A NONCONVEX REGULARIZED APPROACH FOR PHASE RETRIEVAL

Audrey Repetti, Emilie Chouzenoux and Jean-Christophe Pesquet

Université Paris-Est, Lab. d'Informatique Gaspard Monge, UMR CNRS 8049,
Champs-sur-Marne, 77454 Marne-la-Vallée, France
{first.last}@univ-paris-est.fr

ABSTRACT

With the development of new imaging systems delivering large-size data sets, phase retrieval has become recently the focus of much attention. The problem is especially challenging due to its intrinsically nonconvex formulation. In addition, the applicability of many existing solutions may be limited either by their estimation performance or by their computational cost, especially in the case of non-Fourier measurements. In this paper, we propose a novel phase retrieval approach, which is based on a smooth nonconvex approximation of the standard data fidelity term. In addition, the proposed method allows us to employ a wide range of convex separable regularization functions. The optimization process is performed by a block coordinate proximal algorithm which is amenable to solving large-scale problems. An application of this algorithm to an image reconstruction problem shows that it may be very competitive with respect to state-of-the-art methods.

Index Terms— Phase retrieval problem, Nonconvex optimization, Nonsmooth optimization, Proximal methods

1. INTRODUCTION

Phase retrieval is one of the oldest while still most challenging problems encountered in imaging. It basically consists of producing an estimate $\widehat{v}$ of an original multivariate signal $\overline{v}$ from (possibly noisy) magnitude measurements $|H\overline{v}|$ where H is a complex-valued linear operator. Such a problem plays a central role in numerous application fields of image processing, in particular in crystallography [1], optical imaging [2], phase contrast tomography [3], and coherent diffraction imaging [4].

The most popular methods to estimate $\overline{v}$ are probably *Gerchberg-Saxton* (GS) algorithm [5] and its relaxed version *Fienup* algorithm [6]. Initially introduced for the case when H is a Fourier transform matrix, these algorithms alternate projections onto the range of H and onto the nonconvex set of vectors with magnitude equal to $|H\overline{v}|$. Note that,

since the latter constraint set is nonconvex, these algorithms do not benefit from the convergence guarantees of the famous *Projections Onto Convex Set* (POCS) algorithm [7]. Convex relaxations of the phase retrieval problem based on *Semi-Definite-Programming* (SDP) formulations were proposed in [8, 9], giving rise to the *PhaseLift* and *PhaseCut* algorithms, respectively. Both methods have been shown to outperform GS and Fienup, in a certain number of application examples [10].

All the aforementioned methods tend to be sensitive to noise and/or to be less effective in the under-determined case [11]. Such difficulties are mainly due to the ill-posed character of the phase retrieval problem and they can be alleviated by incorporating some a priori information in the reconstruction process. Several algorithms based on SDP relaxation [12], alternating projection [13] and greedy pursuit [11], have been recently introduced, which aim at solving the phase retrieval problem under the prior assumption that the sought signal $\overline{v}$ has a sparse representation in a (possibly redundant) dictionary. However, up to the best of our knowledge, these approaches may become very computationally intensive as the problem dimension increases. This is particularly critical for SDP and alternating projections methods, when the pseudo-inverse of H does not have a closed form, as it usually happens for non-Fourier measurements, or for the greedy method in [11] when the sparsity degree of the data is not sufficient.

In this paper, we introduce a new phase retrieval strategy relying on the minimization of a penalized criterion. This criterion is composed of (*i*) a difference of convex functions (DC function) achieving a smooth approximation of the standard nonconvex nonsmooth least-squares data fidelity term, and (*ii*) a convex non necessarily smooth block separable penalization term. The latter function can be chosen in a flexible manner. For instance, it can be used to incorporate a sparsity assumption on the solution. The resulting nonconvex nonsmooth optimization problem is then solved with the *Block Coordinate Variable Metric Forward-Backward* (BC-VMFB) method that we have recently introduced in [14]. This algorithm is ensured to converge to a critical point of the criterion and it takes advantage of the block separable structure of the regularization term. The proposed method thus benefits from a low computational cost which makes it applicable to large-

scale problems.

The rest of the paper is organized as follows : Section 2 presents the phase retrieval problem and introduces a new objective function. Section 3 describes the proposed BC-VMFB algorithm and emphasizes its convergence properties. Finally, in Section 4, the good performance of our approach is shown via experiments.

2. PROBLEM FORMULATION

2.1. Observation model

Let $\overline{v} \in \mathbb{R}^M$ be the original unknown signal. Given $H \in \mathbb{C}^{S \times M}$ an observation matrix with complex-valued elements, the measurements $z \in [0, +\infty)^S$ are assumed to be related to the original signal through the model

$$z = (z^{(s)})_{1 \leq s \leq S} = |H\overline{v}| + w, \qquad (1)$$

where $|\cdot|$ denotes the componentwise modulus operator, and $w \in [0, +\infty)^S$ is a realization of an additive noise.

It is important to note that Model (1) encompasses the case when the target signal is complex since the observed vector can then be expressed under the form

$$z = \left| [H_{\mathcal{R}} + \mathrm{i}\, H_{\mathcal{I}} \mid -H_{\mathcal{I}} + \mathrm{i}\, H_{\mathcal{R}}] \begin{bmatrix} \overline{v}_{\mathcal{R}} \\ \overline{v}_{\mathcal{I}} \end{bmatrix} \right| + w, \quad (2)$$

with $\mathrm{i}^2 = -1$ and $(\cdot)_{\mathcal{R}}$ (resp. $(\cdot)_{\mathcal{I}}$) designates the real (resp. imaginary) part of its argument.

2.2. Minimization problem

Let $W \in \mathbb{R}^{M \times N}$, $M \leq N$, be a given frame synthesis operator (e.g. a possibly redundant wavelet representation) [15] such that $\overline{v} = W\overline{x}$. Following a synthesis approach, the estimated signal reads $\widehat{v} = W\widehat{x}$ where the frame coefficient vector $\widehat{x} \in \mathbb{R}^N$ is estimated by minimizing the sum of a data fidelity term F and a regularization function R, i.e.

$$G = F + R. \qquad (3)$$

In the context of phase retrieval problems, a usual way to define the data fidelity term F is to consider the nonconvex nonsmooth least-squares criterion [6, 11] :

$$(\forall x \in \mathbb{R}^N) \quad F(x) = \sum_{s=1}^{S} \varphi^{(s)}([HWx]^{(s)}), \qquad (4)$$

where, for every $s \in \{1, \ldots, S\}$, $\varphi^{(s)} = \frac{1}{2}(|\cdot| - z^{(s)})^2$. We propose to replace the latter nonsmooth function by a smooth approximation of it, easier to handle, which is defined as the difference of two convex terms, parametrized by a constant $\delta > 0$:

$$(\forall u \in \mathbb{C})$$
$$\varphi^{(s)}(u) = \frac{1}{2}\left(|u|^2 + (z^{(s)})^2\right) - z^{(s)}\left(|u|^2 + \delta^2\right)^{1/2}.$$

It can be noticed that, in the limit case when $\delta = 0$, the standard nonsmooth fidelity function is recovered.

Regarding the penalization term, we will focus on the case when R is a block separable function. More precisely, let us define $(\mathbb{J}_j)_{1 \leq j \leq J}$ a partition of the frame coefficient index set $\{1, \ldots, N\}$ into $J \geq 2$ subsets with nonzero block-sizes $(N_j)_{1 \leq j \leq J}$. In the following, we assume that the regularization function can be expressed as

$$R(x) = \sum_{j=1}^{J} R_j(x^{(j)}), \qquad (5)$$

where, for every $j \in \{1, \ldots, J\}$, function $R_j \colon \mathbb{R}^{N_j} \to (-\infty, +\infty]$ is proper, lower semicontinuous, convex and continuous on its domain, and $x^{(j)} \in \mathbb{R}^{N_j}$ denotes the j-th block of components of x.

3. PROPOSED ALTERNATING OPTIMIZATION METHOD

3.1. Optimization tools

We first introduce the notion of proximity operator which will be useful to present our algorithm. Let $\psi \colon \mathbb{R}^N \to (-\infty, +\infty]$ be a convex, proper, lower semicontinuous function and let $U \in \mathbb{R}^{N \times N}$ be a symmetric positive definite (SPD) matrix. The proximity operator of ψ at $\xi \in \mathbb{R}^N$ relative to the metric induced by $U \in \mathbb{R}^{N \times N}$ is defined as ([16, Sec. XV.4] and [17])

$$\mathrm{prox}_{U,\psi}(\xi) = \mathrm{argmin}_{\zeta \in \mathbb{R}^N}\, \psi(\zeta) + \frac{1}{2}\|\zeta - \xi\|_U^2.$$

Hereabove, $\|\cdot\|_U$ denotes the weighted norm defined by $\|\cdot\|_U = \langle\cdot, U\cdot\rangle^{1/2}$, where $\langle\cdot, \cdot\rangle$ is the usual Euclidean scalar product. Note that, if U is the identity matrix, then the usual proximity operator introduced in the pioneering work [18] is recovered.

3.2. Proposed algorithm

Although the proposed data fidelity term (4) is a nonconvex function, it presents the advantage of being Lipschitz differentiable. Such a property suggests the use of the *Forward-Backward* (FB) scheme [19] for minimizing the objective function (3). A major advantage of FB algorithm is that it benefits from proven theoretical convergence guarantees, even in the nonconvex case [20, 21]. However, in practice, this algorithm may suffer from slow convergence so that accelerated versions have been proposed in [14, 22–26], still with convergence guarantees in the nonconvex case ([14, 22–24]). In this paper, the block separable structure of R is accounted for by adopting the following BC-VMFB method from [14]. This algorithm combines the advantages of a preconditioning strategy [22] with an alternating minimization scheme :

Algorithm 1 BC-VMFB algorithm.

For every $\ell \in \mathbb{N}$, let $\gamma_\ell \in (0, +\infty)$ and $\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell) \in \mathbb{R}^{N_{j_\ell} \times N_{j_\ell}}$ be a SPD preconditioning matrix.

Initialize with $\boldsymbol{x}_0 \in \mathbb{R}^N$.

Iterations :
 For $\ell = 0, 1, \ldots$
  Select a block index $j_\ell \in \{1, \ldots, J\}$.
  $\tilde{\boldsymbol{x}}_\ell^{(j_\ell)} = \boldsymbol{x}_\ell^{(j_\ell)} - \gamma_\ell \boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell)^{-1} \nabla_{j_\ell} F(\boldsymbol{x}_\ell)$,
  $\boldsymbol{x}_{\ell+1}^{(j_\ell)} = \text{prox}_{\gamma_\ell^{-1} \boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell), R_{j_\ell}} \left(\tilde{\boldsymbol{x}}_\ell^{(j_\ell)} \right)$,
  $\boldsymbol{x}_{\ell+1}^{(\bar{j}_\ell)} = \boldsymbol{x}_\ell^{(\bar{j}_\ell)}$.

In the above algorithm, at each iteration $\ell \in \mathbb{N}$, $\nabla_{j_\ell} F(\boldsymbol{x}_\ell) \in \mathbb{R}^{N_{j_\ell}}$ denotes the partial gradient of F with respect to block $\boldsymbol{x}^{(j_\ell)}$ computed at $\boldsymbol{x}_\ell$, and $\bar{j}_\ell$ is the complementary set of j_ℓ on $\{1, \ldots, J\}$, i.e. $\bar{j}_\ell = \{1, \ldots, J\} \setminus \{j_\ell\}$.

Note that Algorithm 1 is equivalent to the algorithm in [22] for $N_{j_\ell} \equiv N$ and reduces to the standard FB algorithm if, additionally, $\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell)$ reduces to the identity matrix. The convergence of BC-VMFB requires the design of a proper strategy to adjust the preconditioning matrices $(\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell))_{\ell \in \mathbb{N}}$ in an automatic manner. This will be discussed in the next section.

3.3. Choice of the preconditioning matrices

At iteration $\ell \in \mathbb{N}$, let j_ℓ be the chosen index in $\{1, \ldots, J\}$ and let $\boldsymbol{x}_\ell$ be the ℓ-th iterate generated by Algorithm 1. Then, we propose to choose matrix $\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell)$ by following the Majorize-Minimize principle [27], i.e. it is such that the quadratic function

$$Q_{j_\ell}(\cdot, \boldsymbol{x}_\ell) = F(\boldsymbol{x}_\ell) + (\cdot - \boldsymbol{x}_\ell^{(j_\ell)})^\top \nabla_{j_\ell} F(\boldsymbol{x}_\ell)$$
$$+ \frac{1}{2} \| \cdot - \boldsymbol{x}_\ell^{(j_\ell)} \|_{\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell)}^2,$$

majorizes the restriction of F to its j_ℓ-th block on $\mathbb{R}^{N_{j_\ell}}$. We thus have, for every $\boldsymbol{y} \in \mathbb{R}^{N_{j_\ell}}$,

$$F(\boldsymbol{x}_\ell^{(1)}, \ldots, \boldsymbol{x}_\ell^{(j_\ell-1)}, \boldsymbol{y}, \boldsymbol{x}_\ell^{(j_\ell+1)}, \ldots, \boldsymbol{x}_\ell^{(J)}) \leq Q_{j_\ell}(\boldsymbol{y}, \boldsymbol{x}_\ell).$$

According to our result in [14, Prop.4.1], such a majorization condition is fulfilled by the diagonal matrix

$$\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell) = \text{Diag}\left(\boldsymbol{\Omega}_{j_\ell}^\top \boldsymbol{1}_S \right), \tag{6}$$

where $\boldsymbol{1}_S$ is the unit vector on $\mathbb{R}^S$ and the elements of $\boldsymbol{\Omega}_{j_\ell} \in \mathbb{R}^{S \times N_{j_\ell}}$ are given, for every $(s, n) \in \{1, \ldots, S\} \times \mathbb{J}_{j_\ell}$, by

$$\Omega^{(s,n)} = |[\boldsymbol{HW}]_\mathcal{R}^{(s,n)}| \sum_{n'=1}^N |[\boldsymbol{HW}]_\mathcal{R}^{(s,n')}|$$
$$+ |[\boldsymbol{HW}]_\mathcal{I}^{(s,n)}| \sum_{n'=1}^N |[\boldsymbol{HW}]_\mathcal{I}^{(s,n')}|. \tag{7}$$

3.4. Convergence result

One appealing feature of Algorithm 1 is that its convergence is guaranteed by the following result deduced from [14, Thm. 3.1] :

Theorem 1. *Let $(\boldsymbol{x}_\ell)_{\ell \in \mathbb{N}}$ be a sequence generated by Algorithm 1 where $(\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell))_{\ell \in \mathbb{N}}$ are defined as in (6). Assume that*

(i) *G is a coercive function, i.e. $\lim_{\|\boldsymbol{x}\| \to +\infty} G(\boldsymbol{x}) = +\infty$.*

(ii) *R is a semi-algebraic function.* [1]

(iii) *There exists $(\underline{\gamma}, \overline{\gamma}) \in (0, +\infty)^2$ such that, for every $\ell \in \mathbb{N}$, $\underline{\gamma} \leq \gamma_\ell \leq 2 - \overline{\gamma}$.*

(iv) *The blocks are updated according to an essentially cyclic rule, i.e. there exists a constant $K \geq J$ such that, for every $\ell \in \mathbb{N}$, $\{1, \ldots, J\} \subset \{j_\ell, \ldots, j_{\ell+K-1}\}$.*

Then, the sequence $(\boldsymbol{x}_\ell)_{\ell \in \mathbb{N}}$ converges to a critical point $\widehat{\boldsymbol{x}}$ of (3). Moreover, $(G(\boldsymbol{x}_\ell))_{\ell \in \mathbb{N}}$ monotonically converges to $G(\widehat{\boldsymbol{x}})$.

4. EXPERIMENTAL RESULTS

4.1. Observation model

We now demonstrate the practical performance of our algorithm in an image reconstruction problem. Starting from a complex-valued original image $\overline{\boldsymbol{v}} \in \mathbb{C}^M$ (see Fig. 1) with size $M = 128 \times 128$, we generate noisy observations $\boldsymbol{z} \in \mathbb{R}^S$ according to Model (2). The observation matrix $\boldsymbol{H} \in \mathbb{C}^{S \times M}$ results from the composition of two degradation operators, namely a projection matrix modeling $S = 23400$ Radon projections from 128 parallel acquisition lines and 180 angles regularly distributed on $[0, \pi)$, and a complex-valued blur operator. For the latter, a 1D convolution kernel with length 3, acting separately on the acquired sinogram for each angle has been used. It can be noticed that such a model is reminiscent of the linearized version of the phase contrast tomography problem addressed in [28, 29] in the 3D case.

4.2. Regularization term

We choose the synthesis frame operator $\boldsymbol{W}$ so that each vector $\boldsymbol{x} = (x^{(n)})_{1 \leq n \leq N} \in \mathbb{R}^N$, $N = 8M$, is the concatenation of two vectors of $\mathbb{R}^{4M}$, corresponding to an overcomplete Haar decomposition of $\boldsymbol{v}_\mathcal{R}$ (resp. $\boldsymbol{v}_\mathcal{I}$) for one resolution level. An example of frame decomposition of the real part of the original image is displayed in Fig. 2.

The regularization function R is defined as the sum, for $p \in \{1, \ldots, 4M\}$ of terms of the form :

---

1. Semi-algebraicity is a property satisfied by a wide class of functions, which means that their graph is a finite union of sets defined by a finite number of polynomial inequalities.
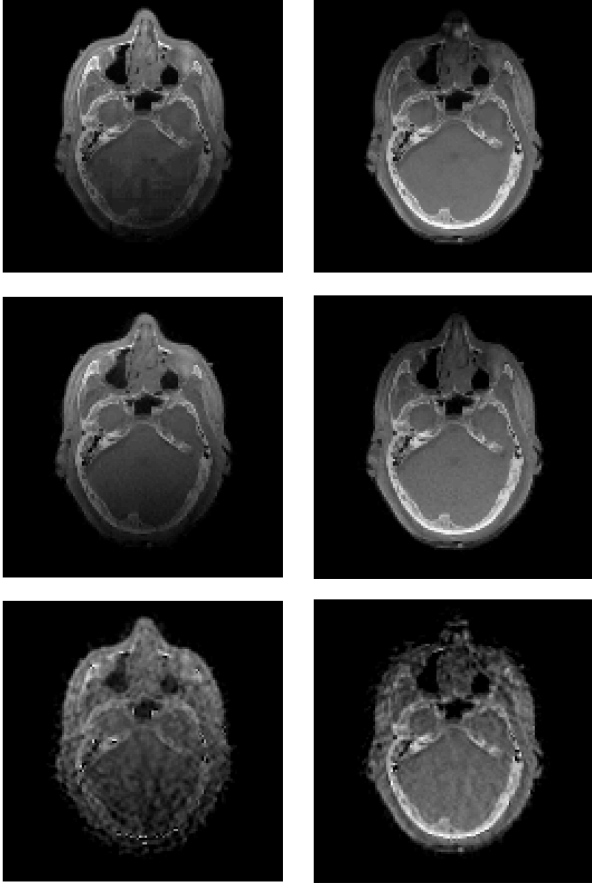
Fig. 1. (top) Real and imaginary parts of the original image $\overline{v}$ and estimated images $\widehat{v}$ using either (middle) Algorithm 1, SNR = 21.27 dB or (bottom) the regularized alternating projection method from [13], SNR = 14.45 dB.
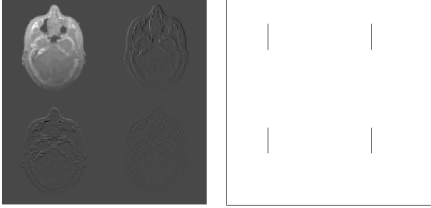

Fig. 2. (left) Example of frame decomposition of the original image. (right) Indices of a block $\boldsymbol{x}^{(j)}$ for $Q = 32$. Only the coefficients corresponding to the real part of the image are displayed.

$$\rho^{(p)}(\boldsymbol{u}^{(p)}) = \begin{cases} \vartheta_p \|\boldsymbol{u}^{(p)} - \boldsymbol{\omega}_p\|_2^{\kappa_p} & \text{if } p \notin \mathbb{E}, \\ 0 & \text{if } p \in \mathbb{E} \text{ and } \boldsymbol{u}^{(p)} = \boldsymbol{0}, \\ +\infty & \text{otherwise}, \end{cases}$$

where $\boldsymbol{u}^{(p)} \in \mathbb{R}^2$ is the p-th pair of frame coefficients corresponding to the real and imaginary part of the image, and $\mathbb{E}$ corresponds to the object background. The sparsity of the detail part of $\widehat{\boldsymbol{x}}$ is promoted by setting $\kappa_p = 1$ and $\vartheta_p = \vartheta^d \in (0, +\infty)$ for the detail subbands of the decomposition. For the approximation coefficients, we choose $\kappa_p = 2$ and $\vartheta_p = \vartheta^a \in (0, +\infty)$. Parameter $\boldsymbol{\omega}_p \in \mathbb{R}^2$ corresponds to a
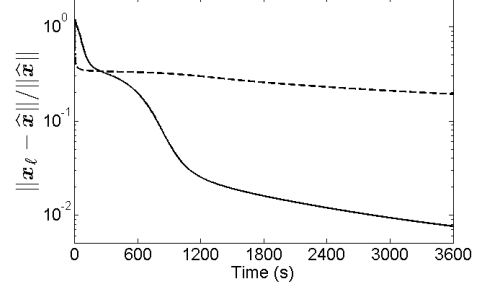

Fig. 3. Convergence profile of Algorithm 1 (solid line) and the non preconditioned variant from [24] (dashed line), Intel(R) Core(TM) i7-3520M @ 2.9GHz using Matlab 7.

mean value parameter on the sought frame coefficient $\boldsymbol{u}^{(p)}$. In our experiments, we take $\boldsymbol{\omega}_p = (0.4, 0.6)$ for the approximation part, and $\boldsymbol{\omega}_p = \boldsymbol{0}$ otherwise. Thus, the resulting criterion G satisfies Assumptions (i)-(iv) of Theorem 1. Note that the proximity operators involved in Algorithm 1 are easy to compute since they have an explicit form [30].

4.3. Simulation results

The magnitude measurement vector is corrupted with an additive real-valued white zero-mean Gaussian noise with variance equal to 0.025 (truncated so as to guarantee the non-negativity of the observations). In our experiments, parameters ϑ^a, ϑ^d and δ are adjusted so as to maximize the signal-to-noise ratio (SNR) between the original image $\overline{v}$ and the reconstructed one $\widehat{v}$. In Algorithm (1), for every j, $\boldsymbol{x}^{(j)}$ corresponds to a vector of given fixed size $8Q$ which gathers 8 blocks extracted from the approximation and detail subbands of both real and imaginary parts of the image, as depicted in Fig. 2. At each iteration $\ell \in \mathbb{N}$, j_ℓ is then randomly chosen so that each block is updated at least once per J iterations. In practice, the reconstruction time was observed to vary significantly with respect to the block-size parameter Q. In this example, the best choice in terms of convergence speed is reached for $Q = 2$. Fig. 1 shows the reconstructed image obtained with our algorithm using this optimal block-size and $\gamma_\ell \equiv 1.9$. We also present in Fig. 1 the reconstruction result obtained using an enhanced version of the regularized alternating projection approach from [13], where the non-inversibility of matrix $\boldsymbol{H}$ has been taken into account. Note that due to the large size of the data, it appeared impossible to run SDP approaches from [8,9,12], neither the greedy method from [11]. Moreover, we obtained poor results using the standard GS algorithm.

Fig. 3 illustrates the decay of $(\|\boldsymbol{x}_\ell - \widehat{\boldsymbol{x}}\|/\|\widehat{\boldsymbol{x}}\|)_\ell$ (in log scale) using the proposed algorithm or the non preconditioned variant from [24]. The optimal solution $\widehat{\boldsymbol{x}}$ has been precomputed for each algorithm, using a large number of iterations. One can observe that the variable metric strategy leads to a significant acceleration of the convergence. It must be noticed that the method from [13] is almost 10 times slower to reach a stabilized SNR value.

5. REFERENCES

[1] R. Harrison, "Phase problem in crystallography," *J. Opt. Soc. Amer. A*, vol. 10, no. 5, pp. 1046–1055, 1993.

[2] A. Walther, "The question of phase retrieval in optics," *J. Modern Opt.*, vol. 10, no. 1, pp. 41–49, 1963.

[3] H. H. Bauschke, P. L. Combettes, and D. R. Luke, "A new generation of iterative transform algorithms for phase contrast tomography," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP 2005)*, Philadelphia, PA, 19-23 Mar. 2005, vol. 4, pp. 89–92.

[4] Y. Shechtman, Y. Eldar, A. Szameit, and M. Segev, "Efficient coherent diffractive imaging for sparsely varying objects," *Optics Express*, vol. 21, no. 5, pp. 6327–6338, 2013.

[5] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.

[6] J. R. Fienup, "Phase retrieval algorithms : A comparison," *Appl. Opt.*, vol. 21, pp. 2758–2769, Aug. 1982.

[7] H. H. Bauschke, P. L. Combettes, and D. R. Luke, "Phase retrieval, error reduction algorithm, and Fienup variants : a view from convex optimization," *J. Opt. Soc. Amer. A*, vol. 19, no. 7, pp. 1334–1345, Jul. 2002.

[8] E. Candès, T. Strohmer, and V. Voroninski, "Phaselift : Exact and stable signal recovery from magnitude measurements via convex programming," *Comm. Pure Appl. Math.*, vol. 66, no. 8, pp. 1241–1274, 2013.

[9] I. Waldspurger, A. d'Aspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," *To appear in Math. Program.*, 2013.

[10] F. Fogel, I. Waldspurger, and A. D'Aspremont, "Phase retrieval for imaging problems," Tech. Rep., Apr. 2013, `http://arxiv.org/abs/1304.7735/`.

[11] Y. Shechtman, A. Beck, and Y. Eldar, "GESPAR : Efficient phase retrieval of sparse signals," *IEEE Trans. Signal Process.*, vol. 62, no. 4, pp. 928–938, Feb 2014.

[12] K. Jaganathan, S. Oymak, and B. Hassibi, "Recovery of sparse 1-D signals from the magnitudes of their Fourier transform," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT 2012)*, Cambridge, MA, 1-6 Jul. 2012, pp. 1473–1477.

[13] S. Mukherjee and C.S. Seelamantula, "An iterative algorithm for phase retrieval with sparsity constraints : application to frequency domain optical coherence tomography," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP 2012)*, Kyoto, Japan, 25-30 Mar. 2012, pp. 553–556.

[14] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "A block coordinate variable metric forward-backward algorithm," Tech. Rep., 2013, `http://www.optimization-online.org/DBHTML/2013/12/4178.html`.

[15] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, Burlington, MA, 3rd edition, 2009.

[16] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms*, Springer-Verlag, New York, 1993.

[17] P. L. Combettes and B. C. Vũ, "Variable metric quasi-Fejér monotonicity," *Nonlinear Anal.*, vol. 78, pp. 17–31, 2013.

[18] J. J. Moreau, "Proximité et dualité dans un espace hilbertien," *Bull. Soc. Math. France*, vol. 93, pp. 273–299, 1965.

[19] P. Tseng, "A modified forward-backward splitting method for maximal monotone mappings," *SIAM J. Control Optim.*, vol. 38, no. 2, pp. 431–446, 1998.

[20] H. Attouch, J. Bolte, and B. F. Svaiter, "Convergence of descent methods for semi-algebraic and tame problems : proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods," *Math. Program.*, vol. 137, pp. 91–129, Feb. 2011.

[21] H. Attouch and J. Bolte, "On the convergence of the proximal algorithm for nonsmooth functions involving analytic features," *Math. Program.*, vol. 116, pp. 5–16, Jun. 2009.

[22] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function," *To appear in J. Optim. Theory Appl.*, 2013.

[23] P. Ochs, Y. Chen, T. Brox, and T. Pock, "iPiano : inertial proximal algorithm for non-convex optimization," *To appear in SIAM J. Imaging Sci.*, 2013.

[24] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *To appear in Math. Program.*, 2013.

[25] Q. Tran-Dinh, A. Kyrillidis, and V. Cevher, "Composite self-concordant minimization," Tech. Rep., 2013, `http://arxiv.org/pdf/1308.2867v1.pdf`.

[26] P. L. Combettes and B. C. Vũ, "Variable metric forward-backward splitting with applications to monotone inclusions in duality," *to appear in Optimization*, 2013.

[27] S. Sotthivirat and J. A. Fessler, "Image recovery using partitioned-separable paraboloidal surrogate coordinate ascent algorithms," *IEEE Trans. Signal Process.*, vol. 11, no. 3, pp. 306–317, 2002.

[28] J. P. Guigay, M. Langer, R. Boistel, and P. Cloetens, "Mixed transfer function and transport of intensity approach for phase retrieval in the Fresnel region," *Opt. Lett.*, vol. 32, no. 12, pp. 1617–1619, Jun 2007.

[29] V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin, "Nonlinear phase retrieval using projection operator and iterative wavelet thresholding," *IEEE Signal Process. Lett.*, vol. 19, no. 9, pp. 579–582, Sept 2012.

[30] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds., pp. 185–212. Springer-Verlag, New York, 2010.