A Penalized Weighted Least Squares Approach For Restoring Data Corrupted With Signal-Dependent Noise

Audrey Repetti, Emilie Chouzenoux and Jean-Christophe Pesquet

Université Paris-Est, Laboratoire d'Informatique Gaspard Monge, France {first.last}@univ-paris-est.fr

INTRODUCTION

★ Primal-dual proximal splitting approach for convex optimization

★ Fast convergence thanks to a preconditioning strategy

★ Restoration of images corrupted with additive Gaussian noise with signal-dependent variance

PROBLEM

$$\underset{\boldsymbol{x}\in\mathcal{H}}{\text{minimize}}\ f(\boldsymbol{x}) = h(\boldsymbol{x}) + g_0(\boldsymbol{x}) + \sum_{j=1}^{J} g_j(\boldsymbol{L}_j\boldsymbol{x}) \qquad (1)$$

- $\mathcal{H}$ and $(\mathcal{G}_j)_{1\leqslant j\leqslant J}$ real Hilbert spaces
- $h\colon \mathcal{H}\to\mathbb{R}$, convex, differentiable with Lipschitzian gradient
- g_0 and $(g_j)_{1\leqslant j\leqslant J}$ proper lsc convex functions defined resp. on $\mathcal{H}$ and $(\mathcal{G}_j)_{1\leqslant j\leqslant J}$
- $\forall j\in\{1,\ldots,J\}$, $\boldsymbol{L}_j\colon\mathcal{H}\to\mathcal{G}_j$ non-zero bounded linear operator

CONVEX OPTIMIZATION TOOLS

For $\boldsymbol{R}$ positive definite self-adjoint linear operator from $\mathcal{H}$ to $\mathcal{H}$:

- **Weighted norm**: $\forall \boldsymbol{x}\in\mathcal{H}$, $\|\boldsymbol{x}\|_{\boldsymbol{R}} = \langle \boldsymbol{x} \mid \boldsymbol{R}\boldsymbol{x}\rangle^{1/2}$
- **Proximal operator**: ψ proper lsc convex function defined on $(\mathcal{H}, \|\cdot\|_{\boldsymbol{R}})$, $\forall \boldsymbol{v}\in\mathcal{H}$, $\text{prox}_{\boldsymbol{R},\psi}(\boldsymbol{v}) = \arg\min_{\boldsymbol{\xi}\in\mathcal{H}} \psi(\boldsymbol{\xi}) + \frac{1}{2}\|\boldsymbol{\xi}-\boldsymbol{v}\|^2_{\boldsymbol{R}}$

ALGORITHM
Preconditioned M+L FBF

Let $(\gamma_k)_{k\in\mathbb{N}}$ be a sequence of $[\varepsilon,(1-\varepsilon)/\tau]$ with $\varepsilon\in(0,1/(\tau+1))$, $\tau=\mu^{(\boldsymbol{Q})}+\sqrt{\sum_{j=1}^{J}\|\boldsymbol{R}_j^{1/2}\boldsymbol{L}_j\boldsymbol{Q}^{1/2}\|^2}$, where $\mu^{(\boldsymbol{Q})}$ is a Lipschitz constant of $\nabla\left(h\circ\boldsymbol{Q}^{1/2}\right)$.

Initialization: Let $\boldsymbol{x}_0\in\mathcal{H}$, and, for every $j\in\{1,\ldots,J\}$, let $\boldsymbol{v}_{j,0}\in\mathcal{G}_j$

Iterations:

For $k=0,\ldots$

$$\begin{aligned}
&\boldsymbol{y}_{1,k} = \boldsymbol{x}_k - \gamma_k\boldsymbol{Q}\big(\nabla h(\boldsymbol{x}_k) + \textstyle\sum_{j=1}^{J}\boldsymbol{L}_j^*\boldsymbol{v}_{j,k}\big)\\
&\boldsymbol{p}_{1,k} = \text{prox}_{\boldsymbol{Q}^{-1},\gamma_k g_0}(\boldsymbol{y}_{1,k}) \qquad (2)\\
&\text{For } j=1,\ldots,J\\
&\quad \boldsymbol{y}_{2,j,k} = \boldsymbol{v}_{j,k} + \gamma_k\boldsymbol{R}_j\boldsymbol{L}_j\boldsymbol{x}_k\\
&\quad \boldsymbol{p}_{2,j,k} = \text{prox}_{\boldsymbol{R}_j^{-1},\gamma_k g_j^*}(\boldsymbol{y}_{2,j,k}) \qquad (3)\\
&\quad \boldsymbol{q}_{2,j,k} = \boldsymbol{p}_{2,j,k} + \gamma_k\boldsymbol{R}_j\boldsymbol{L}_j\boldsymbol{p}_{1,k}\\
&\quad \boldsymbol{v}_{j,k+1} = \boldsymbol{v}_{j,k} - \boldsymbol{y}_{2,j,k} + \boldsymbol{q}_{2,j,k}.\\
&\boldsymbol{q}_{1,k} = \boldsymbol{p}_{1,k} - \gamma_k\boldsymbol{Q}\big(\nabla h(\boldsymbol{p}_{1,k}) + \textstyle\sum_{j=1}^{J}\boldsymbol{L}_j^*\boldsymbol{p}_{2,j,k}\big)\\
&\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \boldsymbol{y}_{1,k} + \boldsymbol{q}_{1,k}.
\end{aligned}$$

⋆ **Low computational cost of (2) and (3) for**

$$\begin{cases} g_0 \text{ separable and } \boldsymbol{Q} \text{ positive diagonal,}\\ \forall j\in\{1,\ldots,J\},\ \boldsymbol{R}_j = \rho_j\,\text{Id},\ \rho_j>0\end{cases}$$

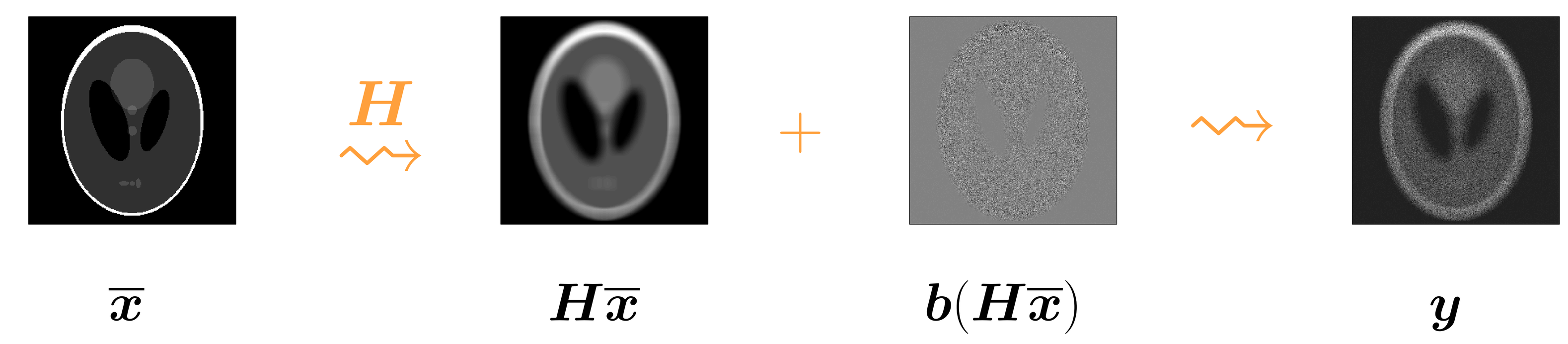
▶ $\begin{cases}\boldsymbol{Q}=\text{Id}\\ \boldsymbol{R}_j=\text{Id}\end{cases}$ ⇒ M+L FBF Algorithm [Combettes and Pesquet 2012]

▶ Appropriate choice for $\boldsymbol{Q}$, $\boldsymbol{R}_j$ ⇒ Acceleration of convergence rate

CONVERGENCE RESULT

Theorem: Under appropriate technical assumptions, there exists $\hat{\boldsymbol{x}}\in\mathcal{H}$ solution to Problem (1) such that $\boldsymbol{x}_k\to\hat{\boldsymbol{x}}$ and $\boldsymbol{p}_{1,k}\to\hat{\boldsymbol{x}}$.

SIGNAL-DEPENDENT NOISE MODEL

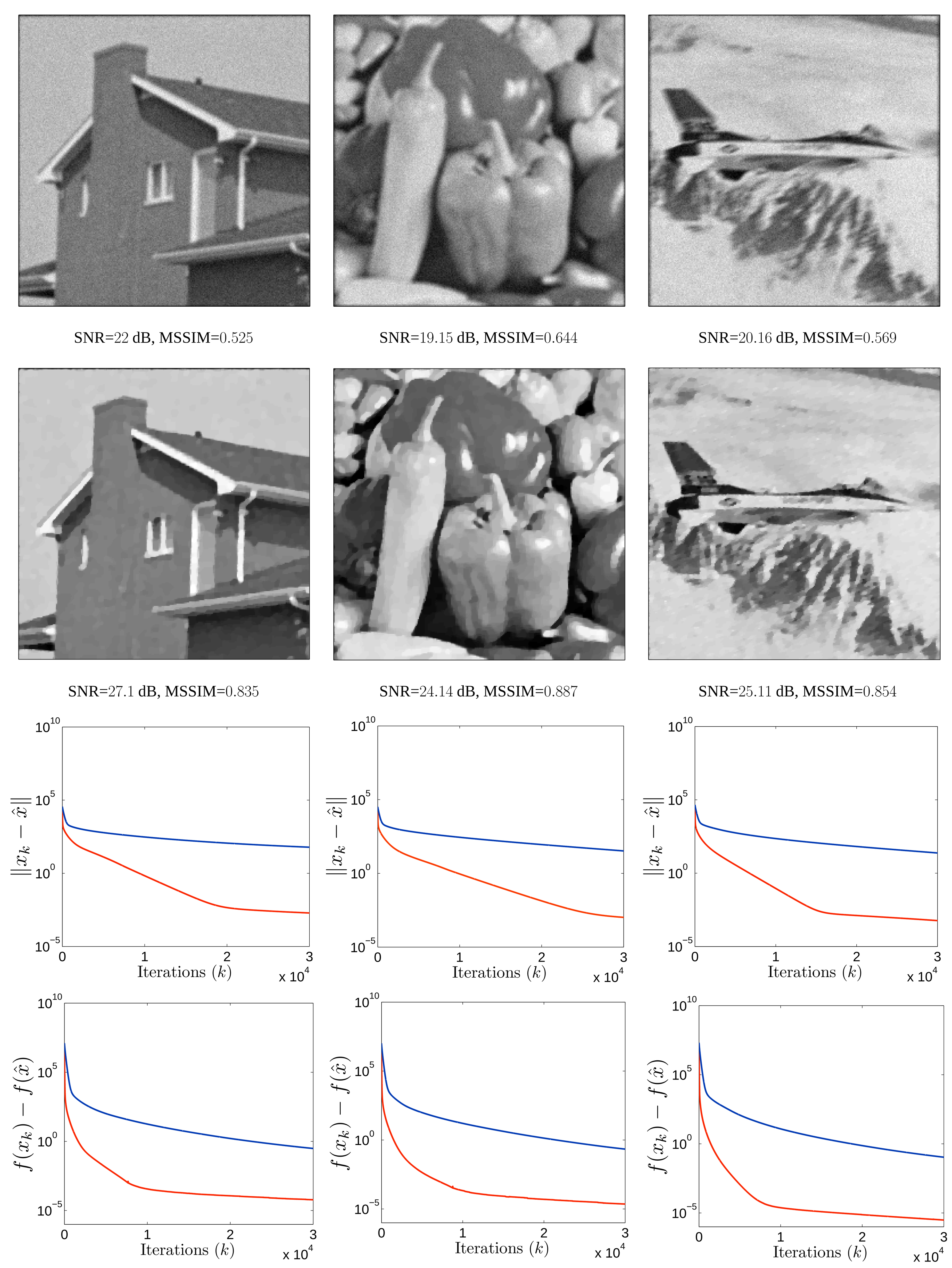


- $\boldsymbol{H}\in\mathbb{R}^{N\times N}$ observation matrix with non-negative elements
- $\boldsymbol{b}(\boldsymbol{H}\overline{\boldsymbol{x}}) = (b_n([\boldsymbol{H}\overline{\boldsymbol{x}}]_n))_{1\leqslant n\leqslant N}$, $b_n\colon[0,+\infty)\to[0,+\infty)\colon z_n\mapsto\sqrt{\alpha z_n+\beta}\,w_n$, with $\alpha\geqslant 0$, $\beta>0$ and $(w_n)_{1\leqslant n\leqslant N}$ realization of $\boldsymbol{W}\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{I}_N)$

⋆ **Data fidelity term:**

$$\forall\boldsymbol{x}\in[0,+\infty)^N,\ h(\boldsymbol{x}) = \frac{1}{2}\sum_{n=1}^{N}\frac{(y_n-[\boldsymbol{H}\boldsymbol{x}]_n)^2}{\alpha[\boldsymbol{H}\boldsymbol{x}]_n+\beta}$$

⋆ $g_0=\iota_{[0,255]^N}$, $J=1$ and $g_1=\lambda\,\text{tv}$, with $\lambda>0$ regularization parameter



From top to bottom: *noisy blurred images ($\alpha=0.1$ and $\beta=50$), restored images, $(\|\boldsymbol{x}_k-\hat{\boldsymbol{x}}\|)_k$ and $(f(\boldsymbol{x}_k)-f(\hat{\boldsymbol{x}}))_k$ using the proposed algorithm and its non preconditioned version.*