A Decoupled Approach for Composite Sparse-plus-Smooth Penalized Optimization

Adrian Jarret, Valérie Costa, Julien Fageot

Laboratory of AudioVisual Communications (LCAV)
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Context and Motivation

Goal: Composite Inverse Problem

Solve for $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^N$ the problem

$$\mathbf{y} = \mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2) + \mathbf{w}$$

$\mathbf{w}$ is AWGN $\quad \mathbf{y}, \mathbf{w} \in \mathbb{R}^L$

Tool: Sparse-plus-Smooth optimization

$$\boxed{\mathcal{V} = \arg\min_{\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^N} \frac{1}{2}\|\mathbf{y} - \mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2)\|_2^2 + \lambda_1 \|\mathbf{L}_1 \mathbf{x}_1\|_1 + \frac{\lambda_2}{2}\|\mathbf{L}_2 \mathbf{x}_2\|_2^2}$$

Classical solvers:

- Coupled proximal solvers (PGD, FB, ...)
  $\mathbf{x} = \begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{bmatrix} \in \mathbb{R}^{2N}$
  ↳ High dimension
- Alternating minimization algorithms
  $\partial_{\mathbf{x}_1}(\cdot), \quad \partial_{\mathbf{x}_2}(\cdot)$
  ↳ Convergence tricky to guarantee

Decoupled Representer Theorem

Requirement

$\ker(\mathbf{A})^\perp$ is invariant by $\mathbf{L}_2^T \mathbf{L}_2$,
i.e., $\mathbf{x} \in \ker(\mathbf{A})^\perp \Rightarrow \mathbf{L}_2^T \mathbf{L}_2 \mathbf{x} \in \ker(\mathbf{A})^\perp$.

Definitions

$$\left.\begin{array}{l} \mathbf{\Lambda_2} = \left(\mathbf{A}\mathbf{A}^T\right)^{-1} \mathbf{A}\mathbf{L}_2^T \mathbf{L}_2 \mathbf{A}^T \\ \mathbf{M}_{\lambda_2} = \lambda_2 \mathbf{\Lambda_2} \left(\mathbf{A}\mathbf{A}^T + \lambda_2 \mathbf{\Lambda_2}\right)^{-1} \end{array}\right\} \in \mathbb{R}^{L \times L}$$

Solution set

$$\boxed{\mathcal{V} = \mathcal{V}_1 \times \mathcal{V}_2}$$

LASSO-type problem 1°

$$\mathbf{x}_1^* \in \mathcal{V}_1 = \arg\min_{\mathbf{x}_1 \in \mathbb{R}^N} P_1(\mathbf{x}_1)$$

$$P_1(\mathbf{x}_1) = \left\{ \begin{array}{r} \frac{1}{2}(\mathbf{y} - \mathbf{A}\mathbf{x}_1)^T \mathbf{M}_{\lambda_2} (\mathbf{y} - \mathbf{A}\mathbf{x}_1) \\ + \lambda_1 \|\mathbf{L}_1 \mathbf{x}_1\|_1 \end{array} \right\}$$

Quadratic problem 2°

$$\mathbf{x}_2^* = \arg\min_{\mathbf{x}_2 \in \mathbb{R}^N} P_2(\mathbf{x}_2)$$

$$P_2(\mathbf{x}_2) = \left\{ \begin{array}{r} \frac{1}{2}\|\mathbf{y} - \tilde{\mathbf{y}} - \mathbf{A}\mathbf{x}_2\|_2^2 \\ + \frac{\lambda_2}{2}\|\mathbf{L}_2 \mathbf{x}_2\|_2^2 \end{array} \right\}$$

Uniqeness properties $\longrightarrow$ $\forall \mathbf{x}_1^* \in \mathcal{V}_1, \quad \mathbf{A}\mathbf{x}_1^* = \tilde{\mathbf{y}} \in \mathbb{R}^L$ $\longrightarrow$ $\mathcal{V}_2 = \left\{\mathbf{A}^T \left(\mathbf{A}\mathbf{A}^T + \lambda_2 \mathbf{\Lambda_2}\right)^{-1} (\mathbf{y} - \tilde{\mathbf{y}})\right\}$

Reconstruction of Simulated Decoupled Composite Problems

Foreground | Background | Measurements | Recovered foreground composite | Recovered foreground single component

Case 1: Image decomposition

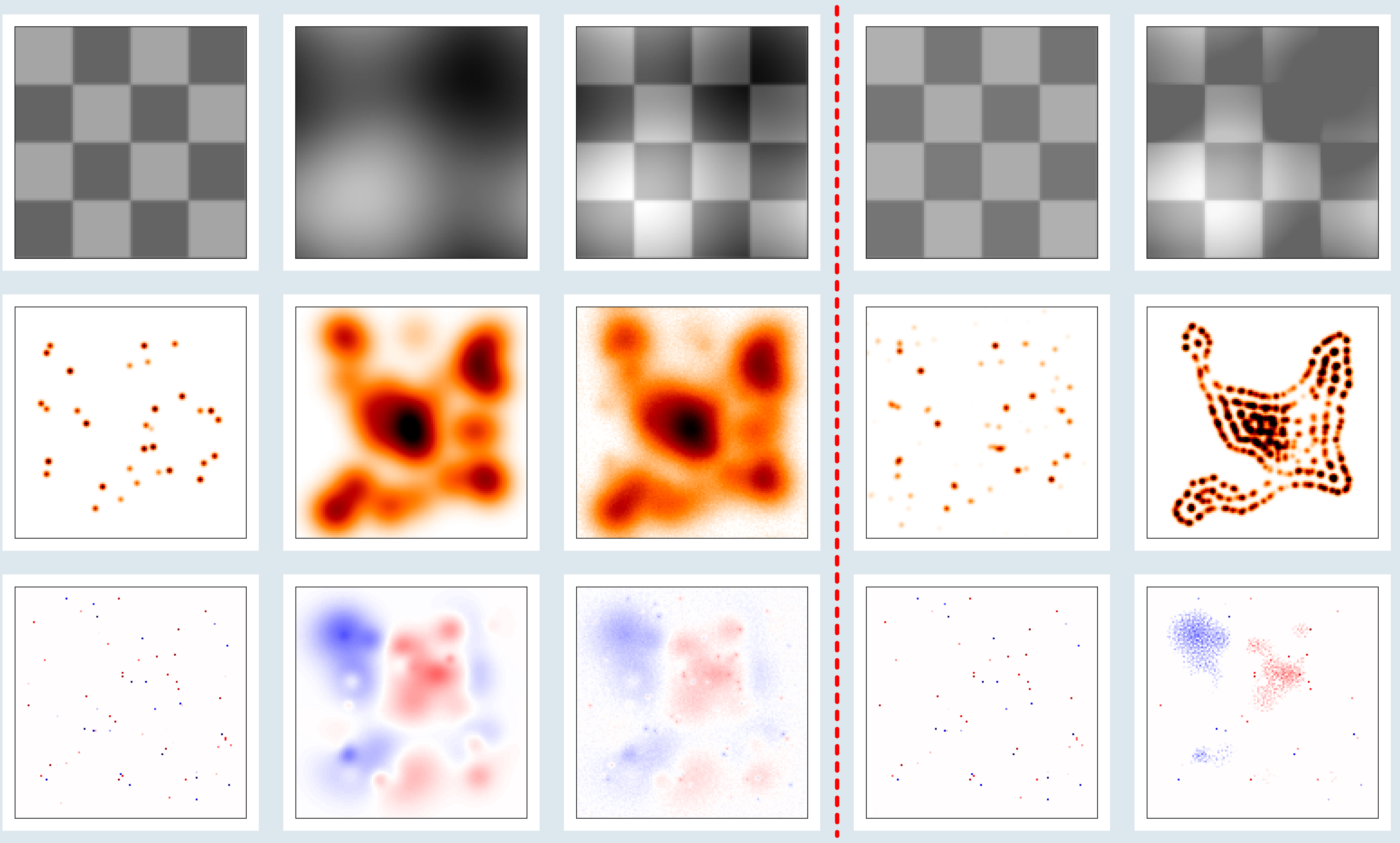
$$\mathbf{A} = \mathbf{I}_N \quad \begin{bmatrix} \mathbf{L}_1 = \nabla \\ \mathbf{L}_2 = \Delta \end{bmatrix}$$

Case 2: Image deconvolution

$$\mathbf{A}\mathbf{x} = \mathbf{g} * \mathbf{x} \quad \|\mathbf{g}\|_1 = 1 \quad \begin{bmatrix} \mathbf{L}_1 = \mathbf{I}_N \\ \mathbf{L}_2 = \Delta \end{bmatrix}$$

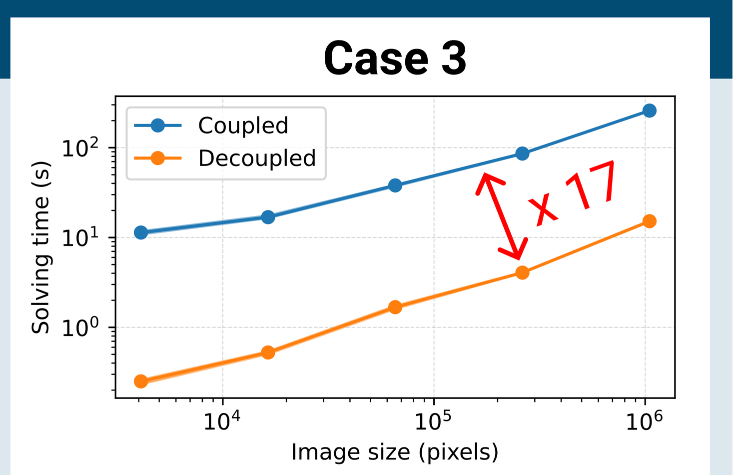
Case 3: Image recovery

$$\mathbf{A}\mathbf{x} = \mathbf{SFx} = \hat{\mathbf{x}}[i_1, \dots, i_L] \quad \begin{bmatrix} \mathbf{L}_1 = \mathbf{I}_N \\ \mathbf{L}_2 = \Delta \end{bmatrix}$$



Time Performances

Time (s)	N	L	**Decoupled**	Coupled	Non-Composite
Case 1	100^2	100^2	**39.0**	98.6	1.1
Case 2	128^2	128^2	**55.8**	421.1	53.8
Case 3	128^2	$\sim 5k$	**0.46**	31.84	1.26



Summary

✓ **Composite problems** are **powerful** tools to model signals with **background**.

✓ The two components are **decoupled** in some cases, enabling **efficient numerical methods** to be developed.

This work was funded by the Swiss National Science Foundation with grant "SESAM - Sensing and Sampling: Theory and Algorithms" (n° 200021 181978/1).