Fast Sky to Sky Interpolation for Radio Interferometric Imaging

Nicolas Monnier, François Orieux, Nicolas Gac, Cyril Tasse, Erwan Raffin, David Guibert

Fast Sky to Sky Interpolation for Radio Interferometric Imaging

Nicolas Monnier*, François Orieux*, Nicolas Gac*, Cyril Tasse†, Erwan Raffin‡, David Guibert‡

*Laboratoire des Signaux et Systèmes, Univ. Paris-Saclay, CNRS, CentraleSupélec, 91 190 Gif-sur-Yvette, France
†GEPI, Observatoire de Paris, CNRS, Univ. Paris Diderot, 92 190 Meudon, France
‡CEPP - Center for Excellence in Performance Programming, Atos Bull, 35 700 Rennes, France

***Abstract*—Reconstruction of radio interferometric images requires the processing of data in Fourier space that dot not have regular coordinates, preventing direct use of the Fast Fourier Transform. The most common solution is to rely on interpolation algorithms, called gridding, that are computationally expensive. In this paper, we propose an algorithmic reinterpretation, named sky to sky method, to reduce the computation cost of the gridding operation and its adjoint, the degridding, when used successively. We analyze the impact of interpolation step size regarding the computation cost and the reconstruction error. We also illustrate this optimization on a full reconstruction with gradient descent and CLEAN algorithm. Finally, we obtain acceleration factors between 1.2 and 16.4 without additional approximation.**

***Index Terms*—Interpolation, inverse problem, gridding, Computation acceleration.**

I. Introduction

Radio interferometric imaging is a technique that reconstructs an image of the sky from a radio observation [1]. Since the data sample incompletely the Fourier space, the reconstruction can be seen as a Fourier synthesis problem that is known to be ill-posed [2]. Moreover, the current and future generations of radio-telescope, such as VLA, LOFAR [3], or SKA [4], are characterized by a higher dynamic range and a higher resolution than the previous generation. These characteristics come with a very large amount of generated data, making our problem large-scaled.

We can find in the literature many different techniques to reconstruct an image of the sky based on the forward problem. A common technique in the radio-astronomy community is the standard CLEAN method and its variants [5], [6], implemented in imagers such as DDFacet or WSClean [7], [8]. Alternatively, we can find techniques based on compressed sensing and convex optimization techniques for image deconvolution [9], [10].These techniques are based on algorithms that repetitively apply forward and adjoint operators. In radio interferometry, the measurements are Fourier coefficients and their coordinates are non-uniform or non-regular. Non-uniform discrete Fourier transform [11] is known to have a high computing cost because of the high number of measurements to process. Thus, a common technique is to use the FFT to interpolate the measure on a regular grid using "gridding" and "degridding" operators. As the model of MRI image reconstruction is close to the Radio interferometric one, MRI uses similar interpolation methods. MRI methods tend to use nuFFT due to the lower amount of data [12], [13], which is currently not feasible for modern radio telescopes. Due to the huge amount of data to process, especially with the upcoming instruments like SKA, the computation of these operators needs to be reduced. In order to address these challenges, we propose a reinterpretation with a new and more efficient implementation of well-established algorithms.

In the first part, we will see the radio interferometric imaging problem and the mandatory gridding and degridding steps. Then, the algorithmic optimization based on the succession of the forward and adjoint operators, defined as the sky to sky (S2S) method, will be presented. Finally, we will present results on time and memory consumption of this optimization, which will be illustrated with the Cotton-Schwab CLEAN algorithm to estimate a simulated sky.

II. Radio Interferometric Imaging

A. Forward problem

A radio interferometer array is a network of antennas that generates measures of the radio emission of the observed sky $\boldsymbol{x}$. Each antenna's pair is defined with a baseline $\boldsymbol{b} = (u, v, w)$ where u, v, and w are the coordinates in unit of wavelength λ. The measurement of one antenna pair of baseline $\boldsymbol{b}$, or so-called *visibility*, is defined as

$$\mathrm{v}(u, v, w) = \iint \frac{x(l, m, n)}{n} e^{-2i\pi(ul+vm+w(n-1))} \,\mathrm{d}l \,\mathrm{d}m \quad (1)$$

where x is the sky brightness distribution using a coordinate system (l, m, and $n = \sqrt{1 - l^2 - m^2}$) that indicates an angular position. If the array of antennas is coplanar ($w \approx 0$) and the FOV is small ($n \approx 1$), each visibility $\mathrm{v}(u, v)$ is the 2D spatial Fourier transform of the sky distribution at frequency (u, v). This result is known as the van Cittert-Zernike theorem [1], and Eq. (1) becomes

$$\mathrm{v}(u, v) = \iint x(l, m) e^{-2i\pi(ul+vm)} \,\mathrm{d}l \,\mathrm{d}m. \quad (2)$$

Since the array cannot cover the full (u, v) plan, radio interferometric imaging aims to recover the sky from incomplete visibility measurements, leading to an ill-posed inverse problem.

After discretization, the true sky can be represented by $N \times N = P$ pixels grid as $\boldsymbol{x} \in \mathbb{R}^P$, and the visibility measurement

as a vector $\boldsymbol{v} \in \mathbb{C}^M$. Relation between $\boldsymbol{x}$ and $\boldsymbol{v}$ is described by the discrete forward problem

$$\boldsymbol{v} = \boldsymbol{S}\boldsymbol{F}\boldsymbol{x} + \boldsymbol{n} \tag{3}$$

where $\boldsymbol{F} \in \mathbb{C}^{P\times P}$ is the Fourier transform, $\boldsymbol{S} \in \mathbb{C}^{M\times P}$ maps M visibilities to P Fourier coefficients (similar to a sampling operator and called "degridding"), and $\boldsymbol{n} \in \mathbb{C}^M$ is an i.i.d. Gaussian noise. Finally, the "dirty image" $\boldsymbol{y} \in \mathbb{R}^P$ is defined as the back propagation of the visibilities, or the application of the adjoint $\boldsymbol{y} = \boldsymbol{F}^\dagger\boldsymbol{S}^\dagger\boldsymbol{v}$ where $\boldsymbol{F}^\dagger$ is the inverse Fourier transform and $\boldsymbol{S}^\dagger \in \mathbb{C}^{P\times M}$ is the "gridding" operator, adjoint of $\boldsymbol{S}$, that interpolate visibilities on regular coordinates.

B. Gridding operator

In most cases, visibilities do not have regular coordinates because they are sampled in time, not in distance. In order to use an FFT, the output of $\boldsymbol{F}$ needs to be interpolated on the visibility coordinates.

The continuous coordinate $\boldsymbol{b}_i = (u_i, v_i)$ associated with the visibility v_i is approximated to the position of the nearest neighbor on a thin grid with steps (Δ_u, Δ_v), such as the new coordinate is

$$\widetilde{\boldsymbol{b}}_i = (n_i\Delta_u, m_i\Delta_v). \tag{4}$$

If K is the oversampling factor, the FFT steps of the grid $\boldsymbol{g}$ are $(K\Delta_u, K\Delta_v)$. In that case, the gridding operator $\boldsymbol{S}^\dagger$ for the FFT coordinate $\boldsymbol{b}_{kl} = (n_kK\Delta_u, m_lK\Delta_v)$ is defined as

$$\begin{aligned}\boldsymbol{g}(\boldsymbol{b}_{kl}) &= \sum_{i=1}^{M} C^\dagger\big(\boldsymbol{b}_k - \widetilde{\boldsymbol{b}}_i\big)\mathrm{v}_i \\ &= \sum_{i=1}^{M} C^\dagger\big(n_kK\Delta_u - n_i\Delta_u, m_lK\Delta_v - m_i\Delta_v\big)\mathrm{v}_i\end{aligned} \tag{5}$$

where $C^\dagger$ is a 2D kernel of size $C_{supp} \times C_{supp}$ in FFT step. The choice and the size of this interpolator to avoid aliasing effects and maximize the accuracy is a well-known subject in the literature, see for instance [14]–[16]. In our case, we used the Kaiser-Bessel function with a support $C_{supp} = 7$.

C. Sky estimation

The sky $\boldsymbol{x}$ reconstruction for incomplete data $\boldsymbol{v}$ is an ill-posed inverse problem. Therefore, interferometric imaging is usually defined as

$$\widehat{\boldsymbol{x}} = \arg\min_x \|\boldsymbol{v} - \boldsymbol{S}\boldsymbol{F}\boldsymbol{x}\|^2 + R(\boldsymbol{x}) \tag{6}$$

where $\|\boldsymbol{v} - \boldsymbol{S}\boldsymbol{F}\boldsymbol{x}\|^2$ is the data fidelity term and R a regularizer [2], [17], [18]. Algorithms used to solve this optimization problem require computing many gradients of the criterion during their iterative processes. The computing cost of a gradient of the data fidelity term

$$\nabla J(\boldsymbol{x}) = \boldsymbol{F}^\dagger\boldsymbol{S}^\dagger(\boldsymbol{v} - \boldsymbol{S}\boldsymbol{F}\boldsymbol{x}) \tag{7}$$

can be very high since it requires the computation of the gridding and degridding. We proposed in this paper a new formulation to reduce the computing cost of the gradient evaluation without additional approximation.

III. Sky to Sky gradient computation method

A. Gridding and degridding decomposition

A way to implement the gridding operator $\boldsymbol{S}^\dagger$ is to decompose it into three sub-operators : accumulation, convolution, and subsampling.

1) **Accumulation $\boldsymbol{A}^\dagger$.**
Let's define $\boldsymbol{g'}$ an oversampled 2D grid of size $KN \times KN = K^2P$ with thin pixel resolution (Δ_u, Δ_v). The pixel value associated with the coordinate $\boldsymbol{b}_p$ is

$$\boldsymbol{g'}(\boldsymbol{b}_p) = \begin{cases}\sum_i \mathrm{v}_i & \text{if } \widetilde{\boldsymbol{b}}_i = \boldsymbol{b}_p, \\ 0 & \text{otherwise,}\end{cases} \tag{8}$$

where $\boldsymbol{g'}$ is a grid of accumulated visibilities that are approximated at the same position on the thin grid. Then, the accumulation operator is $\boldsymbol{A}^\dagger \in \{0,1\}^{K^2P\times M}$ with one 1 per column.

2) **Convolution $\boldsymbol{C}^\dagger$.**
Considering a 2D grid $\widetilde{\boldsymbol{g}}$ of size $KN \times KN$ with thin step resolution (Δ_u, Δ_v) (same size as $\boldsymbol{g'}$). This step is a two-dimensional discrete convolution between the thin grid $\boldsymbol{g'}$ and the kernel $C^\dagger$, such as the value associated to the coordinate $\boldsymbol{b}_p$ is

$$\widetilde{\boldsymbol{g}}(\boldsymbol{b}_p) = \sum_i \sum_j C^\dagger\Big(\boldsymbol{b}_p - \boldsymbol{b}_{ij}\Big)\boldsymbol{g'}(\boldsymbol{b}_{ij}). \tag{9}$$

3) **Sub-sampling $\boldsymbol{O}^\dagger$.**
The FFT grid is coarser by a factor K regarding the oversampled grid. The sub-sampling operator $\boldsymbol{O}^\dagger$ is then a $P \times K^2P$ matrix, filled with one 1 per line and 0 otherwise.

Therefore, the gridding step, illustrated by Fig. 1, can be written as

$$\boldsymbol{S}^\dagger = \boldsymbol{O}^\dagger\boldsymbol{C}^\dagger\boldsymbol{A}^\dagger. \tag{10}$$

In practice, these operators are not instantiated as a matrix but are implicit, being applied by dedicated code.

The degridding operator $\boldsymbol{S}$ is the adjoint of the gridding $\boldsymbol{S}^\dagger$. Thus, from Eq. (10), we can define the degridding decomposition as

$$\boldsymbol{S} = \boldsymbol{A}\boldsymbol{C}\boldsymbol{O}, \tag{11}$$

where $\boldsymbol{O}$ is an oversampling operator filling with zeros, $\boldsymbol{C}$ is a convolution but with a flipped kernel, and $\boldsymbol{A}$ is a sampling–mapping operator that builds M visibilities from K^2P coefficients. This decomposition is illustrated Fig. 2.

B. The gradient computation optimization

The standard method to compute the gradient operator described by Eq. (7) can be developed such as it becomes

$$\begin{aligned}\nabla J(\boldsymbol{x}) &= \boldsymbol{F}^\dagger\boldsymbol{S}^\dagger\boldsymbol{v} - \boldsymbol{F}^\dagger\boldsymbol{S}^\dagger\boldsymbol{S}\boldsymbol{F}\boldsymbol{x} \\ &= \boldsymbol{y} - \boldsymbol{F}^\dagger\boldsymbol{S}^\dagger\boldsymbol{S}\boldsymbol{F}\boldsymbol{x}.\end{aligned} \tag{12}$$

This decomposition, mentioned by the Fast Holographic Deconvolution (FHD) [19], will be referred to as the Sky to Sky method (S2S). We can note that $\boldsymbol{F}\boldsymbol{S}^\dagger\boldsymbol{v}$ is the dirty image $\boldsymbol{y}$,

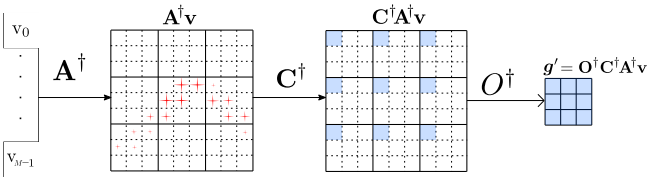

Fig. 1. Gridding decomposition. First, accumulation of visibilities on a 2D thin grid of size $KN \times KM$, then a 2D discrete convolution with a kernel $C^\dagger$, and finally, a K sub-sampling.

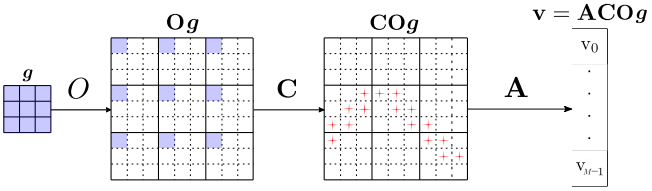

Fig. 2. Degridding decomposition. First, oversampling of a map $\boldsymbol{g}$ with a factor K, then a 2D discrete convolution with a kernel C, and finally, a mapping operator $\boldsymbol{A}$ maps the Fourier coefficients to M visiblities.

which is computed only once. Based on the above development, the succession of degridding and gridding operators can be rewritten as

$$\boldsymbol{S}^\dagger \boldsymbol{S} = \boldsymbol{O}^\dagger \boldsymbol{C}^\dagger \boldsymbol{A}^\dagger \boldsymbol{A} \boldsymbol{C} \boldsymbol{O} \tag{13}$$

with $\boldsymbol{A}^\dagger \boldsymbol{A} = \boldsymbol{A}^*$, and $\boldsymbol{A}^* \in \mathbb{N}^{K^2P \times K^2P}$ a diagonal matrix, such as $\boldsymbol{a} = \text{diag}(\boldsymbol{A}^*)$. The value of each element of $\boldsymbol{a}$ corresponds to the number of visibilities approximated to the corresponding thin coordinate. This decomposition is illustrated by Fig. 3. We define M' as the number of non-zero elements in $\boldsymbol{a}$. With M the number of raw data, practical cases show that $M > M'$ or $M \gg M'$ regarding the oversampling factor K. Consequently, the S2S method avoids processing the $M-$vector of raw visibilities, but instead processes a smaller $M'-$vector. Moreover, unlike FHD [19] who builds and stores $\boldsymbol{S}^\dagger \boldsymbol{S} \in \mathbb{C}^{P \times P}$ as a matrix, we only need to build the sparse vector $\boldsymbol{a}$. Thus, Eq. (13) writes

$$g'(\boldsymbol{b}_k) = \sum_{i'=1}^{M'} C^\dagger \left(\boldsymbol{b}_k - \boldsymbol{b}_{i'} \right) a_{i'} \sum_{i,j} C \left(\boldsymbol{b}_{i'} - \boldsymbol{b}_{ij} \right) g(\boldsymbol{b}_{ij}). \tag{14}$$

For the same oversampling factor, the gradient computed by the S2S method Eq. (12) and the STD method Eq. (7) is the same. Furthermore, the computing cost of the gradient using the S2S method is given by

$$2M'C_{supp}^2 + M' + 2P \log P + P \tag{15}$$

while the computing cost of the standard approach is

$$2MC_{supp}^2 + 2P \log P + M. \tag{16}$$

Hence, for the same K, the presented method is more efficient without additional errors when

$$M' < \frac{2MC_{supp}^2 + M - P}{2C_{supp}^2 + 1}. \tag{17}$$

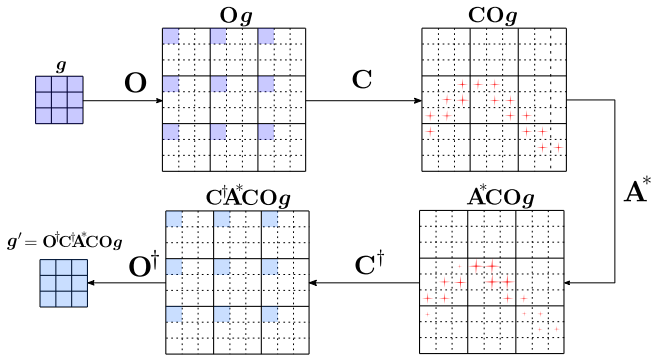

Fig. 3. Degridding and gridding decomposition using the sky to sky method. The visibility vector is not necessary anymore.

IV. Experiments and Results

The experiment aims to show the advantage of using the S2S method to improve the speedup of a gradient evaluation and to show the impact on memory usage regarding the oversampling factor used. Moreover, contrary to FHD [19], we do not need to store the few gigabytes matrix but use an implicit operator. Simulated visibilities have been generated following Eq. (2) for an 8 hours and 64 frequency channels observation of the VLA-D with 1 sec dump time between two samples and a single polarization. All the experiments were performed on a BullSequana X400 system based on Intel(R) Xeon(R) Platinum 8358 processor, computing a first implementation of parallel Cython code.

A. Gradient computation results

For the same oversampling factor K, the STD and S2S methods produce the same result without additional error. However, the S2S method's efficiency depends on K.

1) Oversampling factor error: Visibilities approximated coordinates $\widetilde{\boldsymbol{b}}_i$ depend on K. As shown by [20], a low oversampling factor involves a higher error. The value $K = 63$ is often used in different imaging algorithms that use gridding and degridding operators. We will then take the gradient computed with this oversampling factor with raw data as a reference. Fig. 4 and 5 show the maximum error

$$\text{MaxAE}(K) = \max(\left| \nabla(\boldsymbol{x}_{K=63}) - \nabla(\boldsymbol{x}_{K=j}) \right|) \tag{18}$$

and the mean absolute error between gradients

$$\text{MAE}(K) = \frac{1}{P} \sum_{i=0}^{P} |\nabla(\boldsymbol{x}_{i,K=63}) - \nabla(\boldsymbol{x}_{i,K=j})|. \tag{19}$$

As expected, the error increase as the oversampling factor decrease. Moreover, it is common to use an averaging technique to reduce the amount of data to process, even if it introduces a decorrelation [1] which causes information loss and artifacts. In our case, we averaged the data over 10 seconds to see the error's impact. As expected, the error is higher but merge the raw data curve for low oversampling factors.

2) Time cost: The computation time for raw and averaged data is shown in table I. For any oversampling factor and any data set, the proposed method produces the same result as the STD method with a faster computation time. As expected, the computation time using the STD method is constant over the oversampling factors. On the other hand, the S2S computation time decreases regarding K. For $K = 5$, the S2S method computes $\nabla J(\boldsymbol{x})$ using raw data without additional approximation but with an acceleration factor of 16.3. Moreover, the computation time of raw for K up to 15 using S2S is lower than the STD method for any K using average data.

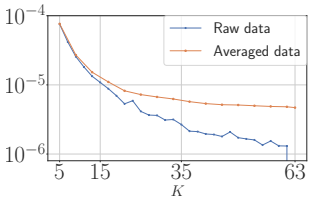

Fig. 4. Mean absolute error.

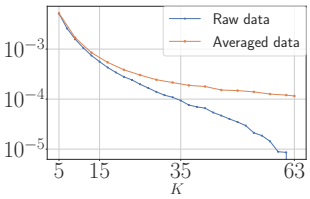

Fig. 5. Maximum absolute error.

K	STD	S2S (speedup)	STD ave.	S2S ave. (speedup)
5	242.9	14.8 (×16.4)	33.7	13.9 (×2.4)
15	242.9	24.1 (×10.0)	33.7	22.1 (×1.6)
35	242.9	43.9 (×5.6)	33.7	27.0 (×1.3)
63	242.9	71.8 (×3.4)	33.7	28.5 (×1.2)

TABLE I
TIME IN SECONDES TO COMPUTE THE GRADIENT.

3) Memory cost: Finally, we compare the memory consumption between the standard and proposed S2S methods. The first method needs to keep in memory the visibilities for all frequency channels and their coordinates on the (u, v) plan. The latter method needs to keep in memory the dirty images for each channel and the sparse vector $\boldsymbol{a}$. The memory consumption for the raw data of the test case is shown in table II. As $M \gg P$, the proposed method allows a lower memory cost. Moreover, with low K, visibilities coordinates approximation on the thin grid is more substantial, and consequently, the sparse vector $\boldsymbol{a}$ is smaller. Results for averaged data are also shown in table II. The memory cost of the proposed method is slightly higher than the standard method because of dirty image storage.

K	STD	S2S	STD ave.	S2S ave.
5	1.94	0.46	0.20	0.45
15	1.94	0.59	0.20	0.55
35	1.94	0.85	0.20	0.63
63	1.94	1.20	0.20	0.65

TABLE II
MEMORY CONSUMPTION NEEDED TO COMPUTE THE GRADIENT IN GB.

B. Sky estimation method results

The estimation of the simulated sky, consisting of 50 points sources with ranged amplitudes from 2×10^{-2} to 1, has been done using the Cotton-Schwab CLEAN [21] algorithm. This iterative algorithm, detailed by [22], is divided into two main parts. The first part, known as *major-cycle*, aims to compute the gradient of the data fidelity term $\nabla J(\boldsymbol{x}^{(k)})$. We implemented the custom S2S method to compute $\nabla J(\boldsymbol{x}^{(k)})$ as Eq. (12). The second part, also known as *minor-cycle*, builds iteratively an estimation $\boldsymbol{x}^{(k+1)}$ from $\nabla J(\boldsymbol{x})^{(k+1)}$ and $\boldsymbol{F}^\dagger \boldsymbol{S}^\dagger \boldsymbol{S} \boldsymbol{F}$ approximated as a convolution. These two steps are repeated until a global stopping criterion is reached.

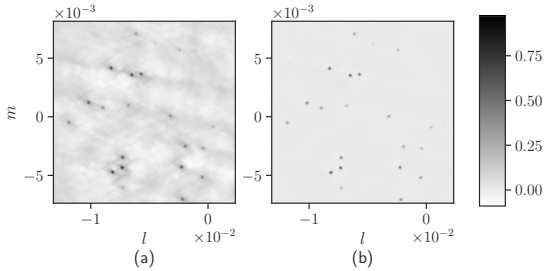

Fig. 6. Zoom on 400 × 400 pixels : (a) Dirty image (b) Restoration.

The panel of Fig. 6 shows a zoom of the dirty image. We performed two major cycles and 2500 minor cycles with a clean gain of 20%. All the initial point sources are identified in the restored image as illustrated on the right panel of Fig. 6. Moreover, table. III shows the time to perform the algorithm for each case. Again, we obtain acceleration factors every time the proposed method is used instead of the STD method. When using the raw data, the maximum speedup is up to 3.44.

K	STD	S2S (speedup)	STD ave.	S2S ave. (speedup)
5	641.2	186.8 (×3.4)	117.3	77.8 (×1.5)
15	641.2	201.1 (×3.2)	117.3	94.0 (×1.2)
35	641.2	240.6 (×2.7)	117.3	103.8 (×1.1)
63	641.2	296.4 (×2.1)	117.3	106.8 (×1.1)

TABLE III
TIME IN SECSONDES TO COMPUTE THE CLEAN ALGORITHM.

V. CONCLUSION

This paper shows a decomposition of the gridding and degridding operators used in radio interferometry imaging problems. We proposed a sky to sky (S2S) method to reduce the computing cost without additional error when these operators are used successively. While the FHD [19] algorithm shows a high acceleration factor at the price of excessive memory consumption, the S2S method allows a reduction of the memory cost while having an acceleration factor up to 16.4 regarding the standard method. The estimation of a sky, simulated with an 8 hours observation with the VLA telescope, has shown an acceleration factor up to 3.4 for the full reconstruction algorithm. Finally, our experiments have shown that the S2S method is always computationally more efficient than the standard method for the same gridding parameters. In future work, we plan to enhance the capabilities of the S2S method with code optimization and a GPU implementation. Moreover, the S2S method makes gradient descent algorithms computationally doable. We plan to explore this field and develop a preconditioner to reduce the number of gradient evaluations required.

ACKNOWLEDGMENT

This work was supported by grants from Région Ile-de-France and DARK-ERA (ANR-20-CE46-0001-01).

REFERENCES

[1] A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and synthesis in radio astronomy*, 2nd ed. Wiley, 2001.

[2] E. Thiébaut and J. Young, "Principles of image reconstruction in optical interferometry: tutorial," *Journal of the Optical Society of America A*, vol. 34, no. 6, p. 904, 2017.

[3] M. P. van Haarlem and al., "LOFAR: The LOw-frequency ARray," *Astronomy & Astrophysics*, vol. 556, p. A2, 2013.

[4] P. Dewdney, P. Hall, R. Schilizzi, and T. Lazio, "The square kilometre array," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1482–1496, 2009.

[5] J. Högbom, "Apperture synthesis with a non-regular distribution of interferometer baselines," *Astronomy and Astrophysics*, vol. Vol. 15, p. p.417, 1974.

[6] B. Clark, "An efficient implementation of the algorithm 'CLEAN'," *Astronomy and Astrophysics*, vol. 89, p. 377, 1980.

[7] C. Tasse, B. Hugo, M. Mirmont, O. Smirnov, M. Atemkeng, L. Bester, M. J. Hardcastle, R. Lakhoo, S. Perkins, and T. Shimwell, "Faceting for direction-dependent spectral deconvolution," *Astronomy & Astrophysics*, vol. 611, p. A87, 2018.

[8] A. R. Offringa and al., "WSClean: an implementation of a fast, generic wide-field imager for radio astronomy," *Monthly Notices of the Royal Astronomical Society*, vol. 444, no. 1, pp. 606–619, 2014.

[9] R. Carrillo, J. McEwen, and Y. Wiaux, "Purify: a new approach to radio-interferometric imaging," *Monthly Notices of the Royal Astronomical Society*, vol. 439, no. 4, pp. 3591–3604, 2014.

[10] Y. Wiaux, L. Jacques, G. Puy, A. M. M. Scaife, and P. Vandergheynst, "Compressed sensing imaging techniques for radio interferometry," *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 3, pp. 1733–1742, 2009.

[11] S. Bagchi and S. K. Mitra, "The nonuniform discrete fourier transform and its applications in filter design: Part i-1-d," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 43, no. 6, pp. 422–433, 1996.

[12] J. Fessler and B. Sutton, "Nonuniform fast fourier transforms using min-max interpolation," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 560–574, Feb. 2003.

[13] C. G. Radhkrishna, P. Weiss, G. Daval-Frérot, A. Massire, A. Vignaud, and P. Ciuciu, "Optimizing full 3D SPARKLING trajectories for high-resolution T2*-weighted Magnetic Resonance Imaging," *IEEE Transactions on Medical Imaging*, p. 17, 2022.

[14] J. A. Fessler, "On NUFFT-based gridding for non-cartesian MRI," *Journal of Magnetic Resonance*, vol. 188, no. 2, pp. 191–195, 2007.

[15] P. Beatty, D. Nishimura, and J. Pauly, "Rapid gridding reconstruction with a minimal oversampling ratio," *IEEE Transactions on Medical Imaging*, vol. 24, no. 6, pp. 799–808, 2005.

[16] P. Thévenaz, T. Blu, and M. Unser, "Image interpolation and resampling," p. 39.

[17] J.-F. Giovannelli and A. Coulais, "Positive deconvolution for superimposed extended source and point sources," *Astronomy & Astrophysics*, vol. 439, no. 1, pp. 401–412, 2005.

[18] H. L. Bester, A. Repetti, S. Perkins, O. M. Smirnov, and J. S. Kenyon, "A practical preconditioner for wide-field continuum imaging of radio interferometric data," 2021.

[19] I. Sullivan and al., "Fast holographic deconvolution: a new technique for precision radio interferometry," *The Astrophysical Journal*, vol. 759, no. 1, p. 17, 2012.

[20] A. R. Offringa, F. Mertens, S. van der Tol, B. Veenboer, B. K. Gehlot, L. V. E. Koopmans, and M. Mevius, "Precision requirements for interferometric gridding in the analysis of a 21 cm power spectrum," *Astronomy & Astrophysics*, vol. 631, p. A12, 2019.

[21] F. R. Schwab, "Relaxing the isoplanatism assumption in self-calibration; applications to low-frequency radio interferometry," *The Astronomical Journal*, vol. 89, p. 1076, 1984.

[22] U. Rau, S. Bhatnagar, M. Voronkov, and T. Cornwell, "Advances in calibration and imaging techniques in radio interferometry," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1472–1481, 2009.