Distributed image reconstruction for very large arrays in radio astronomy

André Ferrari, David Mary, Rémi Flamary and Cédric Richard
Laboratoire Joseph-Louis Lagrange
Université de Nice Sophia-Antipolis, CNRS, Observatoire de la Côte d'Azur
Nice, France
Email: surname.name@unice.fr

Abstract—Current and future radio interferometric arrays such as LOFAR and SKA are characterized by a paradox. Their large number of receptors (up to millions) allow theoretically unprecedented high imaging resolution. In the same time, the ultra massive amounts of samples makes the data transfer and computational loads (correlation and calibration) order of magnitudes too high to allow any currently existing image reconstruction algorithm to achieve, or even approach, the theoretical resolution. We investigate here decentralized and distributed image reconstruction strategies which select, transfer and process only a fraction of the total data. The loss in MSE incurred by the proposed approach is evaluated theoretically and numerically on simple test cases.

I. INTRODUCTION

Since the commissioning of the first large radio interferometers in the 70s and 80s (such as the VLA in the USA and the WSRT) radio astronomy in the range of large wavelengths has grown dramatically, particularly with the development of more and more extended antenna arrays. In the prospect of the most sensitive radio telescope ever built, the SKA which will be operational in the 2020s, several new generation radio telescopes are being built or planned (LOFAR in the Netherlands, ASKAP and the Murchison Widefield Array Australia, e-MERLIN in the UK, e-EVN based in Europe, MeerKAT in South Africa, JVLA the United States).

As an example, LOFAR consists of 48 groups of antennas (stations), among which approximately 35,000 elementary antennas are located in the Netherlands. The "superterp", the heart of LOFAR is a super-station: a cluster of six stations. Eight other stations, totalizing approximately 13,000 antennas are located in the surrounding countries. A project of a new super-station in Nançay (F) is under consideration. Within each station, antennas form a phased array which allows for digital beamforming simultaneously in several directions and frequency bands. The beam-formed data from the stations are centralized at the University of Groningen in the Netherlands where a supercomputer is responsible for the combination of the beam data from all stations. The resulting data are then stored on a cluster of ASTRON, the Netherlands Institute for Radio Astronomy, where the images (and other deliverables) are reconstructed. As a mean of comparison SKA will totalize 2.5 millions antennas, with a square kilometer collecting area distributed over an area of $\approx$ 5,000 km diameter.

Beyond specific objectives that distinguish these new fully digital "software telescopes", they are all characterized by a great flexibility. Another common point is the amount of data which must be transferred to the central computer and processed. It amounts to 1 terabit/second for LOFAR and will be of the order of 14 exabyte/day

for SKA (more than 100 times the global internet traffic). LOFAR uses a 1.5 Blue Gene/P for the data reduction and the computation of correlations. IBM et ASTRON will develop by 2024 a supercomputer to process and store 1 petabytes of data everyday [2].

This correspondence investigates the possibility to distribute the image reconstruction over the super-stations. The main objective is to avoid centralization of the sampled electromagnetic fields acquired by all stations in order to reduce the data transfer and the exponential increase in the calibration and computational load.

Section II recalls the basis of radio astronomy with aperture synthesis and proposes a strategy where each super-station uses all its antenna and one reference signal from the other super-stations. The loss of performances that follows is evaluated on a simple model using the Cramér Rao Lower Bound (CRLB). Section III shows that the image reconstruction problem can be written as a global variable consensus problem with regularization. Numerical simulations illustrate the performances of the proposed approach. A concluding section presents perpectives.

II. APERTURE SYNTHESIS FOR RADIO ASTRONOMY

A. Standard aperture synthesis model

This section provides the basic equations of radio astronomy with multiple sensor array and describes a partial aperture synthesis strategy which aims to reduce data transfer, allowing a decentralized image reconstruction.

To simplify the notations and without loss of generality, we will not make explicit the wavelengths dependence and the Earth rotation and assume punctual antennas. The coordinates of the stations (within each station, a beam is created from the phased array) in a plane perpendicular to the line of sight are denoted as $\boldsymbol{r}_j$ and the map of interest (the "image" of a region of the sky) is $x(\boldsymbol{p})$ where $\boldsymbol{p}$ denotes the angular coordinates on the sky. The fundamental equation of interferometry relates the Fourier transform of the map to the spatial coherency (visibility) of the incoming electromagnetic field. A measurement of the coherency is obtained by correlating the signal acquired by a pair of stations (i, j) properly delayed located at $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$, giving in the noiseless case a point of visibility at spatial frequency $\boldsymbol{u}_\ell = \boldsymbol{r}_j - \boldsymbol{r}_i$:

$$v(\boldsymbol{u}_\ell) = \int x(\boldsymbol{p}) e^{-j2\pi \boldsymbol{u}_\ell^t \boldsymbol{p}} d\boldsymbol{p} \qquad (1)$$

See for example [9], [10] for a comprehensive description of radio astronomy and signal processing related tools.

Computation of $v(\boldsymbol{u}_\ell)$ obviously requires the transfer of signals from stations i and j in the same place. The stations are normally grouped in "super-stations" (e.g. the superterp for LOFAR) accounting for low frequencies ($\|\boldsymbol{r}_i - \boldsymbol{r}_j\|_2$ small). Resolution is

then increased by correlating signals between stations that can be located up to thousands of kilometers from each other.

B. Reduced synthesis aperture model

In general, all possible correlations are computed in order to maximize the (u, v) coverage (set of $\boldsymbol{u}_l$). This requires a centralized system architecture and the resulting measurements correspond to a filtering of the visibility function $v(\boldsymbol{u})$ by the Full Aperture (FA) spatial transfer function

$$\mathcal{A}_F(\boldsymbol{u}) = \left(\sum_k \delta(\boldsymbol{r} - \boldsymbol{r}_k) * \sum_l \delta(\boldsymbol{r} - \boldsymbol{r}_l) \right)(\boldsymbol{u}) \tag{2}$$

$$= \sum_{\ell=1}^{M} w_\ell \delta(\boldsymbol{u} - \boldsymbol{u}_\ell) = \mathcal{A}_L(\boldsymbol{u}) + \mathcal{A}_H^F(\boldsymbol{u}) \tag{3}$$

where the weights w_ℓ count the number of beam pairs measuring the same spatial frequency $\boldsymbol{u}_\ell$ and M is the number of different sampled frequencies. The term $\mathcal{A}_L(\boldsymbol{u})$ is associated to "intra-super-stations" low-frequency correlations and $\mathcal{A}_H^F(\boldsymbol{u})$ to high-frequency "inter-super-stations" correlations. Note that in order to simplify the derivations, we will also denote super-station a single remote station.

In order to reduce the amount of transferred data, we propose to investigate a solution which consists in:

1) exploiting all the low-frequencies by computing locally all the correlations inside each super-station. In this case, the low frequency term associated to the spatial transfer function (denoted as $\mathcal{A}_R(\boldsymbol{u})$) is still $\mathcal{A}_L(\boldsymbol{u})$. Denote as $\mathcal{S}_k$ the set of indices associated to the beams in super-station k.

$$\mathcal{A}_{l,l}(\boldsymbol{u}) = \sum_{m,n \in \mathcal{S}_l} \delta(\boldsymbol{u} - (\boldsymbol{r}_m - \boldsymbol{r}_n)) + \delta(\boldsymbol{u} + (\boldsymbol{r}_m - \boldsymbol{r}_n))$$

$$\mathcal{A}_L(\boldsymbol{u}) = \sum_l \mathcal{A}_{l,l}(\boldsymbol{u})$$

2) Recovering the high frequency information by transferring only one single beam signal from each super-station to all other remote super-stations. If c_k is the index of the reference beam in super-station k transferred to the other super-stations, the resulting sampling pattern of visibilities associated to super-stations k and l is, see Fig. 1:

$$\mathcal{A}_{k,l}(\boldsymbol{u}) = \sum_{m \in \mathcal{S}_k} \delta(\boldsymbol{u} - (\boldsymbol{r}_m - \boldsymbol{r}_l))$$

$$\mathcal{A}_H^R(\boldsymbol{u}) = \sum_{k \neq l} \mathcal{A}_{k,l}(\boldsymbol{u})$$

This solution is motivated by the fact that it allows to correctly sample the high frequencies, at the cost of a reduced SNR, while reducing the number of transfered electromagnetic signals. Other strategies than transferring a single antenna signal are of course possible. A solution which aims to preserve the SNR is to replace the signal indexed by c_k by averaging of neighboring beams. It is important to emphasize that in this case the bias introduced by the averaging of nearby beams must be negligible. This gain in SNR will be denoted as $\rho \geq 1$ in the sequel. Finally, the overall Reduced Aperture (RA) antenna spatial transfer function is:

$$\mathcal{A}_R(\boldsymbol{u}) = \mathcal{A}_L(\boldsymbol{u}) + \rho \mathcal{A}_H^R(\boldsymbol{u}) \tag{4}$$

This strategy reduces the transfer of beams data w.r.t. a centralized processing as long as the number of stations inside each super-station is larger than the number of super-stations: for N_{ss} super-stations of N_s stations each, the first requires $N_{ss}(N_{ss} - 1)$ transfers whereas the second $N_{ss}N_s$.
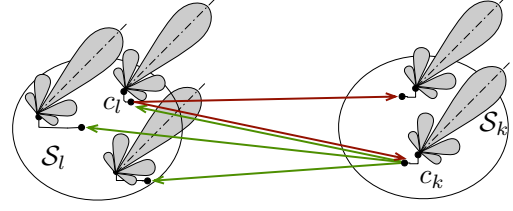

Fig. 1. Reduced Aperture (RA) synthesis using super-stations k and l.

C. Analysis of performances on a simple model

In order to evaluate analytically the loss related to the use of RA synthesis *w.r.t.* to a FA we consider a simple one dimensional case where the map is a shifted Gaussian shape with flux α:

$$x(p) = \frac{\alpha}{\eta\sqrt{2\pi}} e^{-\frac{(p-p_0)^2}{2\eta^2}} \tag{5}$$

The unknown parameters are $\boldsymbol{\theta} = (\alpha, \eta, p_0)$. The visibilities are:

$$v(u_\ell) = \alpha e^{j2\pi u_\ell p_0} e^{-2\pi^2 \eta^2 u_\ell^2} + n_\ell, \ \ell = 1 \dots M \tag{6}$$

where n_ℓ is a measurement noise assumed independent Gaussian circular with $n_\ell \sim \mathcal{N}_c(0, w_\ell^{-1}\sigma^2)$. The coefficient w_ℓ^{-1} takes into account the variance reduction that occurs when the visibility $v(u_\ell)$ is estimated from w_ℓ different baselines.

The elements of the Fisher information matrix $I(\boldsymbol{\theta})$ are computed using the Slepian-Bangs formula [8, p. 293] which gives:

$$I(\boldsymbol{\theta}) = \frac{2}{\sigma^2} \begin{pmatrix} S_0 & -2\pi^2 \alpha S_2 & 0 \\ -2\pi^2 \alpha S_2 & 4\pi^4 \alpha^2 S_4 & 0 \\ 0 & 0 & 4\pi^2 \alpha^2 S_2 \end{pmatrix} \tag{7}$$

$$S_q = \sum_{\ell=1}^{M} w_\ell u_\ell^q e^{-4\pi^2 \eta^2 u_\ell^2} \tag{8}$$

Note that $I(\boldsymbol{\theta})$ is not a function of the source position p_0.

We compare the CRLBs on α, η and p_0 for two spatial transfer functions. In both cases the aperture configuration consists of two uniform sub-apertures separated by D in order to sketch the behaviour of two super-stations:

$$r_k = -D/2 - k\Delta, \ k = 0 \dots L \tag{9}$$

$$r_{L+k+1} = D/2 + k\Delta, \ k = 0 \dots L \tag{10}$$

where $D > (L + 1)\Delta$. In the FA mode, and for $u \geq 0$:

$$\mathcal{A}_L(u) = \sum_{\ell=0}^{L} 2(L + 1 - \ell)\delta(u - \ell\Delta)$$

$$\mathcal{A}_H^F(u) = \sum_{\ell=-L}^{L} (L + 1 - |\ell|)\delta(u - (D + (\ell + L)\Delta))$$

L is assumed even, $L = 2q$ and we consider in the RA mode that $c_k = \pm q$: the reference beam is in the middle of the opposite super-station. As noted above, the low frequency term $A_L(u)$ does not change. The high frequency spatial transfer function is now for $u \geq 0$:

$$\mathcal{A}_H^R(u) = \sum_{\ell=0}^{L} \delta(u - (D + (\ell + q)\Delta))$$

The source width η plays a central role in the estimation. For a point source, $\eta \to 0$, high frequency measurements will bring a lot of information while performances for a very extended source ($\eta \to \infty$) will be independent of the inter-stations visibilities. Figs. 2
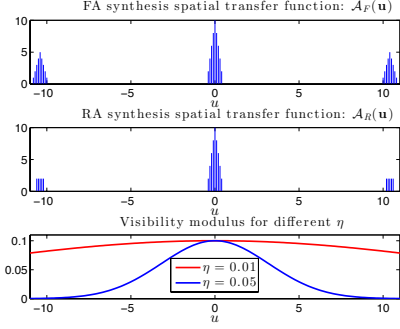
Fig. 2. 1D illustration of $\mathcal{A}_F(u)$ and $\mathcal{A}_R(u)$. Bottom plot shows $|v(u)|$ for 2 characteristic values of η.
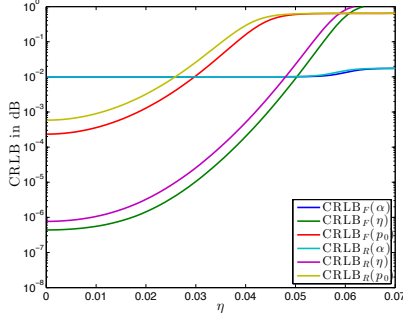

Fig. 3. CRLB for the FA: $\mathrm{CRLB}_F(\cdot)$, and the RA: $\mathrm{CRLB}_R(\cdot)$.

and 3 give the results obtained for a configuration defined by $L = 4$, $D = 10$ and $\Delta = 0.1$. The source parameters are $\alpha = 1$, $p_0 = 0$ and results are given for different values of η. The gain ρ is fixed to $\rho = 2$. Fig. 2 shows the two spatial transfer functions $\mathcal{A}_F(u)$ and $\mathcal{A}_R(u)$. Fig. 3 shows the CRLBs associated to $A_F(u)$ and $A_R(u)$, denoted respectively as CRLB_F and CRLB_R. The thresholding effect when $\eta \approx 0.05$ reflects the shape of the visibility modulus given in Fig. 2: the frequency contribution of the source at the inter-station baseline becomes negligible for $\eta > 0.05$. For $\eta < 0.05$ the order of magnitude of the loss of performances is 4dB. This loss of performance naturally strongly depends on ρ, e.g. when $\rho = 1$ the loss is 13dB.

III. DISTRIBUTED IMAGE RECONSTRUCTION WITH PARTIAL APERTURE SYNTHESIS MODEL

A. Decentralized map reconstruction

The classical model for the map reconstruction is obtained vectorizing the sampled map $\boldsymbol{x} \in (\mathbb{R}^+)^N$ and the visibilities $\boldsymbol{v} \in \mathbb{C}^M$, $M < N$, and reads [5], [6]:

$$\boldsymbol{v} = \boldsymbol{G}\boldsymbol{x} + \boldsymbol{n}, \quad \boldsymbol{G} = \boldsymbol{W}\boldsymbol{T}\boldsymbol{F} \tag{11}$$

where $\boldsymbol{F}$ is the $N \times N$ Fourier transform matrix, $\boldsymbol{W}$ is a diagonal weighting-matrix including various operations (calibration, signal to noise weighting), and $\boldsymbol{T}$ is a 0/1 binary $M \times N$ matrix which codes the sampling of the visibilities in the frequency plane. The noise vector is assumed $\boldsymbol{n} \sim \mathcal{N}_c(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$. With these assumptions, the image restoration problem is usually casted as the general inverse problem:

$$\hat{\boldsymbol{x}} = \arg \min_{\boldsymbol{x} \in (\mathbb{R}^+)^N} \|\boldsymbol{v} - \boldsymbol{G}\boldsymbol{x}\|_2^2 + \Omega(\boldsymbol{x}) \tag{12}$$

where the regularization term $\Omega(\boldsymbol{x})$ can be for example a sparse analysis [3], a sparse synthesis [11] or a hybrid prior [4].
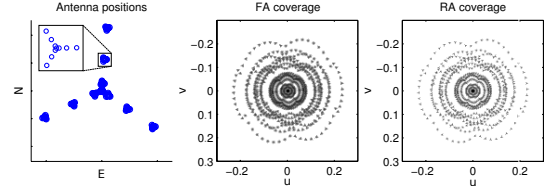

Fig. 4. Left: sensor array configuration (without Earth rotation). Final (u, v) coverage in normalized frequencies for FA, Eq. (3) and RA, Eq. (4).

Denoting $\boldsymbol{T}^\dagger \boldsymbol{v}$ the zero-padded visibilities and $\boldsymbol{y} = \boldsymbol{F}^\dagger \boldsymbol{T}^\dagger \boldsymbol{v}$ the so called "dirty image", (11) can be rewritten as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{e} \tag{13}$$

where $\boldsymbol{H} = \boldsymbol{F}^\dagger \boldsymbol{T}^\dagger \boldsymbol{W} \boldsymbol{T} \boldsymbol{F}$ is a (circulant) convolution matrix. Using $\forall \boldsymbol{z} \ \|\boldsymbol{T}^\dagger \boldsymbol{z}\|_2 = \|\boldsymbol{z}\|_2$ and $\boldsymbol{F}^\dagger \boldsymbol{F} = \boldsymbol{I}$, we have $\|\boldsymbol{v} - \boldsymbol{G}\boldsymbol{x}\|_2 = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2$ and the inverse problem (12) turns to be equivalent to the deconvolution problem

$$\hat{\boldsymbol{x}} = \arg \min_{\boldsymbol{x} \in (\mathbb{R}^+)^N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \Omega(\boldsymbol{x}) \tag{14}$$

The vector $\boldsymbol{v}$ can be partitioned as:

$$\boldsymbol{v}^\dagger = (\boldsymbol{v}_{1,\cdot}^\dagger, \boldsymbol{v}_{2,\cdot}^\dagger, \ldots), \ \boldsymbol{v}_{k,\cdot}^\dagger = (\boldsymbol{v}_{k,1}^\dagger, \boldsymbol{v}_{k,2}^\dagger, \ldots) \tag{15}$$

where $\boldsymbol{v}_{k,k}$ denotes the vector of visibilities obtained by correlating all the beams inside super-station k and vector $\boldsymbol{v}_{k,l}$, $k \neq l$, by correlating beams in station k with reference beam indexed by c_l in station l. More specifically $\boldsymbol{v}_{k,k}$ is associated to $\mathcal{A}_{k,k}(\boldsymbol{u})$ and $\boldsymbol{v}_{k,l}$ is associated to $\mathcal{A}_{k,l}(\boldsymbol{u})$, see sec. II.B. Using the same partitioning for $\boldsymbol{G}$, $\boldsymbol{W}$ and $\boldsymbol{T}$, Eq (12) can be rewritten as :

$$\hat{\boldsymbol{x}} = \arg \min_{\boldsymbol{x} \in (\mathbb{R}^+)^N} \sum_k \|\boldsymbol{v}_{k,\cdot} - \boldsymbol{G}_k \boldsymbol{x}\|_2^2 + \Omega(\boldsymbol{x}) \tag{16}$$

with $\boldsymbol{G}_k = \boldsymbol{W}_k \boldsymbol{T}_k$. Using again the property of the zero padding matrix $\boldsymbol{T}_{k,\cdot}^\dagger$ inside each norm in the sum (16), we obtain:

$$\hat{\boldsymbol{x}} = \arg \min_{\boldsymbol{x} \in (\mathbb{R}^+)^N} \sum_k \|\boldsymbol{y}_{k,\cdot} - \boldsymbol{H}_k \boldsymbol{x}\|_2^2 + \Omega(\boldsymbol{x}) \tag{17}$$

where $\boldsymbol{H}_k = \boldsymbol{F}^\dagger \boldsymbol{T}_k^\dagger \boldsymbol{W}_k \boldsymbol{T}_k \boldsymbol{F}$. Note that $\boldsymbol{H} = \sum_k \boldsymbol{H}_k$ and $\boldsymbol{y} = \sum_k \boldsymbol{y}_k$.

In (16), each sub-problem in the sum amounts to reconstruct $\boldsymbol{x}$ from the intra-super-station and inter-super-station visibilities $\boldsymbol{v}_{k,\cdot}$. In (17), $\boldsymbol{y}_{k,\cdot}$ is a dirty image obtained using only the visibilities $\boldsymbol{v}_{k,\cdot}$ and each sub-problem amounts to reconstruct $\boldsymbol{x}$ from $\boldsymbol{y}_{k,\cdot}$.

Eqs. (16,17) are particularly interesting for the derivation of distributed optimization algorithm, since they correspond to a global variable consensus problem [1], [7]. The next subsections evaluate the impact of the partial aperture models *w.r.t.* the standard model by numerical simulations.

B. Array configuration and aperture synthesis

The shape of the sensor array used in the simulations consists of 10 super-stations of 10 stations each, as shown in Fig. 4. The stations follow a classical Y configuration with different rotations. Ten measurements are performed in a range of 3 hours taking into account the Earth rotation. The RA is obtained by computing for each station the correlation with the center beam of the other super-stations and with $\rho = 2$. For this experiment the ratio $D/(L\Delta)$ is 20 and the visibilities are binned in a 256×256 grid with Shannon sampling. The corresponding (u, v) coverage is shown in Fig. 4.
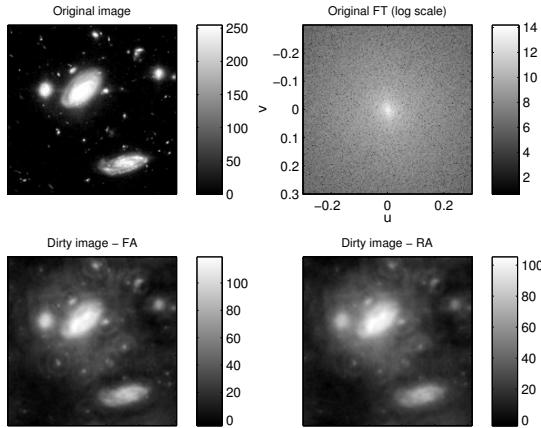
Fig. 5. Original image and its Fourier transform (up). Dirty images observed with the full aperture (down left) and reduced aperture (down right).

Fig. 6. Reconstructed images and error ϵ. Up (down): reconstruction without (with) positivity constraint. Left : FA mode, right : RA mode.

The original image is shown in the upper part of Fig. 5 along with its Fourier transform. Note that the image contains both low and high frequency. The noise of variance σ^2 is such that the measured visibilities SNR is 70 dB. The observed dirty images are also reported. The RA clearly leads to a less detailed dirty image.

C. Distributed image reconstruction

The image reconstruction is performed by solving problem (17) with a regularized global variable consensus ADMM algorithm as described in [1]. The regularization term is $\Omega(\boldsymbol{x}) = \lambda\|\boldsymbol{x}\|_2^2$ with $\lambda = 10^{-6}$. This regularization ensures that the problem is strictly convex and limits the bias.

At each iteration, this algorithm requires to solve a large scale linear problem of size the number of pixels N in each super-station, [1, Eq. (7.6)]. This linear problem can be easily solved in the Fourier domain as discussed in [5]. Note that the quadratic regularization can be included in this step. The consensus step [1, Eq. (7.7)] is simply a projection on the positive orthant. A super-station sends the current image, Lagrangian multiplier and receive the consensus image.

Fig. 6 shows the reconstructed images obtained by solving (17) with and without positivity constraints for both aperture cases, along with the relative norm of the error ϵ. The FA leads obviously to a better reconstruction in both cases. However, while the loss in performance is rather important in the unconstrained case ($\epsilon = 15\%$ error instead of $\epsilon = 23\%$), this gap is significantly reduced with the positivity constraint ($\epsilon = 9.5\%$ error instead of $\epsilon = 12\%$).

IV. CONCLUSION AND PERSPECTIVES

This correspondence investigates a distributed strategy for the image reconstruction problem in radio astronomy when the number of stations inside each super-station is larger than the number of super-stations. It relies on a reduced aperture synthesis where each super-station uses all its beam and a single reference beam from the other super-stations. Part of the missing information for each super-station is then exchanged during the consensus step of the distributed algorithm. The loss of performances, related to the use of a reduced aperture synthesis, is evaluated on the image reconstruction by computer simulations.

The approach proposed in the paper processes all the data after they have been acquired. A natural extension is to reconstruct sequentially the image. Among the benefits of this setup which perfectly fits the operating mode of 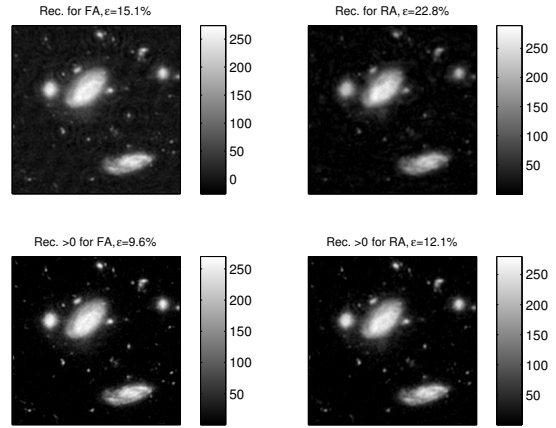interferometers which progressively fill the frequency plane using the Earth rotation, is the possibility to optimize in real-time the observation mode. A straightforward solution is to make a number of iterations of the reconstruction algorithm of section III-C after each measurement and using a "warm start". A much more challenging perspective it to select sequentially the frequency measurements used at each iteration. Whereas this communication relies on a "deterministic" pattern of frequency measurements, a better strategy would be to select at each iteration, among all visibilities, the ones that optimize the reconstruction according to some predefined criterion.

REFERENCES

[1] S. Boyd, N. Parikh, and E. Chu. Distributed Optimization and Statistical Learning Via the Alternating Direction Method of Multipliers. In M. Jordan, editor, *Foundations and Trends in Machine Learning*, page 140. Now Publishers Inc, June 2011.

[2] P. C. Broekema, A.-J. Boonstra, V. C. Cabezas, et al. DOME: towards the ASTRON & IBM center for exascale technology. In *Astro-HPC '12: Proceedings of the 2012 workshop on High-Performance Computing for Astronomy Date*. ACM Request Permissions, June 2012.

[3] R. E. Carrillo, J. D. McEwen, and Y. Wiaux. Sparsity Averaging Reweighted Analysis (SARA): a novel algorithm for radio-interferometric imaging. *Monthly Notices of the Royal Astronomical Society*, 426(2):1223–1234, Oct. 2012.

[4] A. Dabbech, D. Mary, and C. Ferrari. Astronomical image deconvolution using sparse priors: An analysis-by-synthesis approach. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 3665–3668, 2012.

[5] J. F. Giovannelli and A. Coulais. Positive deconvolution for superimposed extended source and point sources. *Astronomy and Astrophysics*, 439(1):401–412, Aug. 2005.

[6] U. Rau, S. Bhatnagar, M. A. Voronkov, and T. J. Cornwell. Advances in calibration and imaging techniques in radio interferometry. *Proceedings of the IEEE*, 97(8):1472–1481, 2009.

[7] A. H. Sayed. Diffusion Adaptation over Networks. In R. Chellapa and S. Theodoridis, editors, *E-Reference Signal Processing*. Elsevier, May 2012.

[8] P. Stoica and R. Moses. *Spectral analysis of signals*. Pearson Prentice Hall, 2005.

[9] A. R. Thompson, J. M. Moran, and J. George W Swenson. *Interferometry and Synthesis in Radio Astronomy*. John Wiley & Sons, Nov. 2008.

[10] A.-J. van der Veen and S. J. Wijnholds. *Signal Processing Tools for Radio Astronomy*. Handbook of Signal Processing Systems. Springer, 2013.

[11] 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*, 395(3):1733–1742, May 2009.