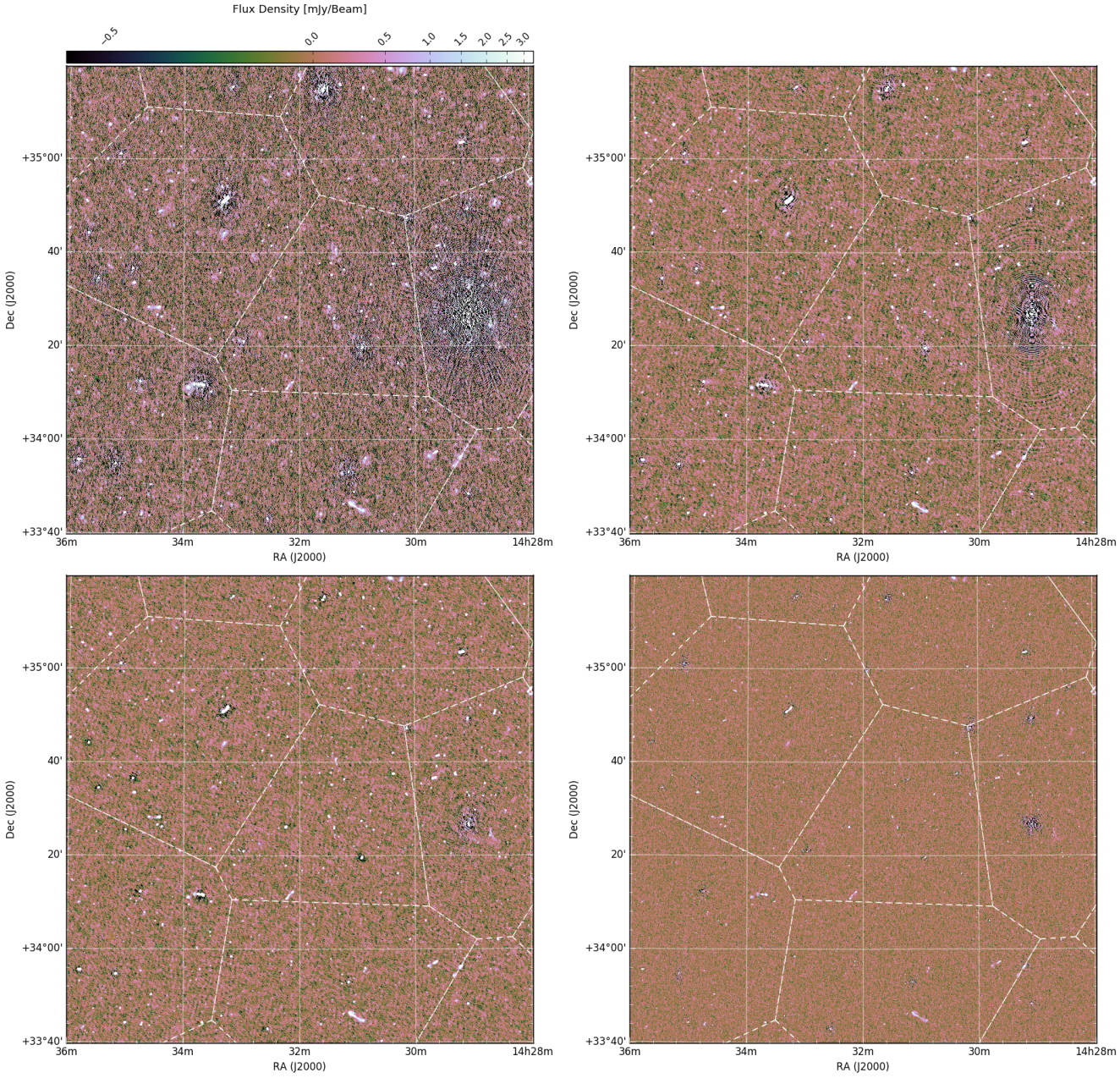

Fig. 2. Some of the images produced at different steps in the DD self-calibration loop implemented as Algorithm 1. The maps correspond (*from left to right, top to bottom*) to Steps 1.1, 1.8, 1.12, and 1.18, respectively. The white lines show the facet locations. The color scale is the same on all panels and is displayed using an inverse hyperbolic sine function to render both the low level artifacts and some bright sources.

opposed to modified, making the DDF-PIPELINE more efficient in terms of memory and disk access.

As explained in Shimwell et al. (2019), the DDF-PIPELINE-V1 is, however, affected by a number of issues. These include (i) artifacts and dynamic range limits around the brightest radio sources, (ii) artificial and diffuse haloes around moderately bright radio sources, and (iii) unmodeled flux absorption.

3. Calibration and imaging robustness

In this section, we describe in detail a DD calibration and imaging algorithm that aims to make the overall RIME imaging and calibration solver more robust against artifacts around the brightest sources (Sect. 3.1) and unmodeled flux absorption (Sects. 3.2 and 3.3). An overview of this approach is shown in Algorithm 1 (the implementation of which is referred to as DDF-PIPELINE-V2), and the corresponding DD self-calibration loop is displayed in Fig. 2. The DDF-PIPELINE-V2 products are described in Sect. 3.4. In Sect. 3.5, we show that DDF-PIPELINE-V2 produces improved images as compared to those previously made with DDF-PIPELINE-V1 (see Appendix A and Shimwell et al. 2019). In Sect. 3.6, we discuss the DDF-PIPELINE-V2 computing efficiency.

3.1. Dynamic range issue

With LOFAR's very large field of view, it is common to observe bright sources within the station's primary beam. Tests we conducted on fields containing bright sources (such as 3C 295, which has a flux density of ∼100 Jy at 150 MHz) have shown

Algorithm 1: Overview of the algorithm implemented in DDF-PIPELINE-V2. The function $\mathcal{I}$ represents the imaging step and takes as input the visibility vector $\mathbf{v}$ together with the beam model $\mathbf{B}_{\Omega_n}$ and kMS-estimated Jones matrices $\mathbf{J}_{\Omega_n}$ at locations Ω_n. The function $\mathcal{K}$ abstracts the DD calibration step, and takes as arguments the visibilities $\mathbf{v}$, the skymodel $\widehat{\mathbf{x}_\nu}$, 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 Ω_n in which to solve for. The extra functions $\mathcal{C}$, $\mathcal{B}$, and $\mathcal{F}$ represent the clustering, bootstrapping and smoothing steps respectively.

Data: Visibilities $\mathbf{v}$ calibrated from DI effects using PREFACTOR.

/* **On 60 LOFAR HBA subbands** */
/* DI initial deconv and clustering */
1.1 $\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}(\mathbf{v}_6, \mathbf{J}_{\Omega_r} = \mathbf{1}, \mathbf{B}_{\Omega_r})$;
1.2 $\Omega_n \leftarrow \mathcal{C}(\widehat{\mathbf{x}_\nu})$;
/* DI calibration and imaging */
1.3 $\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}(\mathbf{v}_6, \mathbf{B}_{\Omega_n})$;
1.4 $\mathbf{v}_6^c \leftarrow \mathcal{K}(\mathbf{v}_6, \widehat{\mathbf{x}_\nu}, \mathbf{B}_{\Omega_n}|\text{full}, \delta t_0, \delta\nu_0, \Omega_0)$;
1.5 $\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}\left(\mathbf{v}_6^c, \mathbf{B}_{\Omega_n}\right)$;
/* Bootstrapping the flux density scale */
1.6 $\mathbf{v}^c \leftarrow \mathcal{B}\left(\mathbf{v}_6^c\right)$;
/* Phase only DD calibration and imaging */
1.7 $\widehat{\mathbf{J}} \leftarrow \varphi \circ \mathcal{F} \circ \mathcal{K}\left(\mathbf{v}_6^c, \widehat{\mathbf{x}_\nu}, \mathbf{B}_{\Omega_n}|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_n\right)$;
1.8 $\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}\left(\mathbf{v}_6^c, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}\right)$;
/* DD calibration and imaging */
1.9 $\widehat{\mathbf{J}} \leftarrow \mathcal{F} \circ \mathcal{K}\left(\mathbf{v}_6^c, \mathbf{B}_{\Omega_n}, \widehat{\mathbf{x}_\nu}|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_n\right)$;
1.10 $\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}\left(\mathbf{v}_6, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}\right)$;
/* DI calibration and imaging */
1.11 $\mathbf{v}_6^c \leftarrow \mathcal{K}\left(\mathbf{v}_6, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}, \widehat{\mathbf{x}_\nu}|\text{full}, \delta t_0, \delta\nu_0, \Omega_0\right)$;
1.12 $\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}\left(\mathbf{v}_6^c, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}\right)$;

/* **On 240 LOFAR HBA subbands** */
/* DD calibration */
1.13 $\widehat{\mathbf{J}} \leftarrow \mathcal{F} \circ \mathcal{K}(\mathbf{v}_{24}, \mathbf{B}_{\Omega_n}, \widehat{\mathbf{x}_\nu}|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_n)$;
/* DI calibration */
1.14 $\mathbf{v}_{24}^c \leftarrow \mathcal{K}\left(\mathbf{v}_{24}, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}, \widehat{\mathbf{x}_\nu}|\text{full}, \delta t_0, \delta\nu_0, \Omega_0\right)$;
/* DD imaging */
1.15 $\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}\left(\mathbf{v}_{24}^c, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}\right)$;
/* DD calibration */
1.16 $\widehat{\mathbf{J}} \leftarrow \mathcal{F} \circ \mathcal{K}\left(\mathbf{v}_{24}^c, \mathbf{B}_{\Omega_n}, \widehat{\mathbf{x}_\nu}|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_n\right)$;
/* Slow DD calibration */
1.17 $\widehat{\mathbf{J}_s} \leftarrow \mathcal{K}\left(\mathbf{v}_{24}^c, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}, \widehat{\mathbf{x}_\nu}|\text{scalar}, 43\text{min}, 2\text{MHz}, \Omega_n\right)$;
/* Final imaging steps */
1.18 $\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}\left(\mathbf{v}_{24}^c, \widehat{\mathbf{J}_s}\widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}\right)$;
1.19 Facet-based astrometric correction (see Shimwell et al. 2019, for details);

that the related residual errors create powerful artifacts that largely dominate the thermal noise, thereby driving a dynamic range limit. As explained in Sect. 2.2, the initial phase calibration is done against TGSS at 150 MHz (Intema et al. 2017).

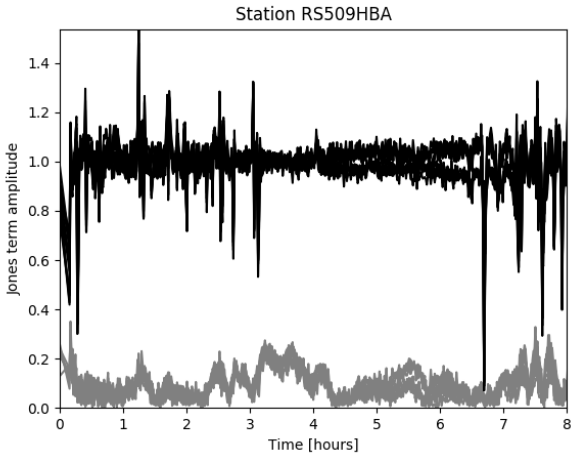

Fig. 3. Amplitude of the diagonal (black) and off-diagonal (gray) terms of the estimated Jones matrices for a remote station using the LOFAR observation synthesized 6″ image as the sky model (self-calibration). If the initial DI calibration and correction by PREFACTOR on the lower resolution TGSS sky model were good enough, the calibration solutions found here would be the unity matrix at all times and frequencies. Therefore, in DDF-PIPELINE-V2, we carry out a few full-Jones DI-only self-calibration steps. This approach seems very efficient in increasing the dynamic range around the brightest radio sources.

However, since LOTSS resolution is much higher than that of TGSS (6″ × 6″vs. 25″ × 25″, respectively), small spatial uncertainties on how the individual bright sources are modeled lead to large Jones matrix errors. Specifically, this effect can be severe when the true point sources are erroneously found to be resolved by TGSS as this leads to large calibration errors for the long baselines. In these situations, our experience shows that the initial DI calibration is not good enough to start doing a DD calibration (due to ill-conditioning issues, it has to be done on larger time-frequency solution intervals). In the following, we study the DI calibration solutions and assess whether they need to be recomputed using a high angular resolution sky model. When using the PREFACTOR DI-calibrated LOTSS data and associated imaging products, this amounts to doing a round of DI self-calibration at the beginning of DDF-PIPELINE-V2.

Solution time and frequency variability is, however, hard to interpret. Indeed, because the RIME 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, they are given with respect to a reference antenna. When Jones matrices are scalar, this amounts to zeroing the phases φ_0 of the reference antenna by subtracting φ_0 from all phases of all antennas. To do this in the general case of non-diagonal Jones matrices, we used a polar decomposition on the Jones matrix $\mathbf{J}_0$ of the reference antenna such that $\mathbf{J}_0 = \mathbf{U}\mathbf{P}_0$, where $\mathbf{U}$ is a unitary matrix[11]. We then applied $\mathbf{U}$ to all Jones matrices as $\mathbf{J}_p \leftarrow \mathbf{U}^H\mathbf{J}_p$. Intuitively, when the Jones matrices are all scalar, the unitary matrix $\mathbf{U}$ is simply $\exp(i\varphi_0)\mathbf{I}$, and that step makes the phases of all $\mathbf{J}_p$ relative to the reference antenna (and specifically zeros the phases of $\mathbf{J}_0$). In the case of nontrivial 2×2 Jones matrices, finding and applying $\mathbf{U}$ has the effect of removing a common rotation from all Jones matrices, and it orthogonalizes them.

We apply this in Fig. 3, where we show the typical DI Jones matrices we can estimate at Step 1.3 for a given remote station

[11] The unitary matrix $\mathbf{U}$ is found by doing a singular value decomposition $\mathbf{J}_0 = \mathbf{W}\mathbf{\Sigma}\mathbf{V}$ and is then built as $\mathbf{U} = \mathbf{W}\mathbf{V}^H$.

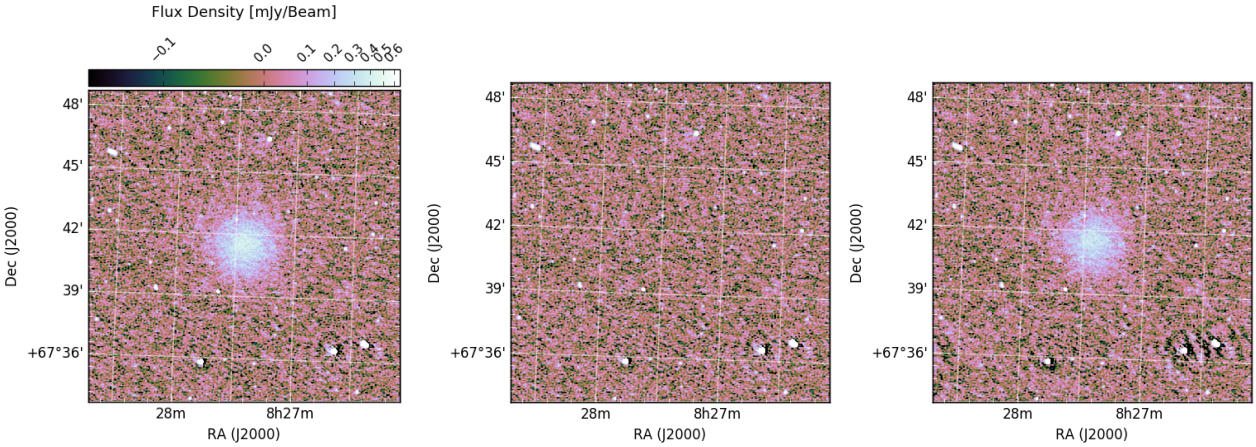

Fig. 4. Simulations conducted to test the robustness of the algorithm described in Sect. 3 and implemented in DDF-PIPELINE-V2. We have simulated an unmodeled extended emission (*left-hand panel*). The emission is absorbed by the DD-calibration step (*central panel*), while it can be partially recovered (*right-hand panel*) by decreasing the effective size of the unknown solutions space (Sects. 3.2 and 3.3).

and frequency, with respect to a reference station in the LOFAR core. They were estimated using kMS and the sky model synthesized by DDFacet from the visibilities corrected by PREFACTOR. 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 polarization 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 the modeled visibilities generated by DDFacet in Step 1.3. As the sky is mostly unpolarized, we assumed $Q = U = V = 0$ Jy in this full-Jones DI calibration step (see Sect. 3.4 for a discussion of polarization related data products). The solution intervals δt_0 and $\delta \nu_0$ along time and frequency are determined such that $n_b \propto (T/\langle|\boldsymbol{x}_\nu|\rangle)^2 \mathrm{Var}\{\mathbf{n}\}$, where n_b is the number of points in the $\delta t_0 \times \delta \nu_0$ time-frequency domain, T is the target solution S/N, and $\mathrm{Var}\{\mathbf{n}\}$ is the variance of the visibilities' noise (see Sob et al. 2020 for a justification).

We note that after the initial DD calibration solutions are obtained in Steps 1.7 and 1.9, a more accurate DI calibration can be performed. Specifically, in the DI calibration Steps 1.11 and 1.14, the model visibilities $\widehat{\mathbf{v}_{\boldsymbol{b}}^{\Sigma}}$ (Eq. (10)) are predicted on any baseline $\boldsymbol{b}$ based on the previously estimated DD-Jones matrices $\widehat{\mathbf{J}}$ (as is done by Smirnov 2011).

3.2. Regularization

The absorption of unmodeled flux by calibration is a well-known issue connected to the calibration of DDEs. Intuitively speaking, when real flux is missing from the modeled sky $\widehat{\mathbf{x}_i}$ of $\mathbf{x}$ at step i, and since the RIME inversion is often ill-posed, the estimates $\widehat{\mathbf{J}_i}$ of $\mathbf{J}$ can be biased in a systematic way. Experience and simulations show that building a new estimate $\widehat{\mathbf{x}_{i+1}}$ from $\widehat{\mathbf{J}_i}$ can be biased in that the unmodeled emission is not and never will be recovered (Fig. 4). This effect is especially severe when the extended emission is poorly modeled or unmodeled since it is detected only by the shortest baselines. Effectively, during the inversion of the RIME system of equations, the DD self-calibration algorithm falls into the wrong (local) minimum.

One way to address this problem is to reduce the effective number of free parameters used to describe the Jones matrices in the $\{pdt\nu\}$-space (see for example Tasse 2014a; Yatawatta 2015; van Weeren et al. 2016; Repetti et al. 2017; Birdi et al. 2020). Forcing the estimated Jones matrix shape to look like that of the real underlying ones improves the conditioning of the inverse problem. In Algorithm 1 (implemented in DDF-PIPELINE-V2), we replaced that normalization method with a smoothing of the kMS-estimated Jones matrices. This function $\mathcal{F}$ updates the Jones matrices $\mathbf{J} \leftarrow \mathcal{F}(\mathbf{J})$ in each direction independently by imposing on them a certain behavior in the time-frequency space (see below), effectively reducing the size of the unknown stochastic process. This can be thought of as a regularization. This is done independently on the phases and amplitudes on the scalar Jones matrices generated in Steps 1.9, 1.13, 1.16. The updated Jones matrices take the analytical form

$$\widehat{\mathbf{J}_{pd,t\nu}} = \widehat{a_{pd,t}}\mathcal{P}\left(t, \widehat{\boldsymbol{\theta}_{pd,\nu}}\right)\exp\left(iK\nu^{-1}\widehat{\Delta_p^T}\right)\mathbf{I}, \tag{12}$$

where $\widehat{\Delta_p^T}$ is the differential TEC (see also Sect. 2.2 and Eq. (8)), $\mathcal{P}$ is a polynomial parametrized by the coefficients in $\boldsymbol{\theta}_{pd,\nu}$ (of size 10), and $\widehat{a_{pd,t}}$ is a scalar meant to describe the loss of correlation due to ionospheric scintillation, as seen in the top left-hand panel of Fig. 5. Typically, for the ∼8-h integration of LOTSS pointings and solving every 30 s and 2 MHz, this parametrization of the Jones matrices reduces the number of free parameters by a factor of $\gtrsim$20.

To assess the recovery of unmodeled 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 haloes of galaxy clusters. After the injection of the artificial extended sources, Steps 1.16 and 1.17 were repeated using the sky model derived in Step 1.18 prior to the injection of the sources. These simulations will be discussed further by Shimwell et al. (in prep.), but briefly, in each simulation the recovered flux of the completely unmodeled emission exceeded 60%. Examples of the injected and recovered emission are shown in Fig 4.

As suggested by the results of the simulations, the effect on real data is in general very satisfactory and allows us to recover a good fraction of 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′

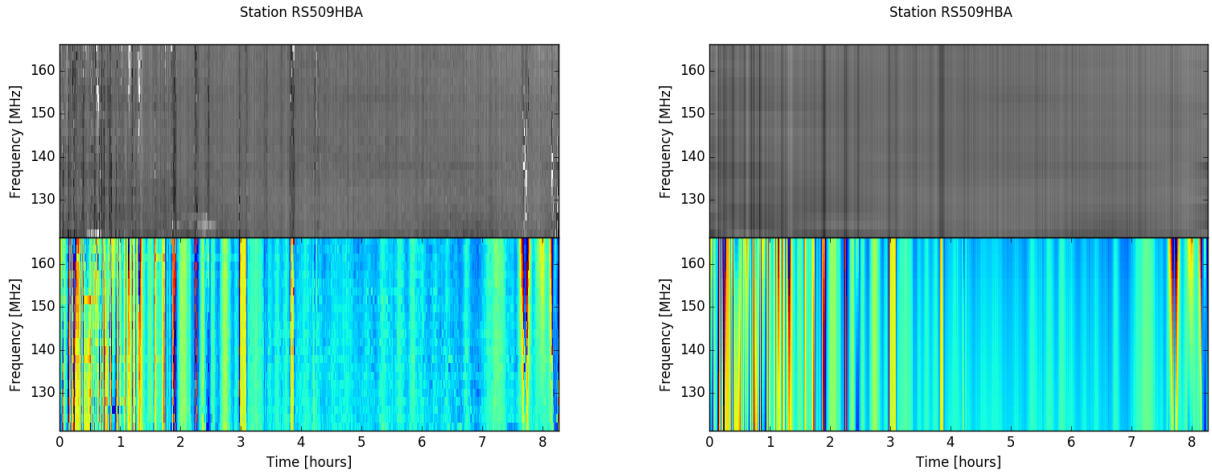

Fig. 5. Amplitude and phase (*top and bottom*, respectively) of a scalar Jones matrix for a given station in a given direction in the example observation. *Left-hand panel*: solution as estimated by the kMS solver. *Right-hand panel*: regularized solution as updated by the $\mathcal{F}$ function. The amplitude color scale ranges from 0 to 1.5.

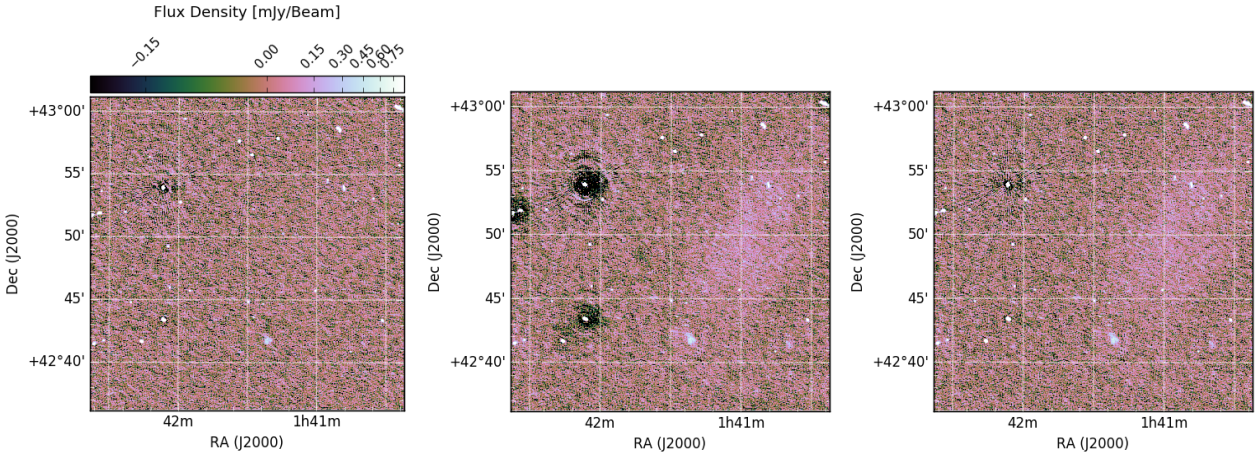

Fig. 6. Effects of two calibration steps on imaging artifacts and flux absorption. *Left-hand panel*: shows that a faint and unmodeled extended emission (on the level of $\sim 0.7\sigma$ here) can be totally absorbed. While regularizing the DD calibration solutions can help in recovering the unmodeled emission (typically after Step 1.16), it can also produce negative imaging artifacts and "holes" around bright sources (*central panel*). *Right-hand panel*: shows that solving the residuals on longer time intervals (Step 1.17) corrects for this issue.

across, with a mean flux density at the peak of only 70% of the local standard deviation.

3.3. Conditioning and solution interval

The additional issue of arcminute-scale negative haloes appearing around bright compact sources (at a level of $\lesssim$1% or the peak) could, however, be seen in ∼10−20% of the LoTSS pointings processed with DDF-PIPELINE-V1. As shown in Fig. 6, we believe this to be connected to the solution regularization itself. This issue is hard to understand in detail because of the non-linearity in the *C*-RIME inversion, but it is likely due to the pointings that show these issues being more severely affected by the incompleteness of the sky model. Specifically, by conducting several experiments we were able to observe that the situation was improved by deconvolving more deeply or by 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 DD calibration steps presented in Algorithm 1, we used solution intervals of 0.5−1 min. Following van Weeren et al. (2016), we added an extra calibration, Step 1.17, where the visibilities are modeled using the latest available sky model together with the smoothed Jones matrices estimated in Step 1.16 and which are defined over the finer time and frequency mesh. Intuitively, since the negative haloes are produced by some systematic effects, the idea is to calibrate for a slowly varying differential effect. The time interval is set to ∼43 min in DDF-PIPELINE-V2, giving 11 solution intervals in the eight-hour LoTSS pointings. This interval has to be long enough to reach a good conditioning for the *C*-RIME inversion and short enough to sample the Jones matrices' remaining physical variations. As shown in the right-hand panel of Fig. 6, this method is very efficient in reducing the negative haloes.

3.4. Data products

3.4.1. Unpolarized flux

Once the estimated DD-Jones matrices and sky model $\widehat{\mathbf{x}_\nu}$ have been obtained at the highest available spatial resolution following the DI and DD self-calibration steps presented in Algorithm 1, additional data products are formed. Users can

adapt the weighting scheme depending on the scientific exploitation of the interferometric data they want to make use of. This is very much tied to how the calibration and deconvolution algorithms work, 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 nonzero). To tackle this issue, the PSF can be modified to make the convolution matrix more diagonal and the deconvolution problem correspondingly better conditioned. This is done at the cost of a lower sensitivity, which in turn can drive systematic errors in the calibration solution estimates because extended emission is poorly modeled on the shorter baselines.

For all these concurring reasons, the faint and extended flux in the highest resolution maps produced by Algorithm 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 degraded the resolution of some of the imaging to allow survey users to choose a resolution based on the broad scientific topic that they need to address. Additionally, we stored the sub-space deconvolution (SSD, see Tasse et al. 2018) masks as residual images so the end user can know if any given source has been deconvolved. With this in mind, the following Stokes I products are generated: (i) a high resolution (6″, 20.000 × 20.000 pixels) wide-bandwidth Stokes I image (Step 1.18), (ii) a low resolution (20″) wide-bandwidth Stokes I image (Step 2b.1), and (iii) a high resolution (6″) Stokes I image in three frequency chunks spread over the whole HBA bandwidth (Step 2b.2).

The DI-calibrated visibilities as well as the final sky models and DD calibration solutions are stored. This allows for additional postprocessing to be made, such as better calibration toward a particular point on the sky (vanWeeren et al. 2020), as well as for reimaging at different resolutions if required.

3.4.2. QUV images

The DDFacet DD-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 DD imaging due to the leakage terms. Indeed, for the problem to be properly addressed, 16 PSFs have to be computed (as there are 16 terms in the quadratic mean of the Mueller matrices). As most of the sources are unpolarized, the leakage terms are properly taken into account in the DD predict (i.e., the forward mapping from sky and Jones matrix estimates to modeled visibilities). Instead of deconvolving the polarized flux, we grid the IQUV residual data. The polarized flux is directly interpretable when the sources are unresolved. Hence, we also generated the following additional products: (i) low resolution (20″) spectral Stokes QU cubes (480 planes - Step 2b.3), (ii) very low resolution Stokes QU cubes (480 planes - Step 2b.4), by cutting the baselines >1.6 km, giving an effective resolution of ∼ 3′, and (iii) a 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 polarized sources and their RMs 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 ∼1.1 rad m^{-2} and an ability to measure RMs of up to ∼450 rad m^{-2} (e.g. O'Sullivan et al. 2020).

The 3′ QU cubes are sensitive to the large-scale polarized emission from the Milky Way, while the 20″ QU cubes are excellent for finding compact polarized sources. However, detailed studies of the polarization and RM structure of resolved extragalactic sources will require the deconvolution of the Q and U data. The DDF-PIPELINE-V2 output provides significantly better performance when correcting for the effect of the instrumental polarization (Fig. 7), which is typically at the level of 1% or less for bright total intensity sources (O'Sullivan et al., in prep.).

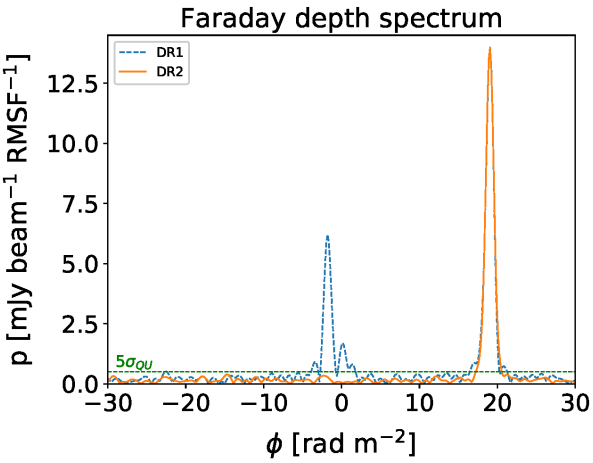

Fig. 7. Faraday depth spectrum, or Faraday dispersion function (FDF), for a radio galaxy in both the DR1 and DR2 datasets, showing the improvement in the suppression of the instrumental polarization signal. The dashed blue line shows the FDF from the DR1 data with a strong instrumental polarization feature near Faraday depths of $\phi \sim 0$ rad m^{-2}, while the solid orange line shows the FDF from the DR2 data in which the instrumental feature is suppressed below the noise level. In both cases, the Faraday depth of the real astrophysical signal is the same.

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

Bright polarized sources are rare in the LoTSS data, with only three sources having a polarized 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 polarized sources, the DDF-PIPELINE-V2 output becomes unreliable for polarized sources. This limitation likely arises from assuming $Q = U = V = 0$ Jy for a field in the DI 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 polarized sources are inserted into existing LoTSS uv-datasets. Possible solutions will be tested in future pipeline developments.

3.5. Comparison between DDF-PIPELINE-V1 and DDF-PIPELINE-V2

Figure 8 shows the comparison between the final high resolution images produced by DDF-PIPELINE-V1 and DDF-PIPELINE-V2 for an eight-hour integration LoTSS pointing (P26HETDEX03). Many processes are involved in the sky reconstruction from radio interferometric data. Imaging and calibration affect the final synthesized maps and introduce complex and systematic residual

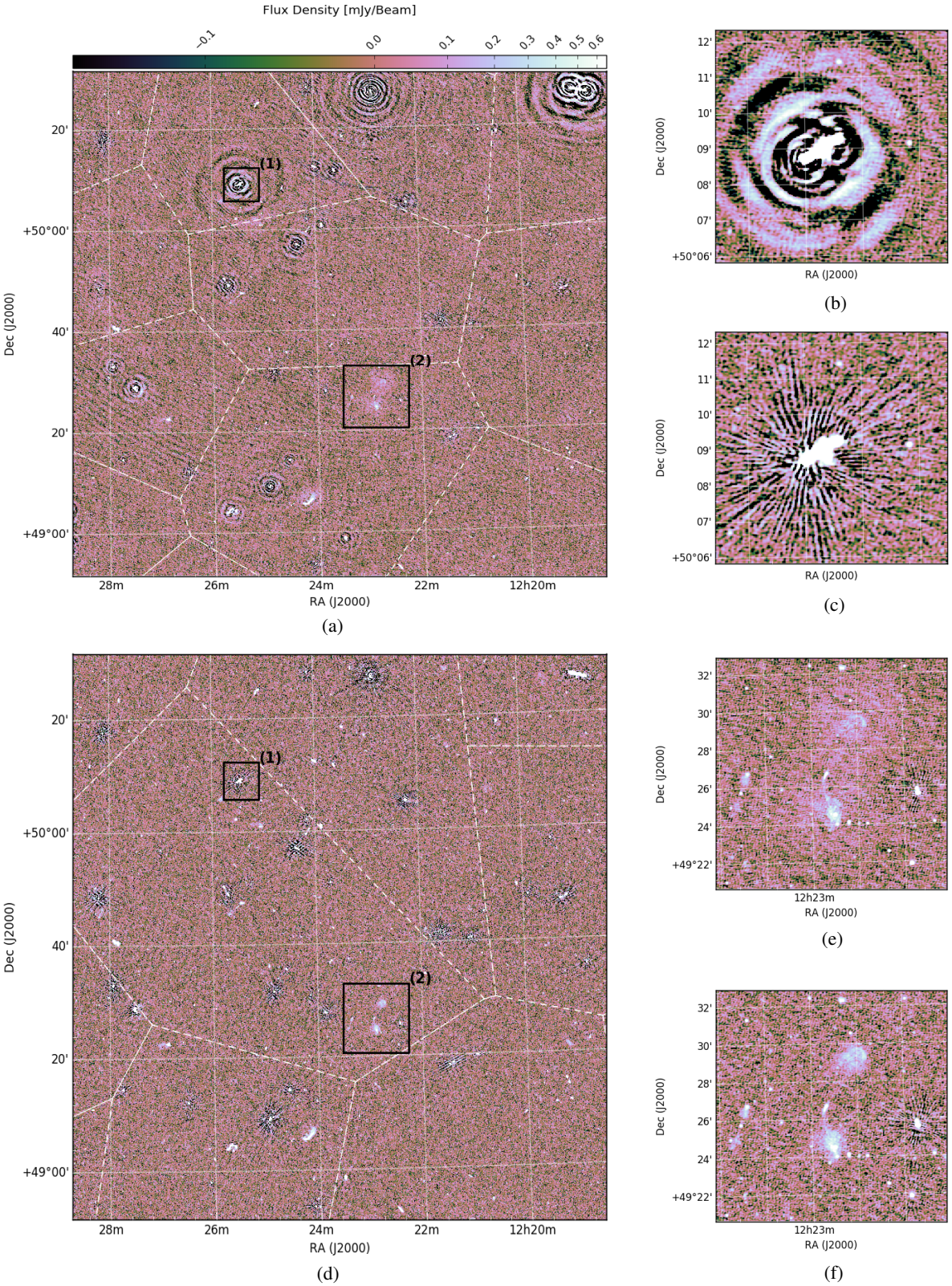

Fig. 8. Differences between the maps produced by Algorithms 0 and 1 from typical eight-hour scans (here the P26HETDEX03 pointing in the Hobby-Eberly Telescope Dark Energy Experiment, HETDEX, field, see Shimwell et al. 2017). The color scale is the same on all panels and is displayed using an inverse hyperbolic sine function to render both the low level artifacts and some bright sources. (*a*) The central part of the P26Hetdex03 8 hours LOFAR-HBA scan as imaged by Alg. 0. (*b*) Region (1) as imaged by Alg. 0. (*c*) Region (1) as imaged by Algorithm 1. (*d*) The central part of the P26Hetdex03 eight-hour LOFAR-HBA scan as imaged by Algorithm 1. (*e*) Region (2) as imaged by Algorithm 0. (*f*) Region (2) as imaged by Algorithm 1.

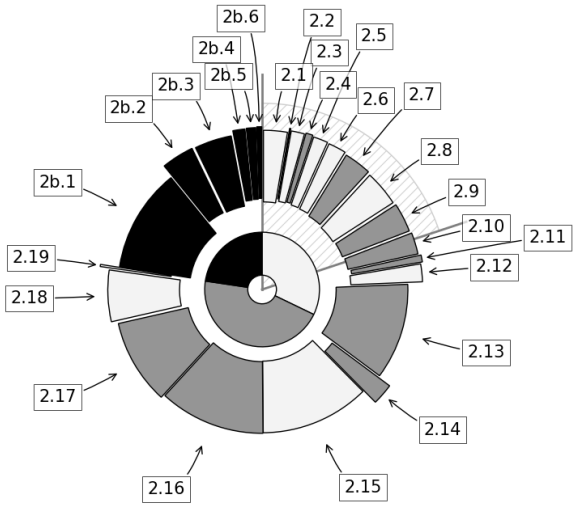

Fig. 9. Nature and ordering of the different steps of Algorithm 1 and how the computing time is distributed across them. The light and dark gray areas represent the imaging and calibration steps, respectively. The black area represents the miscellaneous tasks (additional data products, see Sect. 3.4) that are done once the DI and DD self-calibration loops have completed. These graphs were created from a DDF-PIPELINE-V2 run on a node equipped with 192 GBytes RAM and 2 Intel Xeon Gold 6130 CPU at 2.10 GHz, giving 32 physical compute cores. The dashed area is a quadrant representing a day, while the inner pie shows the total contributions of the imaging, calibration, and miscellaneous tasks.

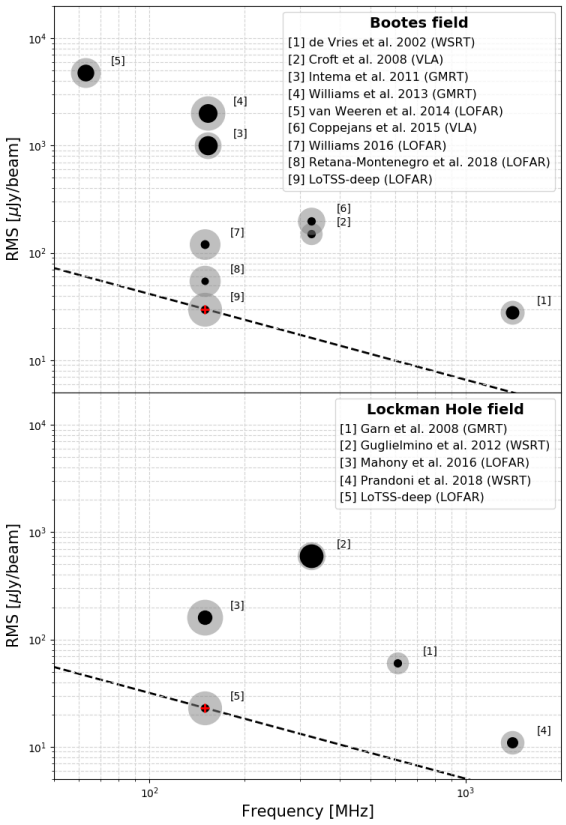

Fig. 10. Sensitivity of the various deep dedicated surveys covering the Boötes (*top*) and Lockman Hole (*bottom*) fields as a function of observing frequency. The resolution of the various surveys corresponds to the radius of the black dot, while the diameter of the corresponding surveyed area is encoded in the size of the gray circle. The LoTSS-Deep Fields pointings are marked with a red cross, and the dashed line corresponds to a source that has a spectral index of −0.7.

errors. It is therefore difficult to find a good and absolute metric to compare the final imaging products.

As discussed in Sect. 3.1, the quality of the initial DI calibration proved to be quite crucial for the feasibility of the following DD calibration and imaging steps. DDF-PIPELINE-V1 was indeed failing at imaging certain fields with very bright sources, while artifacts were present around most moderately bright ones thereby driving the dynamic range limit in large areas. In Figs. 8b and 8c, we show a radio source imaged by DDF-PIPELINE-V1 and DDF-PIPELINE-V2.

Another important issue with the approach we presented in Shimwell et al. (2019) was the presence of a low spatial frequency pattern corresponding to a positive or negative halo around radio sources. Although the effect is complex to analyze, we concluded from various experiments that these systematics were due to the combination of (i) sky model incompleteness, (ii) a *uv*-distance cut used during the calibration, and (iii) the $\mathcal{N}$ normalization function (see Appendix A for details) that we had introduced for DDF-PIPELINE-V1 to be robust against the absorption of extended extended emission. As shown in Figs. 8e and 8f, the approach developed in Sects. 3.2 and 3.3 to conserve unmodeled extended emission and implemented in DDF-PIPELINE-V2 does not produce any significant low spatial frequency systematics.

3.6. DDF-PIPELINE-V2 robustness and performance

As explained above, DDF-PIPELINE-V2 is a high level script that interfaces kMS and DDFacet. Both of the underlying software packages are efficiently parallelized using a custom version of the Python `multiprocessing` package for process-level parallelism and using the SharedArray[12] module. As explained by Tasse et al. (2018), this pythonic approach minimizes the process interconnections for both the kMS and DDFacet 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 1500 h every six-month cycle, which corresponds to approximately two eight-hour pointings (observed simultaneously) each day. The DDF-PIPELINE-V2 compute time is roughly split equally between calibration and imaging tasks (see Fig. 9). The total run time for an eight-hour pointing is ∼5 days (on a node equipped with 192 GBytes RAM and 2 Intel Xeon Gold 6130 CPU at 2.10 GHz, giving 32 physical compute cores) and takes an extra ∼30% of computing time for completion as compared to DDF-PIPELINE-V1. Hence ten 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 1, 2019, over 1000 pointings existed in the archive. Over ∼1000 pointings and ∼12 Petabytes of averaged and compressed LOFAR data (∼40 Petabytes 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

[12] https://pypi.python.org/pypi/SharedArray

Table 1. Overview of the deep field pointings used to synthesize the images on the Boötes and Lockman Hole extragalactic fields.

Project ID	LOFAR Obs. ID	Obs. Date	Start time UTC	Integration time [h]	$\nu_{\rm min}$ [MHz]	$\nu_{\rm max}$ [MHz]	$f_{\rm flag}$	nMS
			Boötes					
LC2_038	L236786	21/07/2014	10:44:00	8.0	120.0	168.7	37.8	25
LC2_038	L243561	15/09/2014	10:22:42	8.0	120.0	168.7	19.2	25
LC4_034	L346004	11/06/2015	16:04:35	8.0	120.2	167.0	10.9	24
LC4_034	L373377	17/09/2015	10:21:18	8.0	120.2	168.9	20.1	25
LC4_034	L374583	24/09/2015	10:09:57	8.0	120.2	168.9	10.6	25
LC4_034	L387597	29/09/2015	09:13:00	8.0	120.2	168.9	27.4	25
LC4_034	L387569	01/10/2015	09:00:00	8.0	120.2	168.9	32.3	25
LC4_034	L400135	10/10/2015	08:46:22	8.0	120.2	168.9	13.6	25
LC4_034	L401825	21/10/2015	08:00:30	8.0	120.2	168.9	9.0	25
LC4_034	L401839	22/10/2015	07:55:23	8.0	120.2	168.9	8.0	25
			Lockman Hole					
LC3_008	L274099	08/03/2015	20:11:00	8.7	120.2	168.9	12.4	25
LC3_008	L281008	14/03/2015	18:26:39	8.7	120.4	169.1	8.1	25
LC3_008	L294287	21/03/2015	19:11:00	8.7	120.2	168.9	18.7	25
LC3_008	L299961	24/03/2015	17:47:20	8.7	120.2	168.9	12.1	25
LC3_008	L340794	25/04/2015	17:08:00	8.7	120.2	168.9	14.5	25
LC3_008	L342938	08/05/2015	14:50:24	8.7	120.2	168.9	16.6	25
LT10_012	L659554	10/07/2018	11:11:00	8.0	120.2	168.9	9.9	25
LT10_012	L659948	12/07/2018	11:08:10	8.0	120.2	168.9	11.9	25
LT10_012	L664320	15/08/2018	08:49:00	8.0	120.2	168.9	11.0	25
LT10_012	L664480	19/08/2018	08:38:46	8.0	120.2	168.9	11.3	25
LT10_012	L667204	12/09/2018	07:06:09	8.0	120.2	168.9	11.2	25
LT10_012	L667218	13/09/2018	07:05:12	8.0	120.2	168.9	10.3	25

Notes. Columns $f_{\rm flag}$ and nMS show the fraction of flagged data and the number of measurement sets present in the archives.

highest-quality multidegree-scale ancillary data across the electromagnetic spectrum: the Boötes, Lockman Hole, ELAIS-N1, and NEP fields. The ultimate aim of the LoTSS-Deep Fields project is to reach noise levels of 10–15 μJy beam^{-1} in each of these fields (requiring ∼500 h of integration). The first LoTSS-Deep Fields data release consists of initial observations in three of these fields: Boötes (∼80 h) and Lockman Hole (∼112 h), presented in the current paper, and ELAIS-N1 (presented by Sabater et al. 2021, for an integration time of ∼170 h in Paper II). This first data release also includes an extensive effort of optical-to-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 multiwavelength data (Kondapally et al. 2021, Paper III). This is supplemented by high-quality photometric redshifts and a characterization of host galaxy properties (Duncan et al. 2021, Paper IV) and source classification (Best et al., in prep.). In order to put the LoTSS-Deep Fields observations into a wider context, in this section we briefly describe the multiwavelength data available on the Boötes and Lockman Hole fields, focusing on the radio coverage (for a more detailed description see Kondapally et al. 2021, Paper III).

4.1.1. Boötes field

The Boötes field is one of the NOAO Deep Wide Field Survey (NDWFS, Jannuzi & Dey 1999) fields covering ∼9.2 deg^2. It contains multiwavelength 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; Kenter et al. 2005; Masini et al. 2020), and optical data (Jannuzi & Dey 1999; Cool 2007; Brown et al. 2007, 2008). At radio frequencies, it has been mapped with the Westerbork Synthesis Radio Telescope (WSRT, see de Vries et al. 2002), the VLA (see Croft et al. 2008; Coppejans et al. 2015), the Giant Metrewave Radio Telescope (GMRT, see Intema et al. 2011; Williams et al. 2013), and LOFAR (van Weeren et al. 2014; Williams et al. 2016; Retana-Montenegro et al. 2018) at various depths, frequencies, resolutions, and covered areas (see Fig. 10 for an overview of the available radio data).

The Boötes pointing data that are presented in this paper are centered on $(\alpha, \delta) = (14{\rm h}32{\rm m}00{\rm s}, +34°30'00'')$ and were observed with the LOFAR-HBA in HBA_DUAL_INNER mode during Cycles 2 and 4, with a bandwidth of 48 MHz (see Table 1). The total integration time of ∼80 h is spread over ten scans of eight hours each.

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 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 (González-Solares et al. 2011), near IR (UKIRT Infrared Deep Sky Survey - Deep Extragalactic Survey; UKIDSS-DXS, see Lawrence et al. 2007), and with the Submillimeter 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

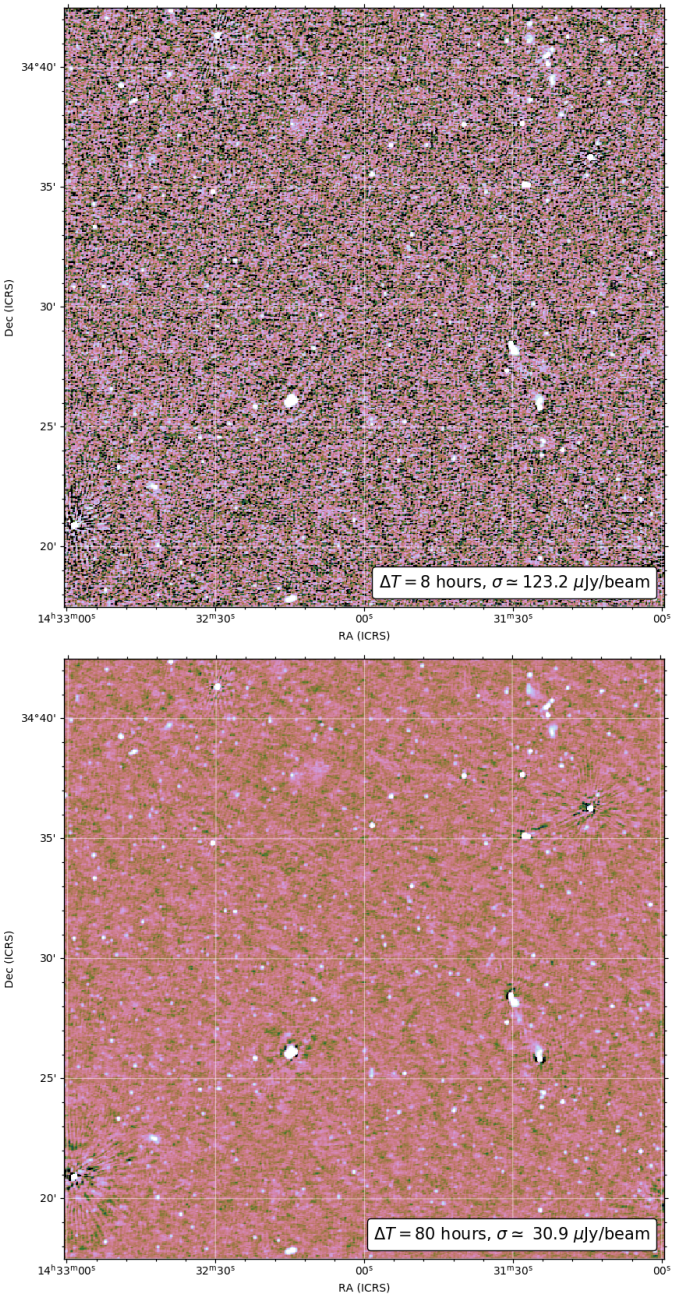

Fig. 11. Restored high resolution image toward the center of the Boötes field for the eight-hour image produced with Algorithm 1 (*top panel*) and the 80-h image produced with Algorithm 2 (*bottom panel*). Both images are thermal noise limited, with the same color scale used on both.

(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. 2009), WSRT (Guglielmino et al. 2012; Prandoni et al. 2018), and LOFAR at 150 MHz (Mahony et al. 2016). Figure 10 presents an overview of the available radio data on the Lockman Hole.

The Lockman Hole observation that we are presenting in this paper consists of 12 pointings of ∼8 h centered on $(\alpha, \delta) = (10\text{h}47\text{m}00\text{s}, +58°05'00'')$ and observed from March 2015 (Cycle 3) to November 2018 (Cycle 4). As for the Boötes field observation, we observed in HBA_DUAL_INNER with ∼48 MHz bandwidth, while the integration time depended on the LOFAR cycle (8.7 h in Cycle 3, 8 h in Cycle 10, see Table 1). The total integration time is ∼100 h.

Algorithm 2: Overview of the algorithm implemented in DDF-PIPELINE-V2. The function $\mathcal{I}$ represents the imaging step and takes as input the visibility vector $\mathbf{v}$ together with the beam model $\mathbf{B}_{\Omega_n}$ and KMS-estimated Jones matrices $\mathbf{J}_{\Omega_n}$ at locations Ω_n. The function $\mathcal{K}$ abstracts the DD calibration step, and takes as arguments the visibilities $\mathbf{v}$, the skymodel $\widehat{\mathbf{x}_\nu}$, 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 Ω_n in which to solve for. The extra functions $\mathcal{C}$, $\mathcal{B}$, and $\mathcal{F}$ represent the clustering, bootstrapping and smoothing steps respectively.

Data: Visibilities $\mathbf{v}$ calibrated from DI 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 $\widehat{\mathbf{x}_\nu}$

/* **On n_p×240 LOFAR HBA subbands** */
/* DD calibration */
2.1 $\widehat{\mathbf{J}} \leftarrow \mathcal{F} \circ \mathcal{K}\left(\mathbf{v}_{n_p\times 24}, \mathbf{B}_{\Omega_n}, \widehat{\mathbf{x}_\nu}|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_n\right)$;
/* DI calibration */
2.2 $\mathbf{v}^c_{n_p\times 24} \leftarrow \mathcal{K}\left(\mathbf{v}_{n_p\times 24}, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}, \widehat{\mathbf{x}_\nu}|\text{full}, \delta t_0, \delta\nu_0, \Omega_0\right)$;
/* DD imaging */
2.3 $\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}\left(\mathbf{v}^c_{n_p\times 24}, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}\right)$;
/* DD calibration */
2.4 $\widehat{\mathbf{J}} \leftarrow \mathcal{F} \circ \mathcal{K}\left(\mathbf{v}^c_{n_p\times 24}, \mathbf{B}_{\Omega_n}, \widehat{\mathbf{x}_\nu}|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_n\right)$;
/* Slow DD calibration */
2.5 $\widehat{\mathbf{J}_s} \leftarrow \mathcal{K}\left(\mathbf{v}^c_{n_p\times 24}, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}, \widehat{\mathbf{x}_\nu}|\text{scalar}, 43\text{min}, 2\text{MHz}, \Omega_n\right)$;
/* Final imaging steps */
2.6 $\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}\left(\mathbf{v}^c_{n_p\times 24}, \widehat{\mathbf{J}_s}\widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}\right)$;
/* Absolute flux density scale correction (see Sabater et al. 2021, for details) */
2.7 $\widehat{\mathbf{x}_\nu} \leftarrow \text{f}_\text{c}\widehat{\mathbf{x}_\nu}$;
2.8 Facet-based astrometric correction (see Shimwell et al. 2019, for details);

4.2. Image synthesis

The Lockman Hole and Boötes field data were reduced using Algorithm 2. In this approach, we first built a wide-band DI+DD self-calibrated sky model $\mathbf{x}_\nu$ from a single wide-band ∼8 -h observation using Algorithm 1. This model was then used to DI+DD calibrate all the n_p pointings (with $n_p = 10$ and $n_p = 12$ for the Boötes and Lockman Hole datasets, respectively) following Algorithm 2. This amounts to repeating Steps 1.13 to 1.18 of Algorithm 1 on a larger dataset. A comparison between the images synthesized from the 8- and 80-h datasets is presented in Fig. 11. On a single node equipped with ∼500 GB of 2.4 GHz RAM and 2 Intel Xeon CPU E5-2660 v4 at 2.00 GHz with 14 physical cores each, Algorithm 2 took ∼21 days to process the 80 h of Boötes data. Figures 12 (further discussed in Sect. 4.3) and 13 show the central parts of the of these deep LOFAR Boötes and Lockman Hole observations.

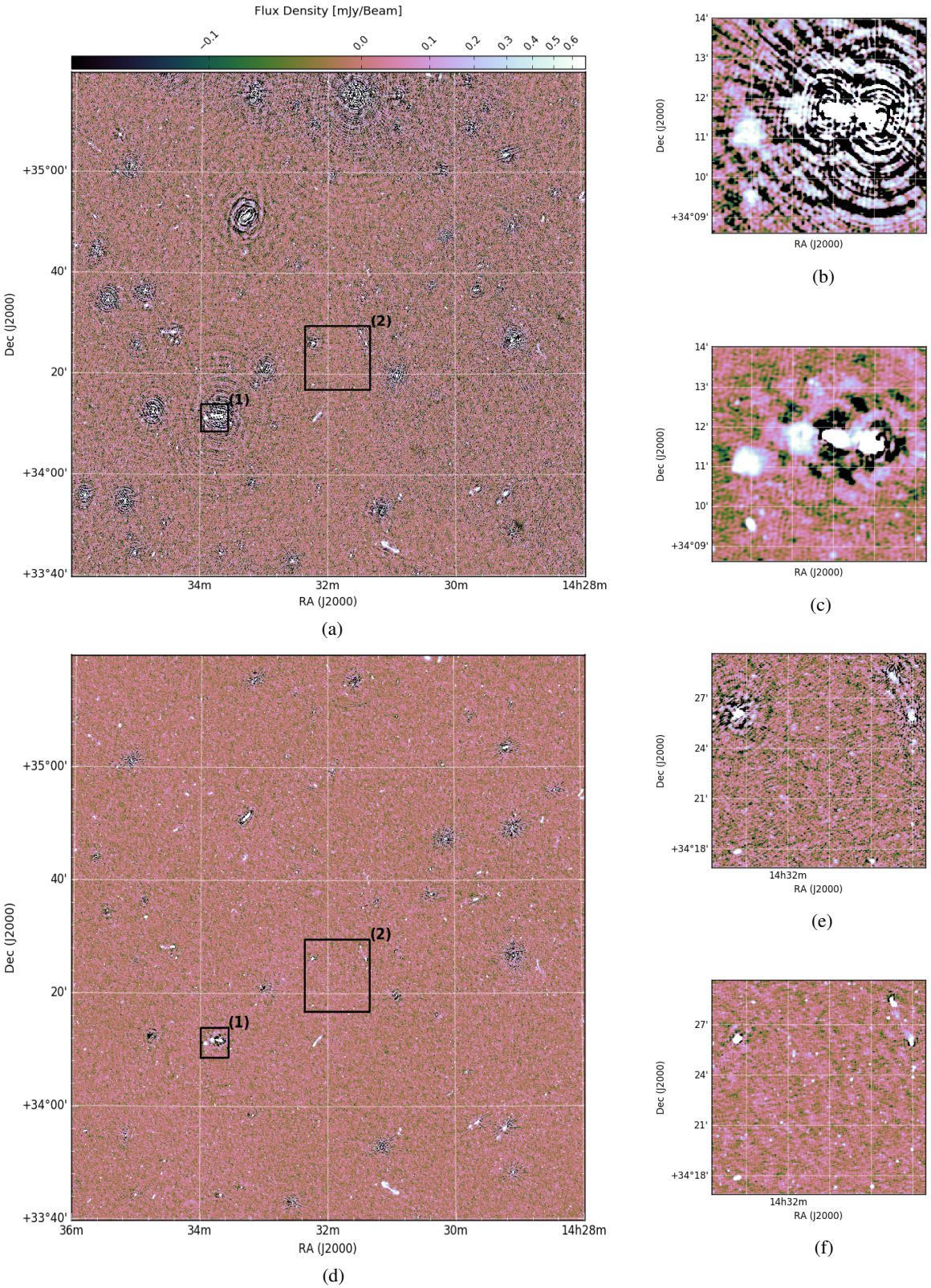

Fig. 12. Comparison between the LOFAR-HBA maps generated at 150 MHz by Retana-Montenegro et al. (2018) and in the current work. The color scale is the same on all panels and is displayed using an inverse hyperbolic sine function to render both the low level artifacts and some bright sources. (*a*) The central $\gtrsim$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 synthesized by kms-ddfacet (this work). (*d*) The same as in 12a, but imaged with Algorithm 2. (*e*) Zoom-in on region (2) of the map synthesized by Retana-Montenegro et al. (2018). (*f*) Zoom-in on region (2) of the map synthesized by kms-ddfacet (this work).

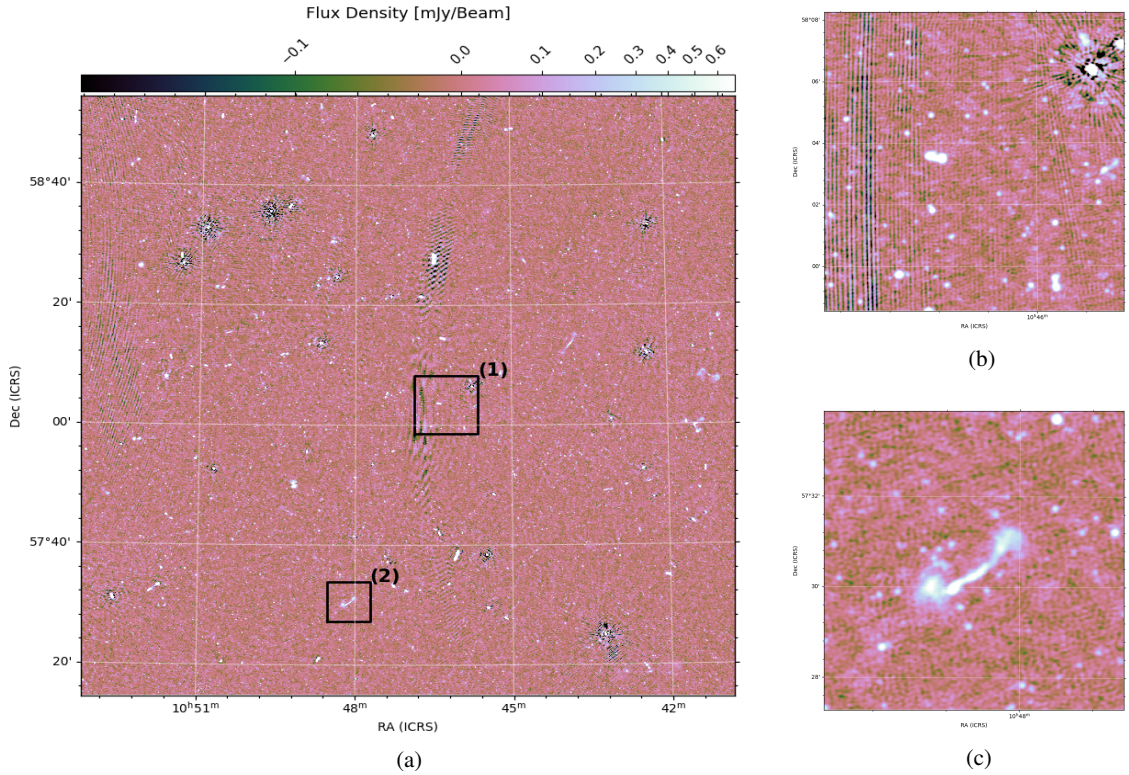

Fig. 13. Central region of the deep LOFAR-HBA maps of the Lockman Hole field generated at 150 MHz. The color scale is the same on all panels and is displayed using an inverse hyperbolic sine function to render both the low level artifacts and some bright sources. The stripy artifact seen in the zoom-out in *panel b* seems to be produced by the residual deconvolution and calibration errors of a few $\gtrsim$10 mJy beam^{-1} bright sources that are a few degrees away from the center of the field. (*a*) The central $\gtrsim$2 deg^2 part of the Lockman Hole field as imaged by Alg. 2. (Sect. 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.

Estimating the noise in radio maps is not straightforward since noise is correlated and non-Gaussian. Furthermore, while the covariance matrix should be entirely described by the PSF, 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 used the statistics of the min{.} estimator (which returns the minimum value of a given sample). Intuitively, while the I-Stokes image max{.} statistics have 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[13], the effective standard deviation is derived.

Figure 14 shows the cumulative distribution of the local noise in the Lockman Hole and Boötes fields maps, reaching $\lesssim$23 and $\lesssim$30 μJy beam^{-1}, respectively. Taking into account the number of pointings with their respective amounts of flagged data, we get total integration times of ~65 and ~88 h on the Boötes and Lockman Hole fields, respectively, giving a theoretical thermal noise difference of a factor of ~1.16, which is compatible with the observed value of ~1.3. Other factors to be taken into account to compare noise properties include the bootstrapping errors, the individual fields' average elevations, and the Galactic noise differences.

[13] The cumulative distribution $\mathcal{F}$ of $Y = \min\{X\}$ with $X \sim \mathcal{N}\{\mu = 0, \sigma = 1\}$ is $\mathcal{F}\{y\} = 1 - \left[\frac{1}{2}\left(1 - \mathrm{erf}\left\{\frac{y}{\sqrt{2}}\right\}\right)\right]^n$, where n is the number of pixels in a given box. Finding y_σ such that $\mathcal{F}\{y_\sigma\} = 1/2$ given the box size gives us a conversion factor from the minimum estimate to the standard deviation.

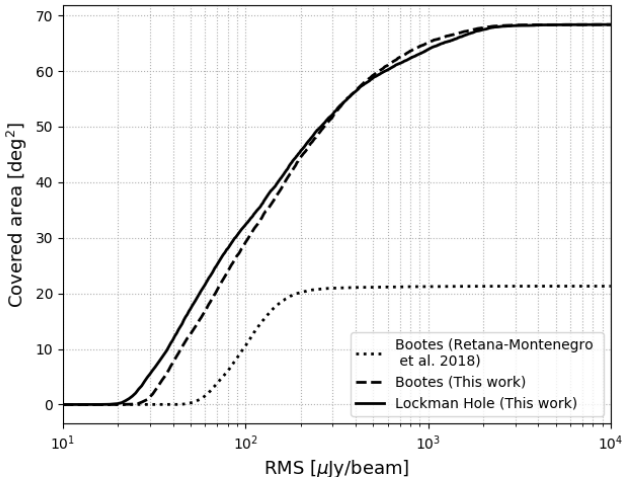

Fig. 14. Cumulative distribution of the local noise estimates in the various maps discussed here. As shown, we have imaged a larger fraction of LOFAR's HBA primary beam than the image presented in Retana-Montenegro et al. (2018).

4.3. Comparison with deep FACTOR/ image synthesis

The image of the Boötes field based on 55 h of LOFAR-HBA data and presented in Retana-Montenegro et al. (2018) reaches an unprecedented noise level image of ~55 μJy beam^{-1} at 150 MHz. To achieve such high sensitivity, Retana-Montenegro et al. (2018) applied third generation calibration and imaging to correct for the DDEs using the FACTOR package (developed by

van Weeren et al. 2016, see Sect. 2 for more detail). Because the set of LOFAR datasets used by Retana-Montenegro et al. (2018) is different[14] the comparison can only be approximate. In Fig. 12 we compare the images produced by Retana-Montenegro et al. (2018) and by Algorithm 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 Algorithm 2 and implemented in DDF-PIPELINE-V2.

4.4. Cataloging

In order to extract astrophysical information, we built a catalog of radio sources from the images produced by Algorithm 2 and the data described in Sect. 4. Even in the apparent flux maps, because of the imperfect calibration and imaging, the LoTSS-Deep Fields images have spatially variable noise; to deal with this issue, we used PyBDSF[15] (Python Blob Detector and Source Finder, see Mohan & Rafferty 2015) since it measures noise locally rather than globally. The sources were detected with 3 and 5σ for the island and peak detection threshold, respectively. The position-dependent noise was estimated using a sliding box algorithm with a size of 40×40 synthesized beams, except around bright sources where the box size was decreased to 15×15 beams to more accurately capture the increased noise in these regions. The columns kept in the final catalog are the source position, peak and integrated flux density, source size and orientation, the associated uncertainties, the estimated local rms at the source position, and a code describing the type of structure fit by PyBDSF. As described in Sabater et al. (2021), 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 Lockman Hole and Boötes fields, respectively. These numbers were estimated from the comparison between the LoTSS-Deep Fields flux densities and a variety of radio data available at other frequencies. The full catalogs cover out to 30% 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 catalogs are available online at the LOFAR survey web page[16], and a thorough analysis of the source catalogs will be presented by Mandal et al. (2021).

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 DD-Jones matrices. Due to the high number of free parameters in that system, and due to the finite amount of data points in the nonlinear RIME system, the inversion can be subject to ill-conditioning and the DD-*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 (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 DD algorithm to both 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. Detailed analyses of the LoTSS-Deep Fields catalogs (including the source counts of the Lockman Hole, Boötes, and ELAIS-N1 fields) are presented in Mandal et al. (2021). In the future, we plan to continue increasing the depths of these fields: The data needed to double the integration time on each field, with a further aim to increase this to 500 h in each field, are already in hand (or are scheduled).

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). M.B. acknowledges support from INAF under PRIN SKA/CTA FORECaST. M.B. 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. M.J.J. 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. P.N.B. and J.S. are grateful for support from the UK STFC via grant ST/R000972/1. M.J.H. acknowledges support from STFC via grant ST/R000905/1. W.L.W. acknowledges support from the ERC Advanced Investigator programme NewClusters 321271. W.L.W. 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). A.B. acknowledges support from the VIDI research programme with project number 639.042.729, which is financed by the Netherlands Organisation for Scientific Research (NWO). I.P. acknowledges support from INAF under the SKA/CTA PRIN "FORECaST" and the PRIN MAIN STREAM "SAuROS" projects M.B. 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. R.K. acknowledges support from the Science and Technology Facilities Council (STFC) through an STFC studentship.

References

Ashby, M. L. N., Stern, D., Brodwin, M., et al. 2009, ApJ, 701, 428
Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559
Bhatnagar, S., & Cornwell, T. J. 2017, AJ, 154, 197
Bhatnagar, S., Cornwell, T. J., Golap, K., & Uson, J. M. 2008, A&A, 487, 419
Biggs, A. D., & Ivison, R. J. 2006, MNRAS, 371, 963
Birdi, J., Repetti, A., & Wiaux, Y. 2020, MNRAS, 492, 3509
Bonnassieux, E., Tasse, C., Smirnov, O., & Zarka, P. 2018, A&A, 615, A66
Brentjens, M. A., & de Bruyn, A. G. 2005, A&A, 441, 1217
Brown, M. J. I., Dey, A., Jannuzi, B. T., et al. 2007, ApJ, 654, 858
Brown, M. J. I., Zheng, Z., White, M., et al. 2008, ApJ, 682, 937
Brunner, H., Cappelluti, N., Hasinger, G., et al. 2008, A&A, 479, 283
Butler, A., Huynh, M., Delhaize, J., et al. 2018, A&A, 620, A3
Ciliegi, P., Zamorani, G., Hasinger, G., et al. 2003, A&A, 398, 901
Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693
Cool, R. J. 2007, ApJS, 169, 21
Coppejans, R., Cseh, D., Williams, W. L., van Velzen, S., & Falcke, H. 2015, MNRAS, 450, 1477

[14] Out of the sets of 7 and 10 observations used in Retana-Montenegro et al. (2018) and in this work respectively, 4 are common, namely L243561, L374583, L400135, L401825.
[15] `https://www.astron.nl/citt/pybdsf`
[16] `https://www.lofar-surveys.org/`

Coppin, K., Chapin, E. L., Mortier, A. M. J., et al. 2006, MNRAS, 372, 1621
Croft, S., van Breugel, W., Brown, M. J. I., et al. 2008, AJ, 135, 1793
de Gasperin, F., Dijkema, T. J., Drabent, A., et al. 2019, A&A, 622, A5
de Ruiter, H. R., Zamorani, G., Parma, P., et al. 1997, A&A, 319, 7
de Vries, W. H., Morganti, R., Röttgering, H. J. A., et al. 2002, AJ, 123, 1784
Dewdney, P. E., Hall, P. J., Schilizzi, R. T., & Lazio, T. J. L. W. 2009, IEEE Proc., 97, 1482
Duncan, K. J., Kondapally, R., Brown, M. J. I., et al. 2021, A&A, 648, A4 (LoTSS SI)
Fomalont, E. B., Kellermann, K. I., Cowie, L. L., et al. 2006, ApJS, 167, 103
Garn, T. S., Green, D. A., Riley, J. M., & Alexander, P. 2010, Bull. Astron. Soc. India, 38, 103
Geach, J. E., Dunlop, J. S., Halpern, M., et al. 2017, MNRAS, 465, 1789
González-Solares, E. A., Irwin, M., McMahon, R. G., et al. 2011, MNRAS, 416, 927
Guglielmino, G., Prandoni, I., Morganti, R., & Heald, G. 2012, Resolving The Sky - Radio Interferometry: Past, Present and Future, 22
Hamaker, J. P. 2000, A&AS, 143, 515
Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, A&AS, 117, 137
Herrera Ruiz, N., O'Sullivan, S. P., Vacca, V., et al. 2020, MNRAS, 463, 2997
Heywood, I., Jarvis, M. J., Baker, A. J., et al. 2016, MNRAS, 460, 4433
Ibar, E., Ivison, R. J., Biggs, A. D., et al. 2009, MNRAS, 397, 281
Intema, H. T., van Weeren, R. J., Röttgering, H. J. A., & Lal, D. V. 2011, A&A, 535, A38
Intema, H. T., Jagannathan, P., Mooley, K. P., & Frail, D. A. 2017, A&A, 598, A78
Jannuzi, B. T., & Dey, A. 1999, ASP Conf. Ser., 193, 258
Jannuzi, B., Weiner, B., Block, M., et al. 2010, BAAS, 42, 513
Kazemi, S., Yatawatta, S., Zaroubi, S., et al. 2011, MNRAS, 414, 1656
Kenter, A., Murray, S. S., Forman, W. R., et al. 2005, ApJS, 161, 9
Kondapally, R., Best, P. N., Hardcastle, M. J., et al. 2021, A&A, 648, A3 (LoTSS SI)
Lacy, M., Baum, S. A., Chandler, C. J., et al. 2020, PASP, 132, 035001
Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599
Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., et al. 2003, PASP, 115, 897
Mahony, E. K., Morganti, R., Prandoni, I., et al. 2016, MNRAS, 463, 2997
Mandal, S., Prandoni, I., Hardcastle, M. J., et al. 2021, A&A, 648, A5 (LoTSS SI)
Martin, C., & GALEX Team. 2005, IAU Symp., 216, 221
Masini, A., Hickox, R. C., Carroll, C. M., et al. 2020, ApJS, 251, 1
Mauch, T., Cotton, W. D., Condon, J. J., et al. 2020, ApJ, 888, 61
Mohan, N., & Rafferty, D. 2015, PyBDSF: Python Blob Detection and Source Finder
Molenaar, G., & Smirnov, O. 2018, Astron. Comput., 24, 45
Murray, S. S., Kenter, A., Forman, W. R., et al. 2005, ApJS, 161, 1
Norris, R. 2010, AAS Meeting Abstracts, 215, 604.05
Offringa, A. R., van de Gronde, J. J., & Roerdink, J. B. T. M. 2012, A&A, 539, A95
Oliver, S. J., Bock, J., Altieri, B., et al. 2012, MNRAS, 424, 1614
O'Sullivan, S., Brüggen, M., Van Eck, C., et al. 2018, Galaxies, 6, 126
O'Sullivan, S. P., Brueggen, M., & Vazza, F. 2020, MNRAS, 495, 2607
Owen, F. N., & Morrison, G. E. 2008, AJ, 136, 1889
Owen, F. N., Morrison, G. E., Klimek, M. D., & Greisen, E. W. 2009, AJ, 137, 4846
Padovani, P. 2016, A&ARv, 24, 13
Pearson, T. J., & Readhead, A. C. S. 1984, ARA&A, 22, 97
Polletta, M. d. C., Wilkes, B. J., Siana, B., et al. 2006, ApJ, 642, 673
Prandoni, I. & Seymour, N. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 67
Prandoni, I., Guglielmino, G., Morganti, R., et al. 2018, MNRAS, 481, 4548
Rengelink, R. B., Tang, Y., de Bruyn, A. G., et al. 1997, A&AS, 124, 259
Repetti, A., Birdi, J., Dabbech, A., & Wiaux, Y. 2017, MNRAS, 470, 3981
Retana-Montenegro, E., Röttgering, H. J. A., Shimwell, T. W., et al. 2018, A&A, 620, A74
Richards, E. A. 2000, ApJ, 533, 611
Sabater, J., Best, P. N., Tasse, C., et al. 2021, A&A, 648, A2 (LoTSS SI)
Schinnerer, E., Carilli, C. L., Scoville, N. Z., et al. 2004, AJ, 128, 1974
Shimwell, T. W., Röttgering, H. J. A., Best, P. N., et al. 2017, A&A, 598, A104
Shimwell, T. W., Tasse, C., Hardcastle, M. J., et al. 2019, A&A, 622, A1
Smirnov, O. M. 2011, A&A, 527, A108
Smirnov, O. M., & Tasse, C. 2015, MNRAS, 449, 2668
Smolčić, V., Novak, M., Bondi, M., et al. 2017a, A&A, 602, A1
Smolčić, V., Novak, M., Delvecchio, I., et al. 2017b, A&A, 602, A6
Sob, U. M., Bester, H. L., Smirnov, O. M., Kenyon, J. S., & Grobler, T. L. 2020, MNRAS, 491, 1026
Tasse, C. 2014a, ArXiv e-prints [arXiv:1410.8706]
Tasse, C. 2014b, A&A, 566, A127
Tasse, C., Röttgering, H., & Best, P. N. 2010, A&A, 525, A127
Tasse, C., Hugo, B., Mirmont, M., et al. 2018, A&A, 611, A87
Van Eck, C. L., Haverkorn, M., Alves, M. I. R., et al. 2018, A&A, 613, A58
van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2
van Weeren, R. J., Williams, W. L., Tasse, C., et al. 2014, ApJ, 793, 82
van Weeren, R. J., Williams, W. L., Hardcastle, M. J., et al. 2016, ApJS, 223, 2
van Weeren, R. J., Shimwell, T. W., Botteon, A., et al. 2020, ArXiv e-prints [arXiv:2011.02387]
Wayth, R. B., Lenc, E., Bell, M. E., et al. 2015, PASA, 32, e025
Williams, W. L., Intema, H. T., & Röttgering, H. J. A. 2013, A&A, 549, A55
Williams, W. L., van Weeren, R. J., Röttgering, H. J. A., et al. 2016, MNRAS, 460, 2385
Yatawatta, S. 2015, MNRAS, 449, 4506
Yatawatta, S., Zaroubi, S., de Bruyn, G., Koopmans, L., & Noordam, J. 2008, ArXiv e-prints [arXiv:0810.5751]
Zwart, J., Wall, J., Karim, A., et al. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 172

Appendix A: LoTSS first data release: overview of DDF-PIPELINE-V1

The data processing strategy of the LoTSS first data release has been extensively described by Shimwell et al. (2019). Since addressing the issues described in Sect. 2 involves making improvements relative to this approach, we give a brief description here of the data reduction strategy in DDF-PIPELINE-V1 (the various steps are outlined in Algorithm 0).

Algorithm 0: Overview of the algorithm implemented in DDF-PIPELINE-V1 to produce the LoTSS-DR1 images. The function $\mathcal{I}$ represents the imaging step and takes as input the visibility vector $\mathbf{v}$ together with the beam model $\mathbf{B}_{\Omega_n}$ and kMS-estimated Jones matrices $\mathbf{J}_{\Omega_n}$ at locations Ω_n. The function $\mathcal{K}$ abstracts the DD calibration step, and takes as arguments the visibilities $\mathbf{v}$, the skymodel $\widehat{\mathbf{x}_\nu}$, 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 Ω_n in which to solve for. The extra functions $\mathcal{C}$, $\mathcal{N}$, and $\mathcal{B}$ represent the clustering, normalisation (see text), and bootstrapping steps respectively.

Data: Visibilities $\mathbf{v}$ calibrated from DI effects using PREFACTOR.
Result: Deconvolved image $\widehat{\mathbf{x}_\nu}$

/* **On 60 LOFAR HBA subbands** */
/* DI initial deconv and clustering */
$\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}(\mathbf{v}_6, \mathbf{B}_{\Omega_r})$;
$\Omega_n \leftarrow \mathcal{C}(\widehat{\mathbf{x}_\nu})$;
/* Phase only DD calibration */
$\widehat{\mathbf{J}} \leftarrow \varphi \circ \mathcal{K}(\mathbf{v}_6, \widehat{\mathbf{x}_\nu}, \mathbf{B}_{\Omega_n}|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_n)$;
/* Absolute flux density scale bootstrapping */
$\widehat{\mathbf{v}} \leftarrow \mathcal{B}(\widehat{\mathbf{v}_6})$;
$\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}\left(\mathbf{v}_6, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}\right)$;
/* DD calibration and imaging */
$\widehat{\mathbf{J}} \leftarrow \mathcal{N} \circ \mathcal{K}(\mathbf{v}_6, \widehat{\mathbf{x}_\nu}, \mathbf{B}_{\Omega_n}|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_n)$;
$\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}\left(\mathbf{v}_6, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}\right)$;

/* **On 240 LOFAR HBA subbands** */
/* Deep DD calibration and imaging */
$\widehat{\mathbf{J}} \leftarrow \mathcal{N} \circ \mathcal{K}(\mathbf{v}_{24}, \widehat{\mathbf{x}_\nu}, \mathbf{B}_{\Omega_n}|\text{scalar}, 1\text{min}, 2\text{MHz}, \Omega_n)$;
$\widehat{\mathbf{x}_\nu} \leftarrow \mathcal{I}\left(\mathbf{v}_{24}, \widehat{\mathbf{J}}\mathbf{B}_{\Omega_n}\right)$;
Facet-based astrometric correction (see Shimwell et al. 2019, for details);

As discussed in Sect. 2, the calibration and imaging problem is non-convex and ill-posed. Beyond the computational issues, the great difficulty of the DDE calibration is sky incompleteness because the DD-$\mathcal{C}$-RIME nonlinear 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 DD deconvolution with DDFacet.

To try to compensate for this effect, in DDF-PIPELINE-V1 (Algorithm 0) we introduced an inner uv-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 that the sky incompleteness was generating some baseline-dependent systematic errors. So for every given direction and solution interval in Shimwell et al. (2019) we were trying to find a gain vector $\boldsymbol{g}$ such that $\boldsymbol{g}\boldsymbol{g}^H \sim \left\langle \boldsymbol{g}_{t\nu}\boldsymbol{g}_{t\nu}^{\mathrm{H}} \right\rangle$ (where $\mathbf{A}^H$ is the Hermitian transpose of matrix $\mathbf{A}$). This amounts to constraining the baseline-dependent error to be solely antenna-dependent. This normalization (described by the function symbol $\mathcal{N}$ in Algorithm 0) was able to recover some extended emission otherwise absorbed in the calibration solution. However, as shown in Fig. 8a 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 that have a bright $\gtrsim$1 Jy source within the primary beam (such as 3C sources), DDF-PIPELINE-V1 was not able to converge.