SPARSITY IN TENSOR OPTIMIZATION FOR OPTICAL-INTERFEROMETRIC IMAGING

Anna Auría⋆, Rafael E. Carrillo⋆, Jean-Philippe Thiran⋆,† and Yves Wiaux‡

⋆Signal Processing Laboratory (LTS5), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
† University of Lausanne (UNIL), CH-1011 Lausanne, Switzerland
‡ Institute of Sensors, Signals, and Systems, Heriot-Watt University, EH14 4AS, UK

ABSTRACT

Image recovery in optical interferometry is an ill-posed nonlinear inverse problem arising from incomplete power spectrum and bispectrum measurements. We review our previous work, which reformulates this nonlinear problem in the framework of tensor recovery and studies two different approaches to solve it: one is nonlinear and nonconvex while the other is linear and convex. We extend the linear convex procedure to account for signal sparsity and we also present numerical simulations that show the improvement in the quality of reconstruction of sparse images when including a sparsity prior.

Index Terms— optical interferometry, tensor optimization, interferometric imaging, phase retrieval.

1. INTRODUCTION

Optical interferometers are being developed to provide high resolution in the optical spectrum. Since their output is not directly an image, tools for image reconstruction are required. Ideally interferometers take measurements that identify with Fourier coefficients of the intensity image $\boldsymbol{x}$ of interest, the so-called visibilities. Each telescope pair at a given observation time identifies a spatial frequency in the image plane and its visibility can be interpreted as the Fourier transform of the image of interest at that particular frequency. When these visibilities are accessible – as for radio interferometry – the problem is linear. At optical wavelengths, atmospherical turbulence induces random phase delays and only power spectrum information can yet be retrieved, together with partial phase information through phase closure or bispectrum measurements [1, 2, 3]. This poses a nonlinear ill-posed inverse problem for image reconstruction in optical interferometry, which is very sensitive to the optimization strategy.

The state-of-the-art method MiRA [4] recovers the image as a maximum a posteriori (MAP) solution of an optimization problem with an objective function of this form $f(\boldsymbol{x}) = f_{\text{data}}(\boldsymbol{x}) + \ell f_{\text{prior}}(\boldsymbol{x})$, for some parameter ℓ to be tuned and with additional positivity and total flux constraints. In [4, 5] sparsity priors have specifically been promoted. MiRA solves the optimization problem by minimizing a nonconvex objective function locally, leading to a solution which is strongly dependent on initialization. The WISARD alternative [6] takes an alternate minimization self-calibration approach that in a first nonconvex step estimates the pseudocomplex visibilities from which the image is subsequently recovered through a linear inverse problem, as for radio interferometry. This overall procedure remains nonconvex and the final solution also depends on the initial guess. In summary, state-of-the-art methods are nonconvex due to the intrinsic data nonlinearity [1], and therefore known to suffer from a strong sensitivity to initialization.

The approaches proposed here are inspired by the convex optimization methods that regularize the linear inverse problem in radio interferometry through sparsity constraints in the framework of compressed sensing [7, 8, 9, 10, 11, 12, 13]. We first review recent work [14], in which we reformulate this nonlinear problem as a linear problem for the supersymmetric rank-1 order-3 tensor formed by the tensor product of the vector representing the image under scrutiny with itself. We summarize two different approaches to solve this problem: a convex approach for tensor recovery with built-in supersymmetry and a nonlinear nonconvex approach formulating the problem for the tensor product of 3 vectors, where supersymmetry is relaxed while the rank-1 constraint is built-in. In contrast with the state of the art though, in this last case the problem is solved through an iterative alternate minimization scheme whose subproblems are all linear and convex, also enforcing reality and positivity of the vectors. We finally build on the work in [14] and study the inclusion of a sparsity prior for the linear approach. Assuming image-space sparsity as a proof of concept leads to solving a sequence of weighted ℓ_1-minimization problems. Numerical simulations show the improvement in the quality of the reconstruction of sparse images when we apply this reweighting scheme compared to the solution without considering sparsity as a prior.

2. DATA MODEL AND TENSOR FORMULATION

Throughout this paper, we keep the same data model and tensor formulation of the optical-interferometric imaging problem as in our previous work [14]. To make the present article self-contained, in this section we recapitulate the description of the notation, the data model and the problem formulation. The reader can find more details in [14].

We adopt a discrete setting where the intensity image of interest is represented by the real and positive vector $\boldsymbol{x} \in \mathbb{R}_+^N$ with components x_i. Its 2D discrete Fourier transform is denoted $\hat{\boldsymbol{x}} \in \mathbb{C}^N$ with components $\hat{x}_i$. We denote $\hat{x}_{-i}$ the component of $\hat{\boldsymbol{x}}$ at the opposite spatial frequency to that associated with $\hat{x}_i$. We consider generic optical-interferometric measurements of the form $\hat{x}_i \hat{x}_j \hat{x}_k$, performed on discrete spatial frequencies, keeping in mind that in a real scenario the model should be able to account for continuous frequencies [15]. In this setting, power spectrum measurements correspond to triplets with $j = -i$ and $k = 0$ ($\hat{x}_0$ stands for the Fourier coefficient at zero frequency). Bispectrum measurements would correspond to triplets whose spatial frequencies associated with $\hat{x}_i$, $\hat{x}_j$, and $\hat{x}_k$ sum to zero, even though we choose to relax this closure constraint. The measurement equation can be written in compact form as:

$$\boldsymbol{y} = \mathcal{V}(\boldsymbol{x}) + \boldsymbol{n}, \tag{1}$$

where $\mathcal{V}$ is a nonlinear operator providing an undersampled set of

triple products of Fourier coefficients of $\boldsymbol{x}$. The measurement vector $\boldsymbol{y} \in \mathbb{C}^M$, with components y_a ($1 \leq a \leq M$) and dimension typically smaller than the signal dimension, $M < N$, is assumed to be affected by a simple noise vector $\boldsymbol{n} \in \mathbb{C}^M$ with i.i.d. Gaussian components n_a. We also assume that the total flux is measured independently and consider a normalized signal such that $\sum_i x_i = \hat{x}_0 = 1$.

The measurement model (1) can be recast as the following linear model for the real and positive supersymmetric rank-1 order-3 tensor $\mathcal{X} = \boldsymbol{x} \circ \boldsymbol{x} \circ \boldsymbol{x} \in \mathbb{R}_+^{N \times N \times N}$:

$$\boldsymbol{y} = \mathcal{T}(\mathcal{X}) + \boldsymbol{n}, \tag{2}$$

where the linear operator $\mathcal{T}$ performs a 2D discrete Fourier transform along each of the 3 dimensions, identified by an operator $\mathcal{F}$, followed by a selection and vectorization operator $\mathcal{M}$ resulting in variable-density undersampling in a 6D Fourier space: $\mathcal{T} = \mathcal{M}\mathcal{F}$. The mask includes the measurement corresponding to the "triple-zero frequency", $\hat{x}_0^3 = 1$, to approximately enforce the flux normalization of the signal. This formulation is proposed as a generalization of the Phase Lift approach [16]. In that framework, quadratic measurements of the form $|\langle \boldsymbol{x}, \boldsymbol{a_i} \rangle|^2$, are seen as linear measurements on the rank-1 matrix $\mathsf{X} = \boldsymbol{x}\boldsymbol{x}^\dagger$ representing the outer product of the signal with itself († stands for the conjugate-transpose operation). In the Phase Lift approach, the ambient dimension of the problem of phase recovery is therefore *lifted* so that it can be formulated as a linear problem. Analogously, we remark that the intrinsic dimension of the tensor formulation (2) is drastically increased. Rank-1 and supersymmetry properties are not explicitly imposed in the general formulation (2), that consequently solves for an unknown $\mathcal{X}$ of size N^3, instead of the original $\boldsymbol{x}$ of size N in (1). In the following sections, the two different regularization schemes for tensor recovery presented in [14] are reviewed and an extension of the linear approach that considers a sparsity prior is proposed.

3. NUCLEAR VERSUS ALTERNATE MINIMIZATION

3.1. Rank-1 Alternate Minimization (AM)

By relaxing the supersymmetry while keeping an explicit rank-1 constraint in (2), we consider the following formulation: $\boldsymbol{y} = \mathcal{T}(\boldsymbol{u_1} \circ \boldsymbol{u_2} \circ \boldsymbol{u_3}) + \boldsymbol{n}$. The measurements can be seen as an undersampled set of products of Fourier coefficients of $\boldsymbol{u_1}$, $\boldsymbol{u_2}$, and $\boldsymbol{u_3}$, bringing back a nonlinear framework. We consider the following nonconvex minimization problem for tensor recovery:

$$\min_{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3} \in \mathbb{R}_+^N} \|\mathcal{T}(\boldsymbol{u_1} \circ \boldsymbol{u_2} \circ \boldsymbol{u_3}) - \boldsymbol{y}\|_2^2. \tag{3}$$

Despite the evident nonlinearity and nonconvexity of (3), an alternate minimization algorithm can be designed thanks to the nonsupersymmetric relaxation solving sequentially for each variable ($\boldsymbol{u_1}$, $\boldsymbol{u_2}$ or $\boldsymbol{u_3}$) while keeping the other two fixed and iterating until convergence, as suggested in [17]. At each iteration, the 3 linear and convex subproblems

$$\min_{\boldsymbol{u_p} \in \mathbb{R}_+^N} \|\mathsf{T}_{(\boldsymbol{u_q}\boldsymbol{u_s})}\boldsymbol{u_p} - \boldsymbol{y}\|_2^2, \tag{4}$$

are therefore solved sequentially for $1 \leq p \neq q \neq s \leq 3$, where the linear operators $\mathsf{T}_{(\boldsymbol{u_q}\boldsymbol{u_s})}$ are defined by $\mathsf{T}_{(\boldsymbol{u_q}\boldsymbol{u_s})}\boldsymbol{u_p} \equiv \mathcal{T}(\boldsymbol{u_p} \circ \boldsymbol{u_q} \circ \boldsymbol{u_s})$. In each subproblem the linear operator is computed using the values of the fixed variables at the current step. The final solution is heuristically chosen to be the mean of the 3 vectors $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, and $\boldsymbol{u}_3$ at convergence. The reader can find further the details on the algorithm used to solve this problem in [14].

3.2. Supersymmetric Nuclear Minimization (NM)

To embed tensor supersymmetry, we substitute the original measurement vector $\boldsymbol{y}$ by its replicated version $\mathcal{R}(\boldsymbol{y})$, and use a symmetrized version $\mathcal{M}_\mathrm{s}$ of the selection mask, ensuring that all permutations of a triplet (i, j, k) are assumed to be measured. The modified inverse problem thus reads as $\boldsymbol{y}_\mathrm{s} = \mathcal{T}_\mathrm{s}(\mathcal{X}) + \boldsymbol{n}_\mathrm{s}$, with $\boldsymbol{y}_\mathrm{s} = \mathcal{R}(\boldsymbol{y})$, $\boldsymbol{n}_\mathrm{s} = \mathcal{R}(\boldsymbol{n})$ and $\mathcal{T}_\mathrm{s} = \mathcal{M}_\mathrm{s}\mathcal{F}$ denoting the symmetrized versions of the measurement vector, noise vector and measurement operator, respectively. Without loss of generality, we assume that the initial selection operator $\mathcal{M}$ contains no redundant measurements, i.e. $i \leq j \leq k$. This ensures that $\mathcal{R}$ is well-defined. Also note that the noise statistics remains unaltered and only concerns the entries before replication.

The convex minimization problem to be defined will be regularized assuming low-rankness, reality and positivity as priors. The low-rankness of the tensor will be promoted through the minimization of the nuclear norm of an appropriately chosen matricization of $\mathcal{X}$, defined trough the operator $\mathcal{C}$, $\mathcal{C}(\mathcal{X}) \in \mathbb{C}^{N \times N}$ with $[\mathcal{C}(\mathcal{X})]_{ij} = \sum_k \mathcal{X}_{ijk}$, which performs the summation over one dimension of the tensor. This summation approach is *a priori* computationally efficient given the reduced size of $\mathcal{C}(\mathcal{X})$ compared to $\mathcal{X}$, as discussed in [14]. Positive-semidefiniteness of $\mathcal{C}(\mathcal{X})$, i.e. positivity of the eigenvalues, is also explicitly added as a convex prior, denoted $\mathcal{C}(\mathcal{X}) \succeq 0$, together with the convex reality and positivity constraints of $\mathcal{X}$: $\mathcal{X} \in \mathbb{R}_+^{N \times N \times N}$. The resulting convex nuclear norm minimization problem (NM) for $\mathcal{X}$ thus reads as:

$$\min_{\mathcal{X} \in S} \|\mathcal{C}(\mathcal{X})\|_* \quad \text{s. t.} \quad \|\boldsymbol{y}_\mathrm{s} - \mathcal{T}_\mathrm{s}(\mathcal{X})\|_2 \leq \epsilon, \tag{5}$$

where $S = S_1 \cap S_2$, with $S_1 = \mathbb{R}_+^{N \times N \times N}$ and $S_2 = \{\mathcal{X} \,|\, \mathcal{C}(\mathcal{X}) \succeq 0\}$ and $\epsilon^2 = 6(2M + 4\sqrt{M})\sigma_n^2/2$ is a likely bound on the symmetrized residual noise term. To extract the sought signal $\boldsymbol{x}_\mathrm{NM}$ from the tensor solution $\mathcal{X}_\mathrm{NM}$ we resort to the generic algorithm proposed by [18] to find the best rank-1 supersymmetric approximation of a supersymmetric tensor $\mathcal{X}_\mathrm{NM}$ in the least square sense. For further details on the algorithm the reader can refer to [14].

3.3. Numerical simulations

In this subsection we review the performance of the NM and AM algorithms through numerical simulations[1]. In all experiments random variable-density sampling in the 6D Fourier space – as described in [14] – is assumed and the input signal-to-noise ratio is defined as $\mathrm{ISNR} = -10\log(\sigma_n^2/e_y^2)$ where $e_y^2 = (1/M)\sum_a^M |y_a|^2$. The signal-to-noise ratio of a reconstruction $\bar{\boldsymbol{x}}$ is defined as $\mathrm{SNR} = -10\log(\|\bar{\boldsymbol{x}} - \boldsymbol{x}\|^2/\|\boldsymbol{x}\|^2)$.

Firstly, we evaluate the reconstruction quality on images constructed from 32 and 64 randomly located spikes. The graphs in Figure 1 represent the SNR curves as a function of undersampling in the range $[0.25, 1]$. A total of 50 and 10 simulations per point are performed for AM and NM respectively, varying the signal, as well as the sampling and noise realizations. The results show a clear superiority of AM relative to NM in terms of average reconstruction quality. Both approaches exhibit nonnegligible variability. The same conclusions are confirmed in Figure 2 in an identical setting on a realistic image representing low-resolution versions of the Eta Carinae star system[2]. The graph shows the SNR curves computed over multiple simulations, varying the sampling and noise realizations. Re-

[1] Code and test data are available at https://github.com/basp-group/co-oi.

[2] Images from [5] downloaded from the JMMC service at apps.jmmc.fr/oidata/shared/srenard/.

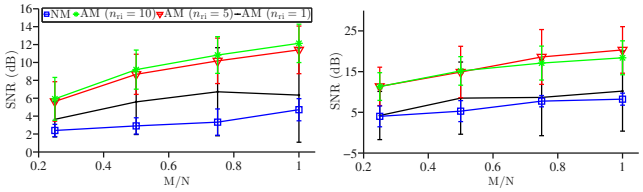

Fig. 1. Reconstruction quality results for synthetic images of size $N = 16^2$ with randomly distributed spikes and ISNR = 30dB for undersampling ratios M/N in the range $[0.25, 1]$. Left panel: 64 spikes. Right panel: 32 spikes. The curves represent the average SNR values over multiple simulations (50 for AM and 10 for NM) and corresponding 1-standard-deviation error bars. Figure borrowed from [14].

constructed images are also reported, providing visual confirmation of the superiority of AM relative to NM over the full undersampling range. On the other hand, the dependency of the nonconvex AM approach on the initial point is clearly illustrated by the $n_{\mathrm{ri}} = 1$ and $n_{\mathrm{ri}} = 5$ curves in Figure 1, confirming the importance of the multiple reinitializations. We observe a saturation between $n_{\mathrm{ri}} = 5$ and $n_{\mathrm{ri}} = 10$ for images of dimension $N = 16^2$. However, tests done in higher dimension images show how this number of necessary reinitializations increases with N, evidencing the instability of AM, as pointed out in [14]. This observation motivates the fact that our study of sparsity priors in the next section is only performed in the context of the purely convex NM approach.

4. NUCLEAR MINIMIZATION WITH SPARSITY

4.1. Algorithm

As a first proof of concept we have chosen to promote the simplest sparsity – in image space – of the signal $\boldsymbol{x}$ of interest, as this can be done simply through adopting a sparsity prior directly on the full tensor $\mathcal{X}$. While ℓ_0-minimization would promote sparsity explicitly, we adopt the common convex relaxation relying on the ℓ_1 norm. Note that a non-weighted ℓ_1 norm is not a meaningful prior function as the tensor values are positive and sum up to unity. In that scenario, we resort to a reweighting scheme consisting in approaching both ℓ_0-minimization on $\mathcal{X}$ and rank minimization on $\mathcal{C}(\mathcal{X})$ by solving a sequence of weighted ℓ_1 and nuclear norm minimization [19, 20], each of which is initialized with the solution of the previous problem.

The weighted-ℓ_1 and nuclear-norm minimization problem (NM-RW) thus reads as:

$$\min_{\mathcal{X}\in S} ||\mathcal{C}(\mathcal{X})||_{*,\mathrm{w}} + \lambda||\mathcal{X}||_{1,\mathrm{w}} \quad \text{s. t.} \quad \|\boldsymbol{y}_{\mathrm{s}} - \mathcal{T}_{\mathrm{s}}(\mathcal{X})\|_2 \leq \epsilon,$$

where S denotes the same set as in (5) and $||\cdot||_{*,w}$ and $||\cdot||_{1,w}$ denote weighted nuclear and ℓ_1 norms respectively. Notice that the weights for the nuclear and the ℓ_1 norm are defined in a different form. In both weighted norms, each element of the vector to be reweighted should essentially be divided by its absolute value in the previous iteration. A stabilization parameter, δ, is necessary to define the weights properly, even when the signal value is zero. In the weighted ℓ_1 norm, each weight is defined as $w_{ijk} = \delta^{(t)}/(\delta^{(t)} + \mathcal{X}^{(t-1)}_{ijk})$, where t indicates the iteration of the reweighting process. λ is set to zero at the first iteration to avoid the use of a non-weighted ℓ_1 norm as a prior, as previously mentioned. In the following iterations, we heuristically set $\lambda^{(t)} = \alpha\|\mathcal{X}^{(t-1)}\|_\infty$,

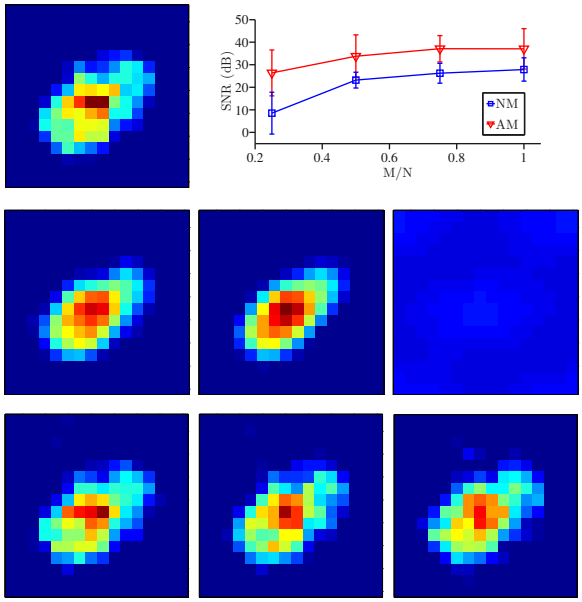

Fig. 2. Eta Carinae star system illustration ($N = 16^2$, ISNR = 30dB). Top row: original image and SNR graph. The curves represent the average SNR values over multiple simulations (50 for AM and 10 for NM) and corresponding 1-standard-deviation error bars. Second and third rows: NM (second) and AM for $n_{\mathrm{ri}} = 5$ (bottom) reconstructions with best SNR for $M = N$ (left), $M = 0.75N$ (centre) and $M = 0.25N$ (right). Figure borrowed from [14].

where $\|\cdot\|_\infty$ denotes the maximum absolute value of the tensor and $0 < \alpha < 1$ is a parameter to be tuned. In order to approximate the rank function through the weighted nuclear norm – i.e. the weighted ℓ_1 norm of the singular values $\sigma_i, i \in \mathbb{R}^N_+$ –, each weight is computed as the inverse of the singular value of $\mathcal{C}(\mathcal{X})$ at the previous iteration, $w_i = \delta^{(t)}/(\delta^{(t)} + \sigma_i^{(t-1)})$. The reweighting process stops when the relative variation between successive solutions $\|\mathcal{X}^{(t)} - \mathcal{X}^{(t-1)}\|_2/\|\mathcal{X}^{(t-1)}\|_2$ is smaller than some bound or after the maximum number of iterations allowed is reached. Finally, the signal is extracted from the tensor using the rank-1 approximation algorithm [18], as mentioned in Section 3.2.

4.2. Numerical Simulations

In this subsection we compare the performance of the linear NM-RW and NM approaches through numerical simulations. In all experiments, the sampling pattern, SNR and ISNR are the same as described in subsection 3.3.

We first evaluate the reconstruction quality on sparse images made of 8 and 16 randomly located spikes. The SNR curves in Figure 3 are built from 10 simulations per point, varying the signal, the sampling and noise realizations. The results show a clear improvement on the SNR when accounting for sparsity. In Figure 4, the effect of the reweighting scheme can be appreciated on an illustration representing a sparsified version of Galaxy M51[2]. Reweighted images (second row) are less blurred and their support is clearly better defined.

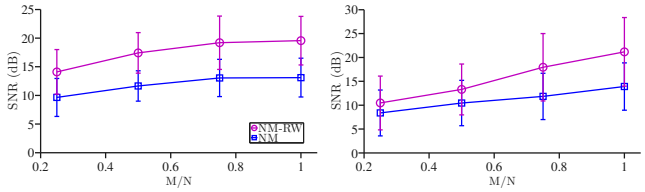

Fig. 3. Reconstruction quality results for synthetic images of size $N = 16^2$ with randomly distributed spikes and ISNR = 30dB for undersampling ratios M/N in the range $[0.25, 1]$. Left panel: 8 spikes. Right panel: 16 spikes. The curves represent the average SNR values over 10 simulations and corresponding 1-standard-deviation error bars.

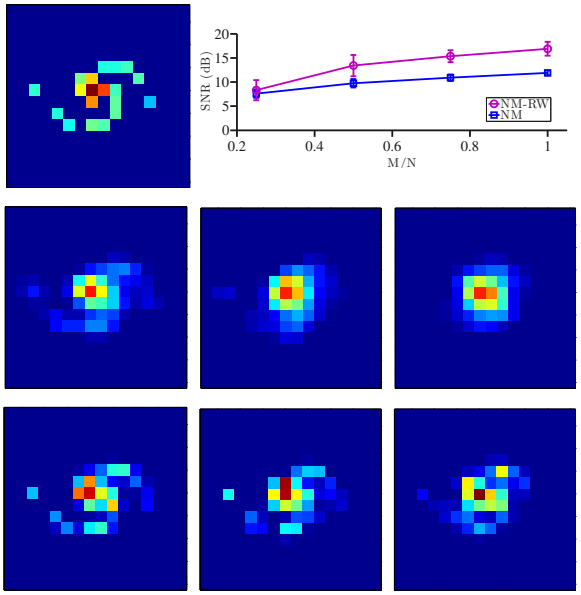

Fig. 4. Sparsified version of M51 Galaxy illustration ($N = 16^2$, ISNR = 30dB). Top row: original image and SNR graph. The curves represent the average SNR values over 10 simulations and corresponding 1-standard-deviation error bars. Second and third rows: NM (second) and NM-RW (bottom) reconstructions with best SNR for $M = N$ (left), $M = 0.75N$ (centre) and $M = 0.25N$ (right).

5. CONCLUSIONS

In this paper we have presented a tensor formulation to solve the optical-interferometric imaging problem. We recall from paper [14] two different approaches to solve it: AM, a nonlinear nonconvex alternate minimization approach and NM, a linear convex approach using nuclear norm minimization as a regularization term. Even if AM provides better image reconstruction quality in low-dimension signals, our simulations suggest that its dependence on the initial point induces an unstable behavior with the increase of N. The underlying convexity of NM ensures essential properties of convergence and independence to initialization, justifying further study, in particular to account for sparsity. Numerical simulations show a clear improvement in the quality of the reconstruction of sparse images when assuming sparsity as a prior. Besides, we point out that, in the present investigations only the simplest case of sparsity – sparsity in image space – has been considered, as a proof of concept. Future work should study the effects of assuming different kinds of sparsity priors, as suggested in recent approaches for radio interferometry [21]. Our methods should also be studied in a more realistic setting with exact power spectrum and bispectrum measurements in the continuous domain and for different noise statistics, and explicitly compared to existing MiRA and WISARD implementations. The linear approaches NM and NM-RW are extremely exacting from a computational standpoint so that software and hardware optimization should also be studied to solve the problem for higher dimension images. Recent results studying the uniqueness of the solution of the phase retrieval problem for sparse signals are presented in [22]. Further research should also analyze our results and formulation in full view of this new theoretical framework.

6. REFERENCES

[1] E. Thiébaut and J. Giovanelli, "Image reconstruction in optical interferometry," *IEEE Signal Processing Magazine*, vol. 21, pp. 97–109, 2010.

[2] H. Thorsteinsson, D. F. Buscher, and J. S. Young, "The bispectrum in model-independent imaging," *Proceedings of SPIE, New Frontiers in Stellar Interferometry*, vol. 5491, pp. 1574–1579, 2004.

[3] J.E. Baldwin and C.A. Haniff, "The application of interferometry to optical astronomical imaging," *Trans. R. Soc. London*, vol. A 360, pp. 969–986, 2002.

[4] E. Thiébaut, "Mira: an effective imaging algorithm for optical interferometry," *Proceedings of SPIE, Optical and Infrared Interferometry*, vol. 7013, pp. 70131I, 2008.

[5] S. Renard, E. Thiébaut, and F. Malbet, "Image reconstruction in optical interferometry: Benchmarking the regularization," *A&A*, vol. 533, pp. A64, 2011.

[6] S. C. Meimon, L. M. Mugnier, and G. Le Besnerais, "Reconstruction method for weak-phase optical interferometry," *Opt. Lett.*, vol. 30, no. 14, pp. 1809–1811, Jul 2005.

[7] Y. Wiaux, L. Jacques, G. Puy, A. M. M. Scaife, and P. Vandergheynst, "Compressed sensing imaging techniques for radio interferometry," *MNRAS*, vol. 395, no. 3, pp. 1733–1742, 2009.

[8] Y. Wiaux, G. Puy, Y. Boursier, and P. Vandergheynst, "Spread spectrum for imaging techniques in radio interferometry," *MNRAS*, vol. 400, no. 2, pp. 1029–1038, 2009.

[9] Y. Wiaux, G. Puy, and P. Vandergheynst, "Compressed sensing reconstruction of a string signal from interferometric observations of the cosmic microwave background," *MNRAS*, vol. 402, no. 4, pp. 2626–2636, 2010.

[10] J. D. McEwen and Y. Wiaux, "Compressed sensing for wide-field radio interferometric imaging," *MNRAS*, vol. 413, no. 2, pp. 1318–1332, 2011.

[11] R. E. Carrillo, J. D. McEwen, and Y. Wiaux, "Sparsity averaging reweighted analysis (sara): a novel algorithm for radio-interferometric imaging," *MNRAS*, vol. 426, no. 2, pp. 1223–1234, 2012.

[12] F. Li, T. J. Cornwell, and F. de Hoog, "Application of compressive sampling to radio astronomy I: Deconvolution," *A&A*, vol. A31, pp. 528–538, 2011.

[13] S. Wenger, M. Magnor, Y. Pihlström, S. Bhatnagar, and U. Rau, "Sparseri: A compressed sensing framework for aperture synthesis imaging in radio astronomy," *Publications of the Astronomical Society of the Pacific*, vol. 122, no. 897, pp. 1367–1374, 2010.

[14] A. Auria, R. E. Carrillo, J.-P. Thiran, and Y. Wiaux, "Tensor optimization for optical-interferometric imaging," *Monthly Notices of the Royal Astronomical Society*, vol. 437, no. 3, pp. 2083–2091, 2014.

[15] R. E. Carrillo, J. D. McEwen, and Y. Wiaux, "PURIFY: a new approach to radio-interferometric imaging," Accepted in MNRAS. Preprint available at http://arxiv.org/abs/1307.4370, 2014.

[16] E. J. Candès, T. Strohmer, and V. Voroninski, "Phase lift: Exact and stable signal recovery from magnitude measurements via convex programming," *Communications on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1241–1274, 2011.

[17] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, "Proximal alternating minimization and projection methods for nonconvex problems," *Mathematics of Operations Research*, vol. 35, no. 3, pp. 438–457, 2010.

[18] E. Kofidis and P. A. Regalia, "On the best rank-1 approximation of higher-order supersymmetric tensors," *SIAM J. Matrix Anal. Appl*, vol. 23, pp. 863–884, 2002.

[19] E. J. Candès, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, 2008.

[20] E. J. Candès, Y. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM J. on Imaging Sciences*, vol. 6, no. 1, pp. 199–225, 2011.

[21] R. E. Carrillo, J. D. McEwen, D. Van De Ville, J.-P. Thiran, and Y. Wiaux, "Sparsity averaging for compressive imaging," *IEEE Signal Processing Letters*, vol. 20, no. 6, pp. 591–594, 2013.

[22] J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli, "Phase retrieval for sparse signals," *arXiv:1308.3058v2*, 2013.