Towards optimal algorithms for the recovery of low-dimensional models with linear rates

---

Vers des algorithmes optimaux pour la reconstruction de modèles de faible dimension

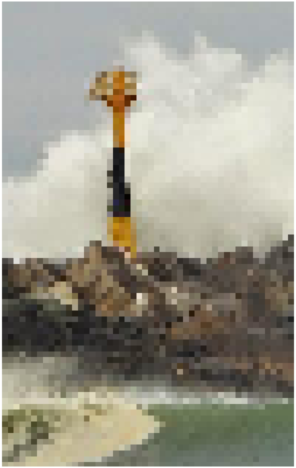
---

SMAI 2025

Yann Traonmilin (CNRS,IMB)

Collaborators: J.-F. Aujol (IMB), R. Gribonval (INRIA Lyon), A. Guennec (IMB), A. Joundi (IMB), A. Newson (ISIR), S. Vaiter (CNRS Nice)

An inverse problem



Low resolution image

High resolution image

A simple formalism

- Observation of $\hat{x} \in \mathbb{R}^N$

$$y = A\hat{x}$$

- Linear inverse problem in high dimension
- Matrix A under-determined, finite number m of measurements
- In this presentation: noiseless case

Dimension reduction and low complexity

- Low-dimensional regularity model Σ (homogeneous set):
$$\hat{x} \in \Sigma$$
- i.e. $\hat{x}$ can be described by few parameters

Examples:

- Image super-resolution: $\nabla \hat{x}$ is a sparse gradient
- Generalized sparsity, low-rank models (matrices, tensors)
- Deep priors (modeled with deep neural networks)

Objective

Given a low dimensional model Σ, what is the best method to recover $\hat{x}$ from y ?

$\Rightarrow$ Problem: define "best"; specify "method"

$\Rightarrow$ Must consider both explicit and *learned* models

$\Rightarrow$ In this talk the "non-convex algorithmic" approach

Accelerated state of the art (convex methods)

The constrained minimization

$$x^* = \arg\min_{x \in \mathbb{R}^N} R(x) \ s.t. \ Ax = y$$

yields a stable and robust estimation of $\hat{x}$ for some operators A, models Σ and convex regularizers R.

- e.g. $\Sigma =$ sparse vectors, $R = \ell^1$-norm, $A =$ Gaussian measurements
- [**T.**, Gribonval, Vaiter, 2024] A theory of optimal convex regularization : maximize compliance between R and Σ
- Application to image decomposition [Guennec, Aujol, **T.**, 2024]

Optimal recovery with non-convex methods?

- Algorithms not necessarily linked with underlying functional (with learned models)
- Idea: build a framework for optimal methods in this context

$\Rightarrow$ Give recovery guarantees for a class of recovery algorithms: Generalized PGD

$\Rightarrow$ Optimize rate of convergence within this class of algorithms

Generalized Projected Gradient Descent

Consider the iterations

$$x_{n+1} = P_\Sigma(x_n) - \mu A^T(AP_\Sigma(x_n) - y).$$

with P_Σ a *generalized* projection onto a set Σ

- Sparse recovery: $P_\Sigma = P_\Sigma^\perp = \mathrm{HT}(\cdot)$, orthogonal PGD = Iterative hard thresholding
- Deep projective prior (auto-encoder): $P_\Sigma = f_D \circ f_E =$ neural network with explicit low-dimensional latent representation
- Deep projective prior (plug and play, diffusion): $P_\Sigma = D =$ denoiser parametrized by a neural network

Restricted isometry constant

The operator B has restricted isometry constant $\delta < 1$ w.r.t Σ if for all $x_1, x_2 \in \Sigma$

$$\|(I - B)(x_1 - x_2)\|_2 \leq \delta \|x_1 - x_2\|_2$$

- Quantifies the "conditioning" of A with respect to Σ through $\delta(\mu A^T A)$
- Classical condition in low-dimensional recovery (close to a necessary condition)

Linear recovery with GPGD

Theorem: [**T.**, Aujol, Guennec, 2024] Suppose $\mu A^T A$ has RIC δ, P_Σ any projection having a **restricted** β-Lipschitz property. Consider GPGD iterations. Then

$$\|x_n - \hat{x}\|_2 \leq (\delta\beta)^n \|x_0 - \hat{x}\|_2.$$

- Conjecture that GPGD are near-optimal in the wider class $x_{n+1} = x_n - \mu(A^T(Ax - y) + g(x_n))$

Restricted β-Lipschitz property

For any $z \in \mathbb{R}^N, x \in \Sigma$,

$$\|P_\Sigma(z) - x\|_2 \leq \beta \|z - x\|$$

- Quantifies global linear convergence rate and identifiability (condition $\delta\beta < 1$)
- Optimal PGD method = minimize $\beta(P_\Sigma)$
- **Theorem:** For homogeneous sets, the orthogonal projection on Σ (if it exists) has a Restricted Lipschitz property $\beta \leq 2$.

Optimality of hard thresholding for sparse recovery?

- $\Sigma = \Sigma_k$, set of k-sparse vectors, $P_\Sigma^\perp(\cdot) = \mathrm{HT}(\cdot)$, orthogonal PGD = Iterative hard thresholding.
- $\mathrm{HT}(\cdot)$ not globally Lipschitz
- **Theorem:** For sparse recovery, $P_\Sigma^\perp(\cdot) = \mathrm{HT}(\cdot)$ has restricted Lipschitz constant $\beta_0 = \sqrt{\frac{3+\sqrt{5}}{2}} \approx 1.618$ (with $\beta \geq 1.567$ for small k) and is **optimal** when considering all possible k
- For a given sparsity, the question is still open (conjecture: surprisingly, not the orthogonal projection in general)

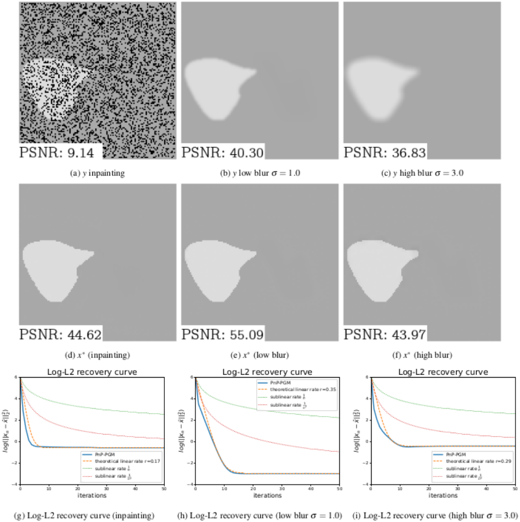
Linear rates for plug and play methods

- Plug and play method = state of the art imaging method
- Use a general purpose *learned* denoiser D (e.g. a U-NET) as a projection operator

$$x_{n+1} = D\left(x_n - \mu A^T(Ax_n - y)\right)$$

- If D is **restricted** Lipschitz with respect to $\Sigma = \mathrm{Fix}(D)$ and A has the RIP w.r.t Σ then we have linear convergence to $\hat{x} \in \mathrm{Fix}(D)$.

Experiments



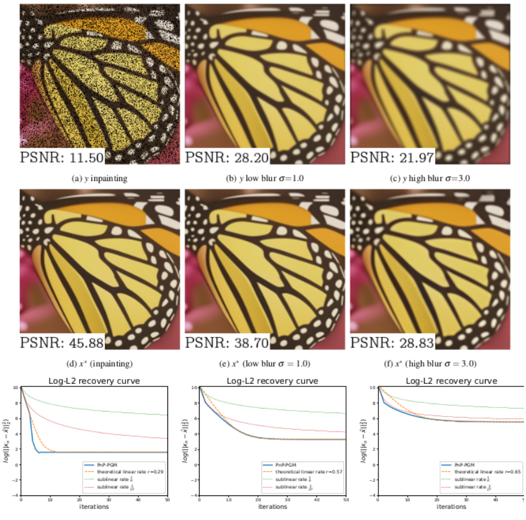
(a) y inpainting (b) y low blur $\sigma = 1.0$ (c) y high blur $\sigma = 3.0$

(d) x^* (inpainting) (e) x^* (low blur) (f) x^* (high blur)

(g) Log-L2 recovery curve (inpainting) (h) Log-L2 recovery curve (low blur $\sigma = 1.0$) (i) Log-L2 recovery curve (high blur $\sigma = 3.0$)

Linear recovery of a fixed point of the PNP-PGM method with a general purpose DRUNET denoiser (synthetic images)

Experiments



(a) y inpainting (b) y low blur σ=1.0 (c) y high blur σ=3.0

(d) x^* (inpainting) (e) x^* (low blur $\sigma = 1.0$) (f) x^* (high blur $\sigma = 3.0$)

(g) Log-L2 recovery curve (inpainting) (h) Log-L2 recovery curve (low blur $\sigma = 1.0$) (i) Log-L2 recovery curve (high blur $\sigma = 3.0$

Linear recovery of a fixed point of PNP-PGM method with a general purpose DRUNET denoiser (natural images)

Optimal recovery of low dimensional models

- Performance guarantees for general class of algorithms, definition of an optimality criterion
- A flexible framework bridging low-dimensional recovery theory and learning-based approaches
- See also [Joundi, Newson, **T.**, 2025]: orthogonal regularization of the learned projection yielding improved recovery

Future work

- Stability analysis
- Generalize class of algorithms

Thanks !

- Towards optimal algorithms for the recovery of low dimensional-models, **T.**, Aujol and Guennec, preprint, 2024.
- Stochastic orthogonal regularization for deep projective priors. Joundi, Newson, **T.**, preprint, 2025.
- A theory of optimal convex regularization for low-dimensional recovery, **T.**, Gribonval and Vaiter, Information and Inference, 2024
- Adaptive Parameter Selection For Gradient-sparse + Low Patch-rank Recovery: Application To Image Decomposition. Guennec, Aujol, **T.**, EUSIPCO 2024.