A NONCONVEX REGULARIZED APPROACH FOR PHASE RETRIEVAL

Audrey Repetti, Emilie Chouzenoux, and Jean-Christophe Pesquet
IGM and UMR-CNRS 8049, Paris-Est Univ. France
first.last@univ-paris-est.fr

State of the art

✓ Existing phase retrieval methods:

- POCS algorithms (e.g. Gerchberg-Saxton, Fienup, sparse Fienup, ...)
- Semi-Definite Programming formulations (e.g. PhaseLift, PhaseCut, ...)
- Greedy pursuit algorithm (e.g. GESPAR, ...)

✗ May be limited either by their estimation performance or by their computational cost, especially in the case of non-Fourier measurements.

Contributions

★ Smooth approximation of the usual least-squares criterion.

★ Flexible choices for the regularization function.

★ Proposition of an efficient minimization algorithm:

- $\rightsquigarrow$ ability to handle high dimensional problems
- $\rightsquigarrow$ convergence ensured for nonconvex problems
- $\rightsquigarrow$ variable metric strategy for accelerated convergence.

OBSERVATION MODEL

Observation measurements $\boldsymbol{z} = (z^{(s)})_{1\le s\le S} \in [0,+\infty[^S$:

$$\boldsymbol{z} = |\boldsymbol{H}\overline{\boldsymbol{v}}| + \boldsymbol{w}$$

- $\overline{\boldsymbol{v}} \in \mathbb{R}^M$ $\rightsquigarrow$ original unknown image
- $\boldsymbol{H} \in \mathbb{C}^{S\times M}$ $\rightsquigarrow$ observation matrix
- $\boldsymbol{w} \in [0,+\infty[^S$ $\rightsquigarrow$ realization of a positive additive noise

If $\overline{\boldsymbol{v}}$ is complex: $\overline{\boldsymbol{v}} = \overline{\boldsymbol{v}}_{\mathcal{R}} + \mathrm{i}\,\overline{\boldsymbol{v}}_{\mathcal{I}}$

$\rightsquigarrow$ $\boldsymbol{z} = |(\boldsymbol{H}_{\mathcal{R}} + \mathrm{i}\,\boldsymbol{H}_{\mathcal{I}})(\overline{\boldsymbol{v}}_{\mathcal{R}} + \mathrm{i}\,\overline{\boldsymbol{v}}_{\mathcal{I}})| + \boldsymbol{w}$

$\rightsquigarrow$ $\boldsymbol{z} = |\underbrace{[\boldsymbol{H}_{\mathcal{R}} + \mathrm{i}\,\boldsymbol{H}_{\mathcal{I}} \mid -\boldsymbol{H}_{\mathcal{I}} + \mathrm{i}\,\boldsymbol{H}_{\mathcal{R}}]}_{\text{Complex}} \underbrace{\begin{bmatrix}\overline{\boldsymbol{v}}_{\mathcal{R}}\\ \overline{\boldsymbol{v}}_{\mathcal{I}}\end{bmatrix}}_{\text{Real}}| + \boldsymbol{w}$

Synthesis approach: $\overline{\boldsymbol{v}} = \boldsymbol{W}\overline{\boldsymbol{x}}$, where $\boldsymbol{W} \in \mathbb{R}^{M\times N}$ is a frame synthesis operator ($M \le N$).

$\rightsquigarrow$ Reconstruction problem: Find $\widehat{\boldsymbol{x}} \in \mathbb{R}^M$ from $\boldsymbol{z} = |\boldsymbol{H}\boldsymbol{W}\overline{\boldsymbol{x}}| + \boldsymbol{w}$.

MINIMIZATION PROBLEM

$$\text{Find } \widehat{\boldsymbol{x}} \in \operatorname{Argmin}\ (F+R) \tag{1}$$

Data fidelity term: $\rightsquigarrow$ smooth nonconvex function

$$(\forall \boldsymbol{x} \in \mathbb{R}^N) \quad F(\boldsymbol{x}) = \sum_{s=1}^{S} \varphi^{(s)}([\boldsymbol{H}\boldsymbol{W}\boldsymbol{x}]^{(s)})$$

where

$$(\forall u \in \mathbb{C}) \quad \varphi^{(s)}(u) = \tfrac{1}{2}\left(|u|^2 + (z^{(s)})^2\right) - z^{(s)}\left(|u|^2 + \delta^2\right)^{1/2}, \qquad \delta > 0.$$

Regularization term: $\rightsquigarrow$ block separable structure

- $(\forall \boldsymbol{x} \in \mathbb{R}^N)$ $\boldsymbol{x} = (\underbrace{\boldsymbol{x}^{(1)}}_{\in\mathbb{R}^{N_1}}, \ldots, \underbrace{\boldsymbol{x}^{(J)}}_{\in\mathbb{R}^{N_J}})$.
- $R(\boldsymbol{x}) = \sum_{j=1}^{J} R_j(\boldsymbol{x}^{(j)})$, where $(\forall j \in \{1,\ldots,J\})$ R_j is proper, l.s.c., convex and continuous on its domain.

PROPOSED ALGORITHM

Let $\boldsymbol{x}_0 \in \operatorname{dom} R$ and $(\underline{\gamma}, \overline{\gamma}) \in]0,+\infty[^2$.
For $\ell = 0, 1, \ldots$

$$\left\lfloor \begin{array}{l} \text{Let } j_\ell \in \{1,\ldots,J\} \text{ and } \gamma_\ell \in [\underline{\gamma}, 2-\overline{\gamma}]. \\ \boldsymbol{x}_{\ell+1}^{(j_\ell)} = \operatorname{prox}_{\gamma_\ell^{-1}\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell), R_{j_\ell}}\left(\boldsymbol{x}_\ell^{(j_\ell)} - \gamma_\ell\,(\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell))^{-1}\nabla_{j_\ell}F(\boldsymbol{x}_\ell)\right), \\ \boldsymbol{x}_{\ell+1}^{(\overline{\jmath}_\ell)} = \boldsymbol{x}_\ell^{(\overline{\jmath}_\ell)}. \end{array}\right.$$

- $\operatorname{prox}_{\gamma_\ell^{-1}\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell), R_{j_\ell}} = \underset{\boldsymbol{y}\in\mathbb{R}^{N_{j_\ell}}}{\operatorname{argmin}}\ R_{j_\ell}(\boldsymbol{y}) + \frac{1}{2\gamma_\ell}\|\boldsymbol{y} - \cdot\|^2_{\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell)}$
  with $\|\cdot\|^2_{\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell)} = (\cdot)^\top \boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell)(\cdot)$.
- $\nabla_{j_\ell}F(\boldsymbol{x}_\ell) \in \mathbb{R}^{N_{j_\ell}}$ is the partial gradient of F w.r.t. $\boldsymbol{x}^{(j_\ell)}$ computed at $\boldsymbol{x}_\ell$.
- $\overline{\jmath}_\ell = \{1,\ldots,J\} \setminus \{j_\ell\}$ and $\boldsymbol{x}^{(\overline{\jmath}_\ell)} = (\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(\overline{\jmath}_\ell - 1)}, \boldsymbol{x}^{(\overline{\jmath}_\ell + 1)}, \ldots, \boldsymbol{x}^{(J)})$.
- $\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell) \in \mathbb{R}^{N_{j_\ell}\times N_{j_\ell}}$ is an SPD matrix such that

$$(\forall \boldsymbol{y} \in \mathbb{R}^{N_{j_\ell}}) \quad F(\boldsymbol{x}_\ell^{(1)}, \ldots, \boldsymbol{x}_\ell^{(j_\ell-1)}, \boldsymbol{y}, \boldsymbol{x}_\ell^{(j_\ell+1)}, \ldots, \boldsymbol{x}_\ell^{(J)}) \le F(\boldsymbol{x}_\ell) + (\boldsymbol{y} - \boldsymbol{x}_\ell^{(j_\ell)})^\top \nabla_{j_\ell}F(\boldsymbol{x}_\ell) + \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{x}_\ell^{(j_\ell)}\|^2_{\boldsymbol{A}_{j_\ell}(\boldsymbol{x}_\ell)}.$$

CONVERGENCE RESULT

Assumptions

- $*$ G is a coercive function, i.e. $\lim_{\|\boldsymbol{x}\|\to+\infty} G(\boldsymbol{x}) = +\infty$.
- $*$ R is a semi-algebraic function.
- $*$ The blocks are updated according to an essentially cyclic rule, i.e. there exists a constant $K \ge J$ such that, for every $\ell \in \mathbb{N}$, $\{1,\ldots,J\} \subset \{j_\ell, \ldots, j_{\ell+K-1}\}$.

Convergence theorem

- ▶ $(\boldsymbol{x}_\ell)_{\ell\in\mathbb{N}}$ converges to a critical point $\widehat{\boldsymbol{x}}$ of (1).
- ▶ $\big(G(\boldsymbol{x}_\ell)\big)_{\ell\in\mathbb{N}}$ is a nonincreasing sequence converging to $G(\widehat{\boldsymbol{x}})$.

▶ **Observation matrix:** $\boldsymbol{H} \in \mathbb{C}^{S\times M}$ is the composition of:

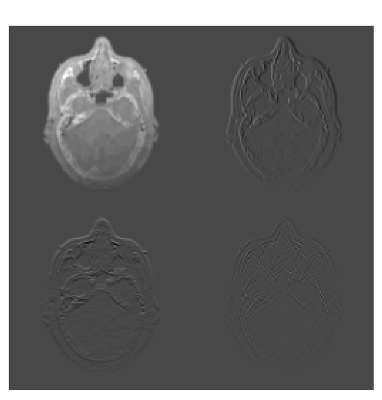
- a matrix modeling S parallel Radon projections
- a complex-valued blur operator.

▶ **Synthesis frame operator:** $\boldsymbol{W} \in \mathbb{R}^{M\times N}$, $N = 8M$, such that $\boldsymbol{x} = (\boldsymbol{x}_{\mathcal{R}}, \boldsymbol{x}_{\mathcal{I}}) \in \mathbb{R}^{4M} \times \mathbb{R}^{4M}$ with $\boldsymbol{x}_{\mathcal{R}}$ (resp. $\boldsymbol{x}_{\mathcal{I}}$) is an overcomplete Haar decomposition of $\boldsymbol{v}_{\mathcal{R}}$ (resp. $\boldsymbol{v}_{\mathcal{I}}$) for one resolution level.

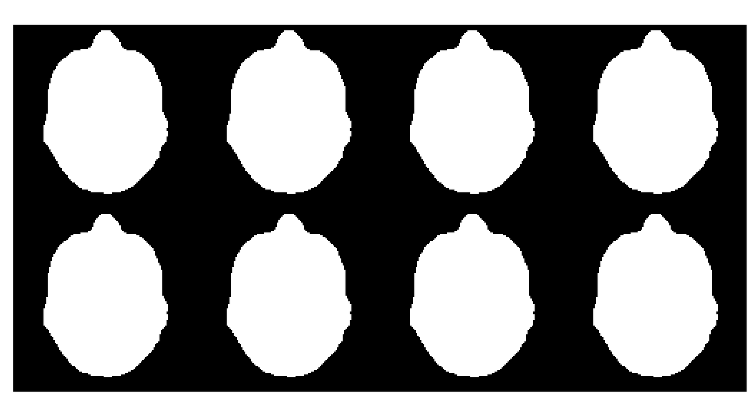
▶ **Regularization function:**
$R(\boldsymbol{x}) = \sum_{p=1}^{4M} \varrho^{(p)}(x_{\mathcal{R}}^{(p)}, x_{\mathcal{I}}^{(p)})$, where

$$(\forall p \in \mathbb{E}) \quad \varrho^{(p)} = \iota_{\{(0,0)\}},$$
$$(\forall p \notin \mathbb{E}) \quad \varrho^{(p)} = \theta_p\|\cdot - \boldsymbol{\omega}_p\|_2^{\kappa_p},$$

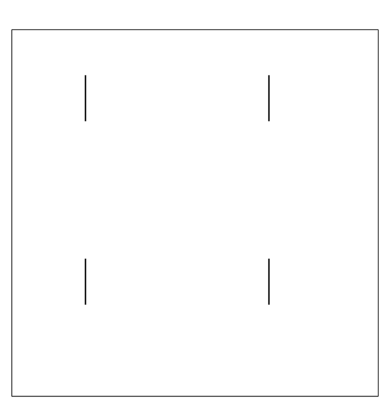
with $\kappa_p \in \{1,2\}$, $\theta_p \in]0,+\infty[$ and $\boldsymbol{\omega}_p \in \mathbb{R}^2$.

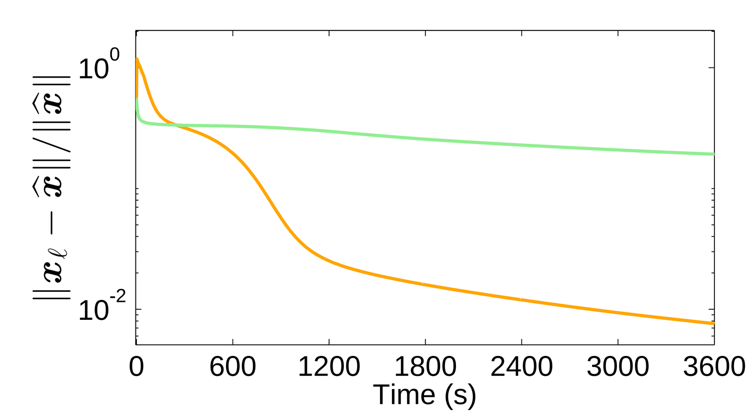


$\overline{\boldsymbol{x}}_{\mathcal{R}}$

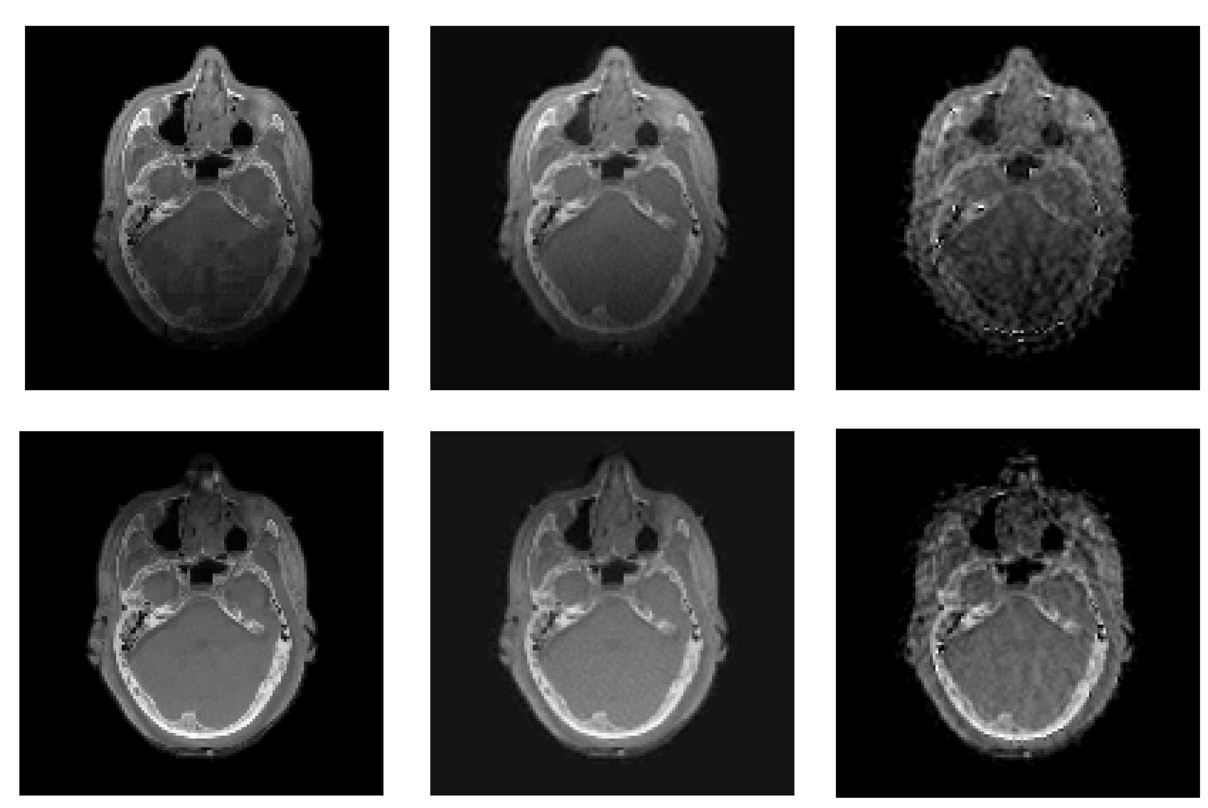


Object background $\mathbb{E}$





(left) Indices of a block $\boldsymbol{x}_{\mathcal{R}}^{(j)} \in \mathbb{R}^{4Q}$ for $Q = 32$.
(right) Convergence profile of the proposed algorithm and its non-preconditioned variant from **[Bolte *et al.*–2014]**.



(left) Real and imaginary parts of the original image $\overline{\boldsymbol{v}} \in \mathbb{C}^M$, with $M = 128\times 128$, and estimated images $\widehat{\boldsymbol{v}}$ using either (middle) the proposed method, SNR $= 21.27$ dB or (right) the regularized alternating projection method from **[Mukherjee *et al.*–2012]**, SNR $= 14.45$ dB.