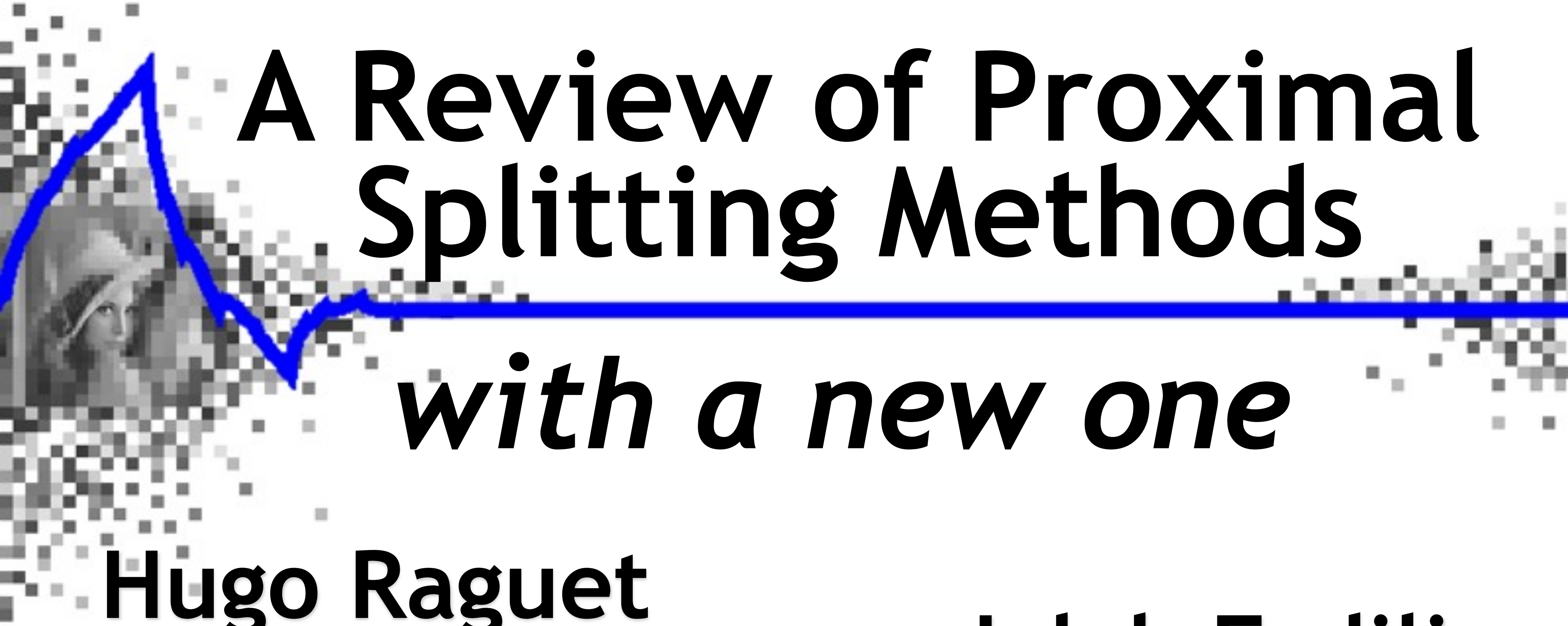


A Review of Proximal Splitting Methods

with a new one

Hugo Raguet
Gabriel Peyré

Jalal Fadili

www.numerical-tours.com

Convex Optimization

Setting: $G : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$

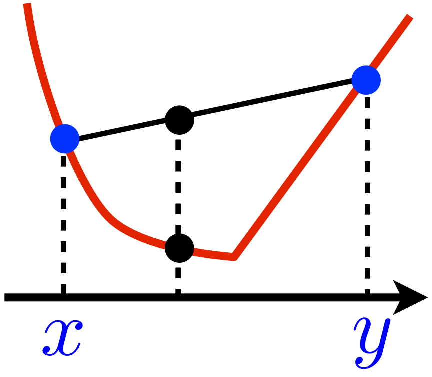
$\mathcal{H}$: Hilbert space. Here: $\mathcal{H} = \mathbb{R}^N$.

Problem: $\min_{x \in \mathcal{H}} G(x)$

Convex Optimization

Setting: $G : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$

$\mathcal{H}$: Hilbert space. Here: $\mathcal{H} = \mathbb{R}^N$.



$$\boxed{\textit{Problem:} \quad \min_{x \in \mathcal{H}} G(x)}$$

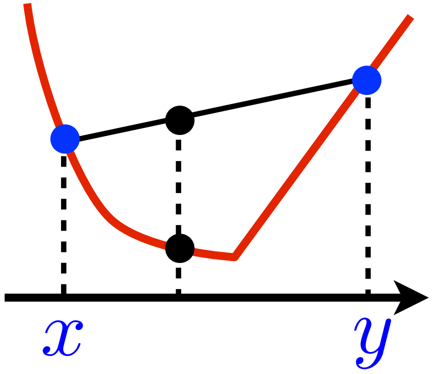
Class of functions:

Convex: $G(tx + (1-t)y) \leqslant tG(x) + (1-t)G(y) \qquad t \in [0,1]$

Convex Optimization

Setting: $G : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$

$\mathcal{H}$: Hilbert space. Here: $\mathcal{H} = \mathbb{R}^N$.

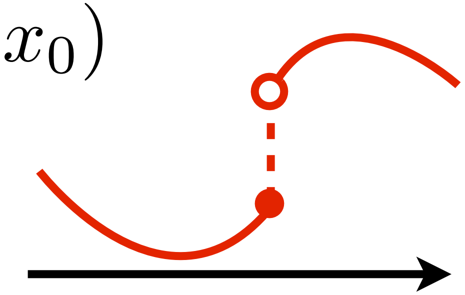


Problem: $\min_{x \in \mathcal{H}} G(x)$

Class of functions:

Convex: $G(tx + (1-t)y) \leqslant tG(x) + (1-t)G(y) \qquad t \in [0,1]$

Lower semi-continuous: $\liminf_{x \to x_0} G(x) \geqslant G(x_0)$



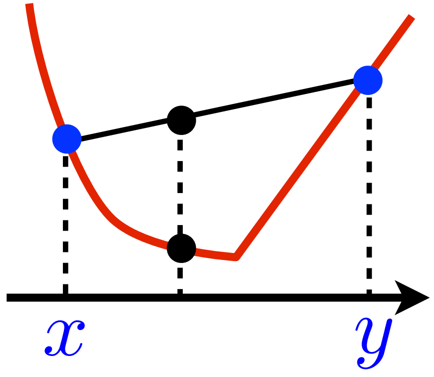
Proper: $\{x \in \mathcal{H} \setminus G(x) \neq +\infty\} \neq \emptyset$

Convex Optimization

Setting: $G : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$

$\mathcal{H}$: Hilbert space. Here: $\mathcal{H} = \mathbb{R}^N$.

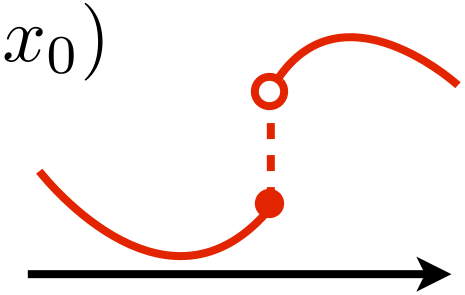
Problem: $\min_{x \in \mathcal{H}} G(x)$



Class of functions:

Convex: $G(tx + (1-t)y) \leqslant tG(x) + (1-t)G(y) \quad t \in [0,1]$

Lower semi-continuous: $\liminf_{x \to x_0} G(x) \geqslant G(x_0)$



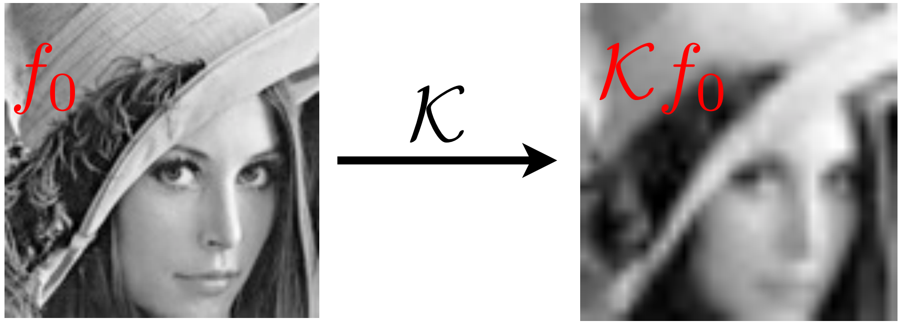
Proper: $\{x \in \mathcal{H} \setminus G(x) \neq +\infty\} \neq \emptyset$

Indicator: $$\iota_{\mathcal{C}}(x) = \begin{cases} 0 & \text{if} \quad x \in \mathcal{C}, \\ +\infty & \text{otherwise.} \end{cases}$$

($\mathcal{C}$ closed and convex)

Example: ℓ^1 Regularization

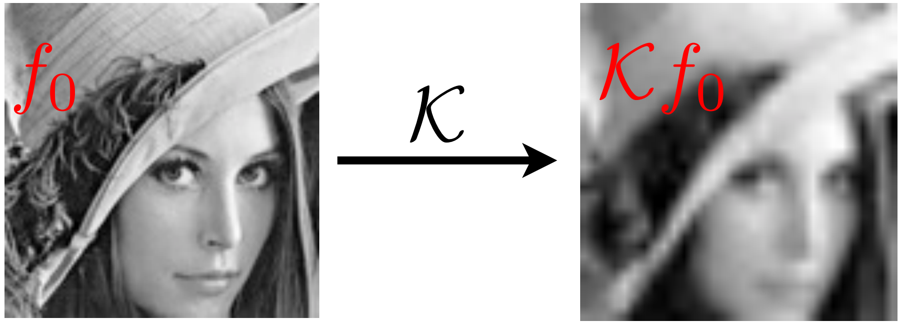
Inverse problem: measurements $y = \mathcal{K} f_0 + w$



$$\mathcal{K} : \mathbb{R}^N \rightarrow \mathbb{R}^P, \quad P \leqslant N$$

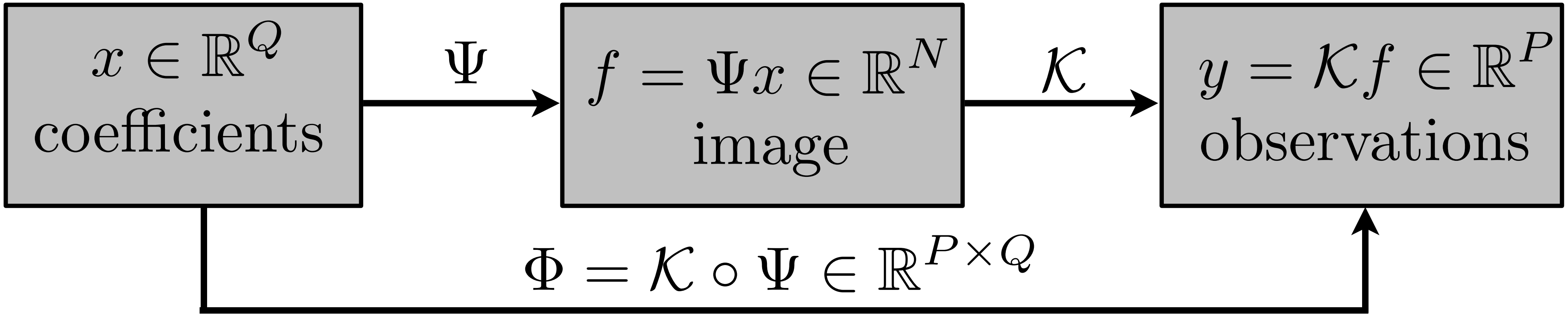
Example: ℓ^1 Regularization

Inverse problem: measurements $y = \mathcal{K} f_0 + w$



$$\mathcal{K} : \mathbb{R}^N \rightarrow \mathbb{R}^P, \quad P \leqslant N$$

Model: $f_0 = \Psi x_0$ sparse in dictionary $\Psi \in \mathbb{R}^{N \times Q}, Q \geqslant N$.



$x \in \mathbb{R}^Q$
coefficients

Ψ

$f = \Psi x \in \mathbb{R}^N$
image

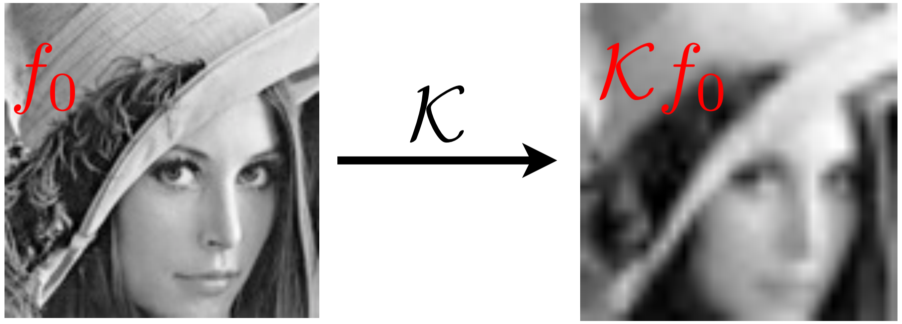
$\mathcal{K}$

$y = \mathcal{K} f \in \mathbb{R}^P$
observations

$\Phi = \mathcal{K} \circ \Psi \in \mathbb{R}^{P \times Q}$

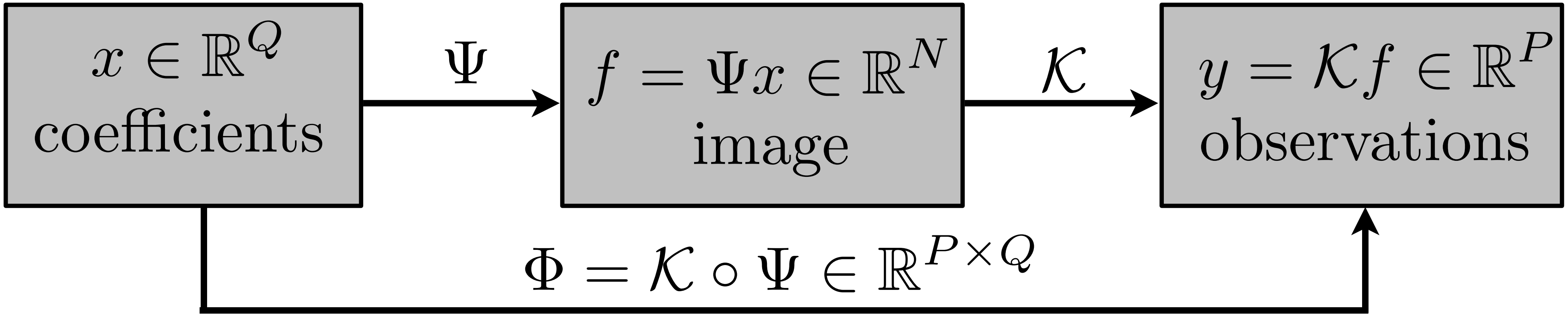
Example: ℓ^1 Regularization

Inverse problem: measurements $y = \mathcal{K} f_0 + w$



$\mathcal{K} : \mathbb{R}^N \rightarrow \mathbb{R}^P, \quad P \leqslant N$

Model: $f_0 = \Psi x_0$ sparse in dictionary $\Psi \in \mathbb{R}^{N \times Q}, Q \geqslant N$.



$x \in \mathbb{R}^Q$
coefficients

Ψ

$f = \Psi x \in \mathbb{R}^N$
image

$\mathcal{K}$

$y = \mathcal{K} f \in \mathbb{R}^P$
observations

$\Phi = \mathcal{K} \circ \Psi \in \mathbb{R}^{P \times Q}$

Sparse recovery: $f^\star = \Psi x^\star$ where $x^\star$ solves

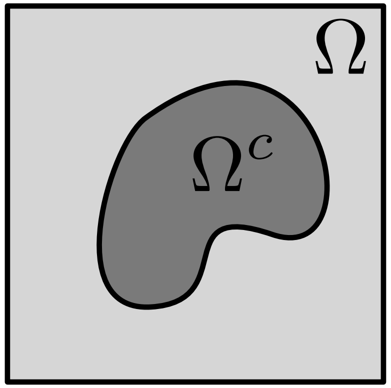
$$\min_{x \in \mathbb{R}^N} \underset{\text{Fidelity}}{\boxed{\frac{1}{2}\|y - \Phi x\|^2}} + \underset{\text{Regularization}}{\boxed{\lambda \|x\|_1}}$$

Example: ℓ^1 Regularization

Inpainting: masking operator $\mathcal{K}$

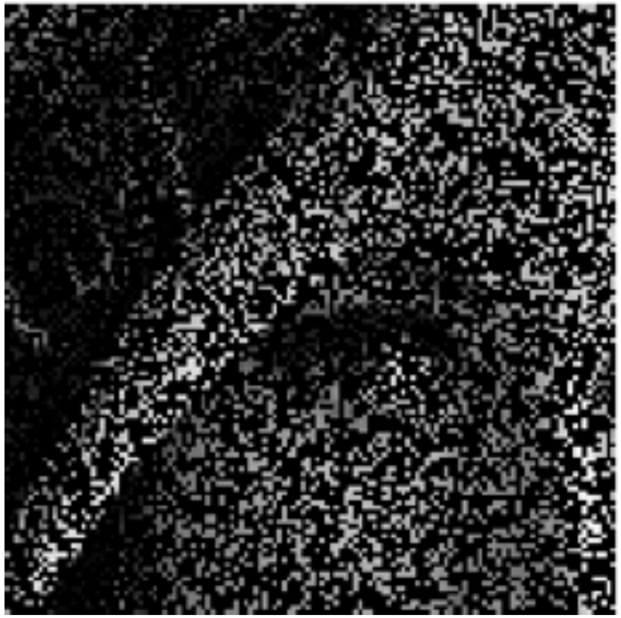
$$(\mathcal{K}f)_i = \begin{cases} f_i & \text{if} \quad i \in \Omega, \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathcal{K} : \mathbb{R}^N \to \mathbb{R}^P \qquad P = |\Omega|$$



$\Psi \in \mathbb{R}^{N \times Q}$ translation invariant wavelet frame.

Orignal $f_0 = \Psi x_0$



$y = \Phi x_0 + w$

Recovery $\Psi x^\star$

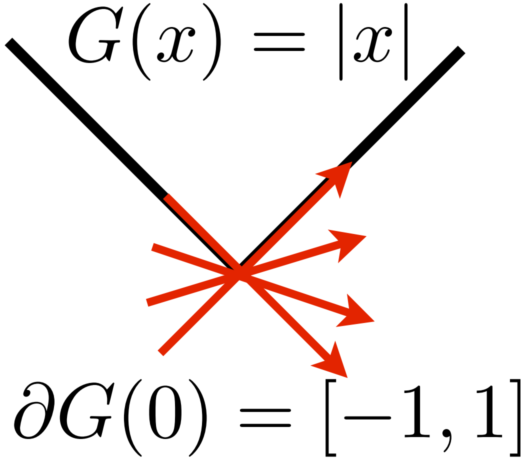
Overview

- Subdifferential Calculus
- Proximal Calculus
- Forward Backward
- Douglas Rachford
- Generalized Forward-Backward

Sub-differential

Sub-differential:

$$\partial G(x) = \{u \in \mathcal{H} \setminus \forall\, z,\; G(z) \geqslant G(x) + \langle u,\; z - x\rangle\}$$



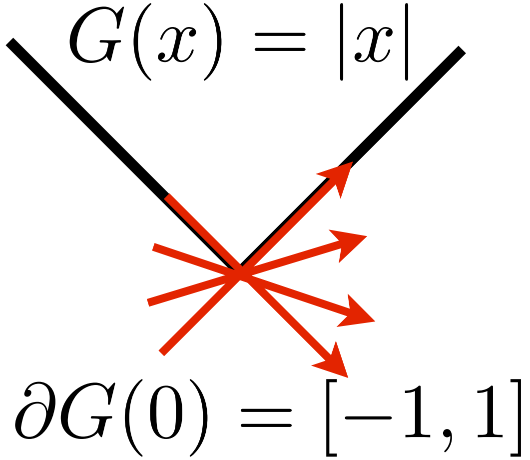
Sub-differential

Sub-differential:

$$\partial G(x) = \{u \in \mathcal{H} \setminus \forall\, z,\; G(z) \geqslant G(x) + \langle u,\; z - x \rangle\}$$

Smooth functions:

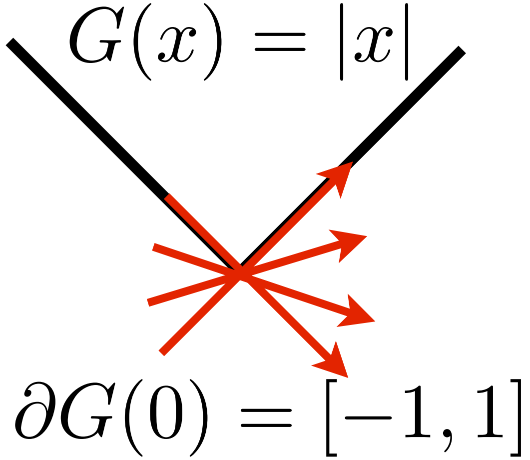
If F is C^1, $\partial F(x) = \{\nabla F(x)\}$



Sub-differential

Sub-differential:

$$\partial G(x) = \{u \in \mathcal{H} \setminus \forall\, z,\; G(z) \geqslant G(x) + \langle u,\; z - x\rangle\}$$



Smooth functions:

If F is C^1, $\partial F(x) = \{\nabla F(x)\}$

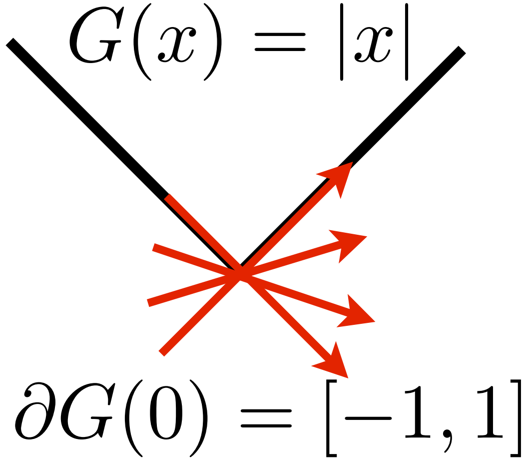
First-order conditions:

$$x^\star \in \underset{x \in \mathcal{H}}{\operatorname{argmin}}\; G(x) \quad\Longleftrightarrow\quad 0 \in \partial G(x^\star)$$

Sub-differential

Sub-differential:

$$\partial G(x) = \{u \in \mathcal{H} \setminus \forall\, z,\; G(z) \geqslant G(x) + \langle u,\; z - x\rangle\}$$

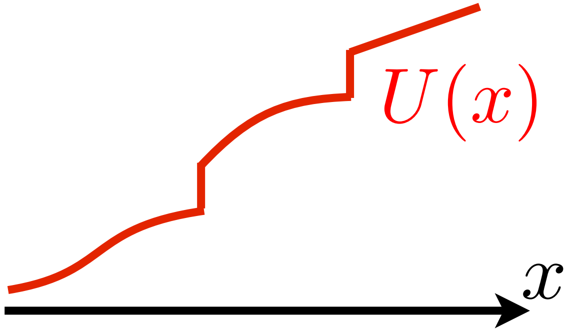


Smooth functions:

$$\text{If } F \text{ is } C^1,\; \partial F(x) = \{\nabla F(x)\}$$

First-order conditions:

$$x^\star \in \underset{x \in \mathcal{H}}{\text{argmin}}\; G(x) \quad \Longleftrightarrow \quad 0 \in \partial G(x^\star)$$



Monotone operator: $\quad U(x) = \partial G(x)$

$$\forall\, (u, v) \in U(x) \times U(y), \quad \langle y - x,\; v - u\rangle \geqslant 0$$

Example: ℓ^1 Regularization

$$x^\star \in \underset{x \in \mathbb{R}^Q}{\text{argmin}}\; G(x) = \frac{1}{2}\|y - \Phi x\|^2 + \lambda \|x\|_1$$

$$\partial G(x) = \Phi^*(\Phi x - y) + \lambda \partial \|\cdot\|_1(x)$$

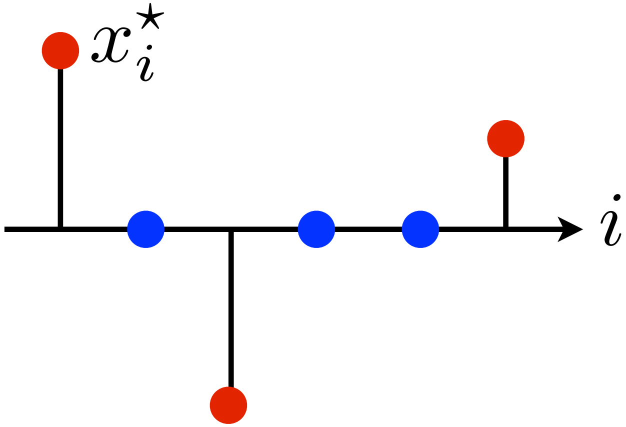
$$\partial \|\cdot\|_1(x)_i = \begin{cases} \text{sign}(x_i) & \text{if} \quad x_i \neq 0, \\ [-1, 1] & \text{if} \quad x_i = 0. \end{cases}$$

Example: ℓ^1 Regularization

$$x^\star \in \underset{x \in \mathbb{R}^Q}{\text{argmin}}\ G(x) = \frac{1}{2}\|y - \Phi x\|^2 + \lambda \|x\|_1$$

$$\partial G(x) = \Phi^*(\Phi x - y) + \lambda \partial \| \cdot \|_1(x)$$

$$\partial \| \cdot \|_1(x)_i = \begin{cases} \text{sign}(x_i) & \text{if} \quad x_i \neq 0, \\ [-1, 1] & \text{if} \quad x_i = 0. \end{cases}$$



Support of the solution: ●

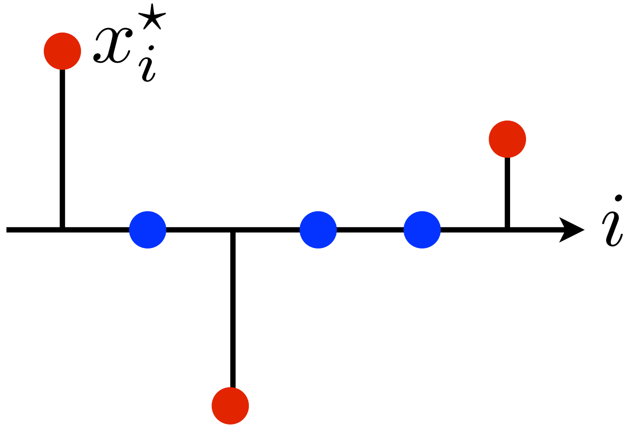
$$I = \{ i \in \{0, \ldots, N-1\} \setminus x_i^\star \neq 0 \}$$

Example: ℓ^1 Regularization

$$x^\star \in \underset{x \in \mathbb{R}^Q}{\operatorname{argmin}} G(x) = \frac{1}{2}\|y - \Phi x\|^2 + \lambda \|x\|_1$$

$$\partial G(x) = \Phi^*(\Phi x - y) + \lambda \partial \| \cdot \|_1(x)$$

$$\partial \| \cdot \|_1(x)_i = \begin{cases} \operatorname{sign}(x_i) & \text{if} \quad x_i \neq 0, \\ [-1, 1] & \text{if} \quad x_i = 0. \end{cases}$$



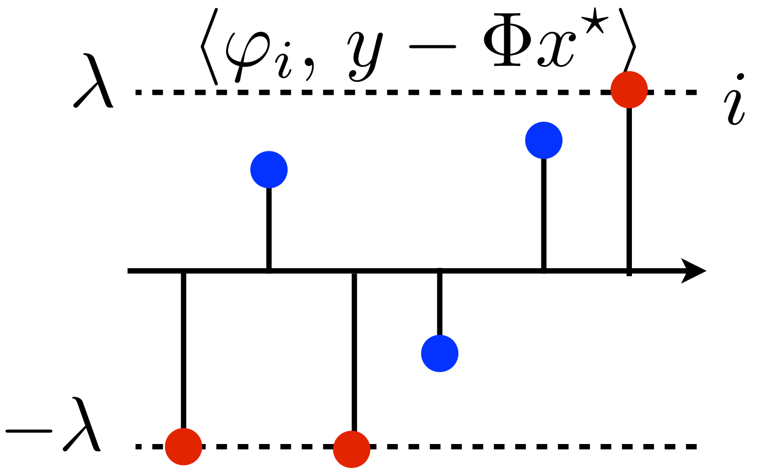
Support of the solution: ●

$$I = \{ i \in \{0, \ldots, N-1\} \setminus x_i^\star \neq 0 \}$$

First-order conditions:

$$\exists s \in \mathbb{R}^N, \quad \Phi^*(\Phi x^\star - y) + \lambda s = 0$$

$$\begin{cases} s_I = \operatorname{sign}(x_I), \\ \|s_{I^c}\|_\infty \leqslant 1. \end{cases}$$



Overview

- Subdifferential Calculus
- **Proximal Calculus**
- Forward Backward
- Douglas Rachford
- Generalized Forward-Backward

Proximal Operators

Proximal operator of G:

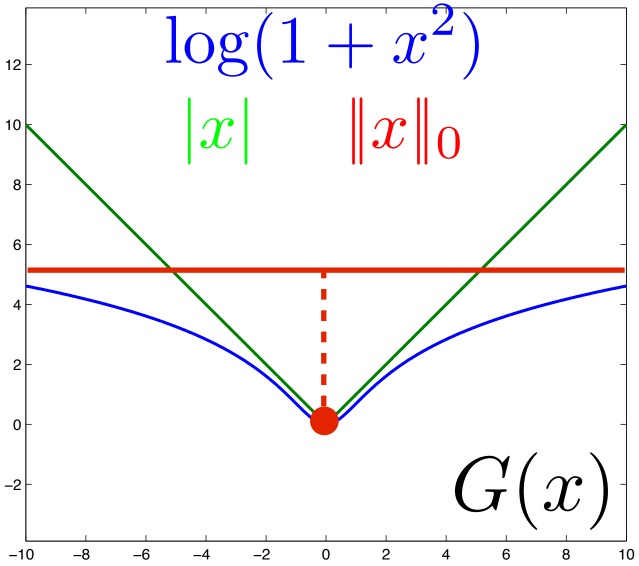
$$\mathrm{Prox}_{\gamma G}(x) = \underset{z}{\mathrm{argmin}} \ \frac{1}{2}\|x - z\|^2 + \gamma G(z)$$

Proximal Operators

Proximal operator of G:

$$\mathrm{Prox}_{\gamma G}(x) = \underset{z}{\operatorname{argmin}} \frac{1}{2}\|x - z\|^2 + \gamma G(z)$$

$$G(x) = \|x\|_1 = \sum_i |x_i|$$



$$G(x) = \|x\|_0 = |\,\{i \setminus x_i \neq 0\}\,|$$

$$G(x) = \sum_i \log(1 + |x_i|^2)$$

Proximal Operators

Proximal operator of G:

$$\mathrm{Prox}_{\gamma G}(x) = \underset{z}{\mathrm{argmin}}\ \frac{1}{2}\|x-z\|^2 + \gamma G(z)$$

$$G(x) = \|x\|_1 = \sum_i |x_i|$$

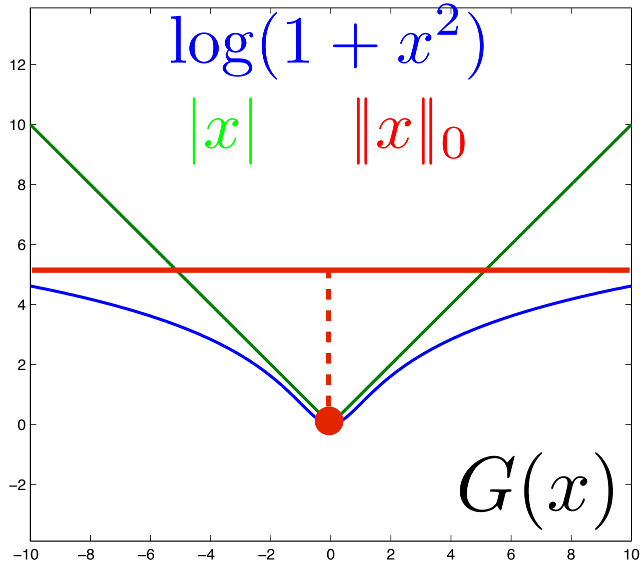
$$\mathrm{Prox}_{\gamma G}(x)_i = \max\left(0, 1 - \frac{\gamma}{|x_i|}\right) x_i$$

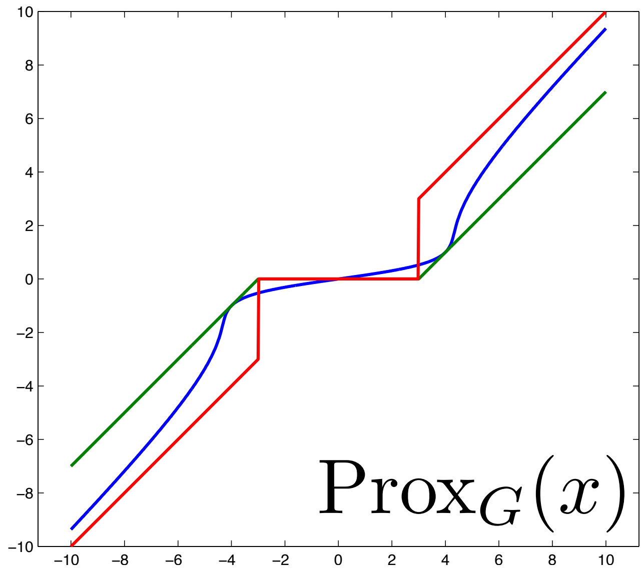
$$G(x) = \|x\|_0 = |\,\{i \setminus x_i \neq 0\}\,|$$

$$\mathrm{Prox}_{\gamma G}(x)_i = \begin{cases} x_i & \text{if} \quad |x_i| \geqslant \sqrt{2\gamma}, \\ 0 & \text{otherwise.} \end{cases}$$

$$G(x) = \sum_i \log(1 + |x_i|^2)$$

$\longrightarrow$ 3rd order polynomial root.





Proximal Calculus

Separability: $G(x) = G_1(x_1) + \ldots + G_n(x_n)$

$$\mathrm{Prox}_G(x) = (\mathrm{Prox}_{G_1}(x_1), \ldots, \mathrm{Prox}_{G_n}(x_n))$$

Proximal Calculus

Separability: $G(x) = G_1(x_1) + \ldots + G_n(x_n)$

$$\mathrm{Prox}_G(x) = (\mathrm{Prox}_{G_1}(x_1), \ldots, \mathrm{Prox}_{G_n}(x_n))$$

Quadratic functionals: $G(x) = \frac{1}{2}\|\Phi x - y\|^2$

$$\begin{aligned}\mathrm{Prox}_{\gamma G} &= (\mathrm{Id} + \gamma \Phi^* \Phi)^{-1} \Phi^* \\ &= \Phi^* (\mathrm{Id} + \gamma \Phi \Phi^*)^{-1}\end{aligned}$$

Proximal Calculus

Separability: $G(x) = G_1(x_1) + \ldots + G_n(x_n)$

$$\mathrm{Prox}_G(x) = (\mathrm{Prox}_{G_1}(x_1), \ldots, \mathrm{Prox}_{G_n}(x_n))$$

Quadratic functionals: $G(x) = \frac{1}{2}\|\Phi x - y\|^2$

$$\begin{aligned}\mathrm{Prox}_{\gamma G} &= (\mathrm{Id} + \gamma \Phi^* \Phi)^{-1} \Phi^* \\ &= \Phi^* (\mathrm{Id} + \gamma \Phi \Phi^*)^{-1}\end{aligned}$$

Composition by tight frame: $A \circ A^* = \mathrm{Id}$

$$\mathrm{Prox}_{G \circ A}(x) = A^* \circ \mathrm{Prox}_G \circ A + \mathrm{Id} - A^* \circ A$$

Proximal Calculus

Separability: $G(x) = G_1(x_1) + \ldots + G_n(x_n)$

$$\mathrm{Prox}_G(x) = (\mathrm{Prox}_{G_1}(x_1), \ldots, \mathrm{Prox}_{G_n}(x_n))$$

Quadratic functionals: $G(x) = \frac{1}{2}\|\Phi x - y\|^2$

$$\mathrm{Prox}_{\gamma G} = (\mathrm{Id} + \gamma \Phi^* \Phi)^{-1} \Phi^*$$

$$= \Phi^* (\mathrm{Id} + \gamma \Phi \Phi^*)^{-1}$$

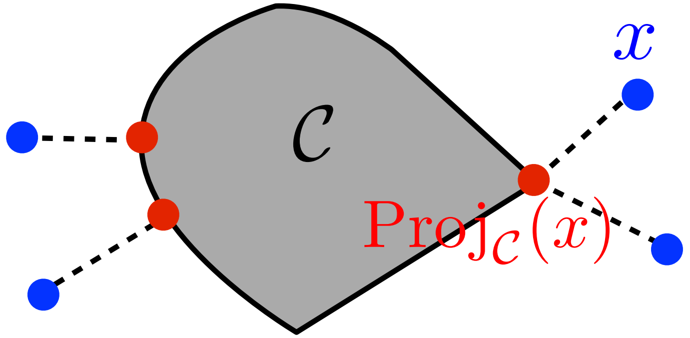
Composition by tight frame: $A \circ A^* = \mathrm{Id}$

$$\mathrm{Prox}_{G \circ A}(x) = A^* \circ \mathrm{Prox}_G \circ A + \mathrm{Id} - A^* \circ A$$

Indicators: $G(x) = \iota_{\mathcal{C}}(x)$

$$\mathrm{Prox}_{\gamma G}(x) = \mathrm{Proj}_{\mathcal{C}}(x)$$

$$= \underset{z \in \mathcal{C}}{\mathrm{argmin}}\ \|x - z\|$$



Prox and Subdifferential

Resolvant of ∂G*:*

$$z = \mathrm{Prox}_{\gamma G}(x) \quad \Longleftrightarrow \quad 0 \in z - x + \gamma\partial G(z)$$

$$\Longleftrightarrow \quad x \in (\mathrm{Id} + \gamma\partial G)(z) \quad \Longleftrightarrow \quad z = (\mathrm{Id} + \gamma\partial G)^{-1}(x)$$

Inverse of a set-valued mapping:

$$\text{where } x \in U(y) \Longleftrightarrow y \in U^{-1}(x)$$

$$\mathrm{Prox}_{\gamma G} = (\mathrm{Id} + \gamma\partial G)^{-1} \quad \text{is a single-valued mapping}$$

Prox and Subdifferential

Resolvant of ∂G:

$$z = \mathrm{Prox}_{\gamma G}(x) \quad \Longleftrightarrow \quad 0 \in z - x + \gamma \partial G(z)$$

$$\Longleftrightarrow \quad x \in (\mathrm{Id} + \gamma \partial G)(z) \quad \Longleftrightarrow \quad z = (\mathrm{Id} + \gamma \partial G)^{-1}(x)$$

Inverse of a set-valued mapping:

$$\text{where } x \in U(y) \Longleftrightarrow y \in U^{-1}(x)$$

$$\mathrm{Prox}_{\gamma G} = (\mathrm{Id} + \gamma \partial G)^{-1} \quad \text{is a single-valued mapping}$$

Fix point: $\quad x^\star \in \underset{x}{\mathrm{argmin}}\, G(x)$

$$\Longleftrightarrow \quad 0 \in \partial G(x^\star) \quad \Longleftrightarrow \quad x^\star \in (\mathrm{Id} + \gamma \partial G)(x^\star)$$

$$\Longleftrightarrow \quad x^\star = (\mathrm{Id} + \gamma \partial G)^{-1}(x^\star) = \mathrm{Prox}_{\gamma G}(x^\star)$$

Gradient and Proximal Descents

Gradient descent: $x^{(\ell+1)} = x^{(\ell)} - \gamma_\ell \nabla G(x^{(\ell)})$ **[explicit]**

G is C^1 and ∇G is L-Lipschitz

Theorem: If $0 < \gamma_\ell < 2/L$, $x^{(\ell)} \rightarrow x^\star$ a solution.

Gradient and Proximal Descents

Gradient descent: $x^{(\ell+1)} = x^{(\ell)} - \gamma_\ell \nabla G(x^{(\ell)})$ **[explicit]**

G is C^1 and ∇G is L-Lipschitz

Theorem: If $0 < \gamma_\ell < 2/L$, $x^{(\ell)} \to x^\star$ a solution.

Sub-gradient descent: $x^{(\ell+1)} = x^{(\ell)} - \gamma_\ell v^{(\ell)}, \quad v^{(\ell)} \in \partial G(x^{(\ell)})$

Theorem: If $\gamma_\ell \sim 1/\ell$, $x^{(\ell)} \to x^\star$ a solution.

$\longrightarrow$ Problem: slow.

Gradient and Proximal Descents

Gradient descent: $x^{(\ell+1)} = x^{(\ell)} - \gamma_\ell \nabla G(x^{(\ell)})$ **[explicit]**

G is C^1 and ∇G is L-Lipschitz

Theorem: If $0 < \gamma_\ell < 2/L$, $x^{(\ell)} \rightarrow x^\star$ a solution.

Sub-gradient descent: $x^{(\ell+1)} = x^{(\ell)} - \gamma_\ell v^{(\ell)}, \quad v^{(\ell)} \in \partial G(x^{(\ell)})$

Theorem: If $\gamma_\ell \sim 1/\ell$, $x^{(\ell)} \rightarrow x^\star$ a solution.

$\longrightarrow$ Problem: slow.

Proximal-point algorithm: $x^{(\ell+1)} = \mathrm{Prox}_{\gamma_\ell G}(x^{(\ell)})$ **[implicit]**

Theorem: If $\gamma_\ell \geqslant c > 0$, $x^{(\ell)} \rightarrow x^\star$ a solution.

$\longrightarrow$ $\mathrm{Prox}_{\gamma G}$ hard to compute.

Overview

- Subdifferential Calculus

- Proximal Calculus

- **Forward Backward**

- Douglas Rachford

- Generalized Forward-Backward

Proximal Splitting Methods

Solve $\min_{x \in \mathcal{H}} E(x)$

Problem: $\mathrm{Prox}_{\gamma E}$ is not available.

Proximal Splitting Methods

Solve $\min_{x \in \mathcal{H}} E(x)$

Problem: $\mathrm{Prox}_{\gamma E}$ is not available.

Splitting: $$E(x) = \underset{\text{Smooth}}{\boxed{F(x)}} + \sum_i \underset{\text{Simple}}{\boxed{G_i(x)}}$$

Proximal Splitting Methods

Solve $\displaystyle \min_{x \in \mathcal{H}} E(x)$

Problem: $\mathrm{Prox}_{\gamma E}$ is not available.

Splitting: $$E(x) = \underset{\text{Smooth}}{\boxed{F(x)}} + \sum_i \underset{\text{Simple}}{\boxed{G_i(x)}}$$

Iterative algorithms using: $\nearrow \nabla F(x)$

$\searrow \mathrm{Prox}_{\gamma G_i}(x)$

Forward-Backward: $\xrightarrow{\text{solves}}$ $F + G$

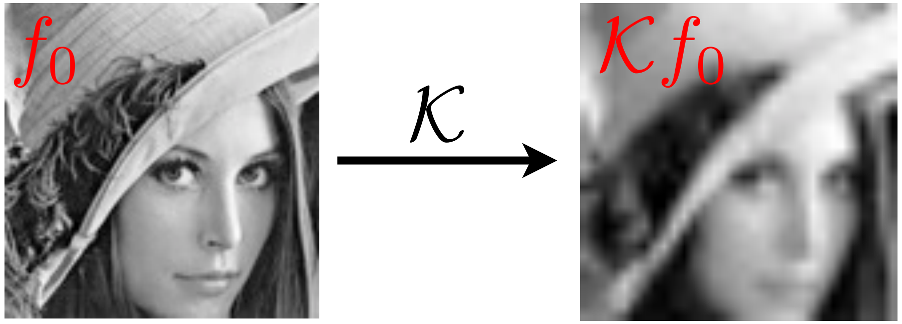
Douglas-Rachford: $\longrightarrow$ $\sum G_i$

Primal-Dual: $\longrightarrow$ $\sum G_i \circ A$

Generalized FB: $\longrightarrow$ $F + \sum G_i$

Smooth + Simple Splitting

Inverse problem: measurements $y = \mathcal{K} f_0 + w$



$$\mathcal{K} : \mathbb{R}^N \rightarrow \mathbb{R}^P, \quad P \leqslant N$$

Model: $f_0 = \Psi x_0$ sparse in dictionary Ψ.

Sparse recovery: $f^\star = \Psi x^\star$ where $x^\star$ solves

$$\min_{x \in \mathbb{R}^N} \underset{\text{Smooth}}{\boxed{F(x)}} + \underset{\text{Simple}}{\boxed{G(x)}}$$

Data fidelity: $F(x) = \frac{1}{2} \|y - \Phi x\|^2 \qquad \Phi = \mathcal{K} \circ \Psi$

Regularization: $G(x) = \|x\|_1 = \sum_i |x_i|$

Forward-Backward

Fix point equation:

$$x^\star \in \underset{x}{\operatorname{argmin}}\; F(x) + G(x) \iff 0 \in \nabla F(x^\star) + \partial G(x^\star)$$

$$\iff (x^\star - \gamma \nabla F(x^\star)) \in x^\star + \gamma \partial G(x^\star)$$

$$\iff x^\star = \operatorname{Prox}_{\gamma G}(x^\star - \gamma \nabla F(x^\star))$$

Forward-Backward

Fix point equation:

$$x^\star \in \underset{x}{\operatorname{argmin}}\, F(x) + G(x) \iff 0 \in \nabla F(x^\star) + \partial G(x^\star)$$

$$\iff (x^\star - \gamma \nabla F(x^\star)) \in x^\star + \gamma \partial G(x^\star)$$

$$\iff x^\star = \operatorname{Prox}_{\gamma G}(x^\star - \gamma \nabla F(x^\star))$$

Forward-backward: $$x^{(\ell+1)} = \operatorname{Prox}_{\gamma G}\left(x^{(\ell)} - \gamma \nabla F(x^{(\ell)}) \right)$$

Forward-Backward

Fix point equation:

$$x^\star \in \underset{x}{\operatorname{argmin}}\ F(x) + G(x) \iff 0 \in \nabla F(x^\star) + \partial G(x^\star)$$

$$\iff (x^\star - \gamma \nabla F(x^\star)) \in x^\star + \gamma \partial G(x^\star)$$

$$\iff x^\star = \operatorname{Prox}_{\gamma G}(x^\star - \gamma \nabla F(x^\star))$$

Forward-backward: $x^{(\ell+1)} = \operatorname{Prox}_{\gamma G}\left(x^{(\ell)} - \gamma \nabla F(x^{(\ell)})\right)$

Projected gradient descent: $G = \iota_{\mathcal{C}}$

Forward-Backward

Fix point equation:

$$x^\star \in \underset{x}{\text{argmin}}\; F(x) + G(x) \iff 0 \in \nabla F(x^\star) + \partial G(x^\star)$$

$$\iff (x^\star - \gamma \nabla F(x^\star)) \in x^\star + \gamma \partial G(x^\star)$$

$$\iff x^\star = \text{Prox}_{\gamma G}(x^\star - \gamma \nabla F(x^\star))$$

Forward-backward: $$x^{(\ell+1)} = \text{Prox}_{\gamma G}\left(x^{(\ell)} - \gamma \nabla F(x^{(\ell)}) \right)$$

Projected gradient descent: $G = \iota_{\mathcal{C}}$

Theorem: Let ∇F be L-Lipschitz.

If $\gamma < 2/L$, $x^{(\ell)} \to x^\star$ a solution of $(\star)$

Forward-Backward

Fix point equation:

$$x^\star \in \underset{x}{\text{argmin}}\ F(x) + G(x) \iff 0 \in \nabla F(x^\star) + \partial G(x^\star)$$

$$\iff (x^\star - \gamma \nabla F(x^\star)) \in x^\star + \gamma \partial G(x^\star)$$

$$\iff x^\star = \text{Prox}_{\gamma G}(x^\star - \gamma \nabla F(x^\star))$$

Forward-backward: $x^{(\ell+1)} = \text{Prox}_{\gamma G}\left(x^{(\ell)} - \gamma \nabla F(x^{(\ell)})\right)$

Projected gradient descent: $G = \iota_{\mathcal{C}}$

Theorem: Let ∇F be L-Lipschitz.

If $\gamma < 2/L$, $x^{(\ell)} \to x^\star$ a solution of $(\star)$

$\longrightarrow$ Multi-step accelerations (Nesterov, Beck-Teboule).

Example: L1 Regularization

$$\min_x \frac{1}{2}\|\Phi x - y\|^2 + \lambda\|x\|_1 \quad \Longleftrightarrow \quad \min_x F(x) + G(x)$$

$$F(x) = \frac{1}{2}\|\Phi x - y\|^2$$

$$\nabla F(x) = \Phi^*(\Phi x - y) \qquad L = \|\Phi^*\Phi\|$$

$$G(x) = \lambda\|x\|_1$$

$$\mathrm{Prox}_{\gamma G}(x)_i = \max\left(0, 1 - \frac{\gamma\lambda}{|x_i|}\right) x_i$$

Forward-backward $\Longleftrightarrow$ Iterative soft thresholding

Overview

- Subdifferential Calculus
- Proximal Calculus
- Forward Backward
- **Douglas Rachford**
- Generalized Forward-Backward

Douglas Rachford Scheme

$$\min_x \underset{\text{Simple}}{\boxed{G_1(x)}} + \underset{\text{Simple}}{\boxed{G_2(x)}} \qquad (\star)$$

Douglas-Rachford iterations:

$$z^{(\ell+1)} = \left(1 - \frac{\alpha}{2}\right) z^{(\ell)} + \frac{\alpha}{2} \mathrm{RProx}_{\gamma G_2} \circ \mathrm{RProx}_{\gamma G_1}(z^{(\ell)})$$

$$x^{(\ell+1)} = \mathrm{Prox}_{\gamma G_2}(z^{(\ell+1)})$$

Reflexive prox:

$$\mathrm{RProx}_{\gamma G}(x) = 2\mathrm{Prox}_{\gamma G}(x) - x$$

Douglas Rachford Scheme

$$\min_x \boxed{G_1(x)} + \boxed{G_2(x)} \qquad (\star)$$

Simple Simple

Douglas-Rachford iterations:

$$z^{(\ell+1)} = \left(1 - \frac{\alpha}{2}\right) z^{(\ell)} + \frac{\alpha}{2} \text{RProx}_{\gamma G_2} \circ \text{RProx}_{\gamma G_1}(z^{(\ell)})$$

$$x^{(\ell+1)} = \text{Prox}_{\gamma G_2}(z^{(\ell+1)})$$

Reflexive prox:

$$\text{RProx}_{\gamma G}(x) = 2\text{Prox}_{\gamma G}(x) - x$$

Theorem: If $0 < \alpha < 2$ and $\gamma > 0$,

$x^{(\ell)} \rightarrow x^\star$ a solution of $(\star)$

DR Fix Point Equation

$$\min_x G_1(x) + G_2(x) \quad \Longleftrightarrow \quad 0 \in \partial(G_1 + G_2)(x)$$

$$\Longleftrightarrow \quad \exists z,\ z - x \in \partial(\gamma G_1)(x) \quad \text{and} \quad x - z \in \partial(\gamma G_2)(x)$$

$$\Longleftrightarrow \quad x = \mathrm{Prox}_{\gamma G_1}(z) \quad \text{and} \quad \boxed{(2x - z) - x \in \partial(\gamma G_2)(x)}$$

DR Fix Point Equation

$$\min_x G_1(x) + G_2(x) \quad \Longleftrightarrow \quad 0 \in \partial(G_1 + G_2)(x)$$

$$\Longleftrightarrow \quad \exists z,\ z - x \in \partial(\gamma G_1)(x) \quad \text{and} \quad x - z \in \partial(\gamma G_2)(x)$$

$$\Longleftrightarrow \quad x = \mathrm{Prox}_{\gamma G_1}(z) \quad \text{and} \quad \boxed{(2x - z) - x \in \partial(\gamma G_2)(x)}$$

$$\boxed{\quad} \quad \Longleftrightarrow \quad x = \mathrm{Prox}_{\gamma G_2}(2x - z) = \mathrm{Prox}_{\gamma G_2} \circ \mathrm{RProx}_{\gamma G_1}(z)$$

$$\Longleftrightarrow \quad z = 2\mathrm{Prox}_{\gamma G_2} \circ \mathrm{RProx}_{\gamma G_1}(y) - (2x - z)$$

$$\Longleftrightarrow \quad z = 2\mathrm{Prox}_{\gamma G_2} \circ \mathrm{RProx}_{\gamma G_1}(z) - \mathrm{RProx}_{\gamma G_1}(z)$$

$$\Longleftrightarrow \quad z = \mathrm{RProx}_{\gamma G_2} \circ \mathrm{RProx}_{\gamma G_1}(z)$$

$$\Longleftrightarrow \quad z = \left(1 - \frac{\alpha}{2}\right) z + \frac{\alpha}{2} \mathrm{RProx}_{\gamma G_2} \circ \mathrm{RProx}_{\gamma G_1}(z)$$

Example: Constrainted L1

$$\min_{\Phi x = y} \|x\|_1 \quad \Longleftrightarrow \quad \min_x G_1(x) + G_2(x)$$

$$G_1(x) = i_{\mathcal{C}}(x), \qquad \mathcal{C} = \{x \setminus \Phi x = y\}$$

$$\text{Prox}_{\gamma G_1}(x) = \text{Proj}_{\mathcal{C}}(x) = x + \Phi^* (\Phi \Phi^*)^{-1} (y - \Phi x)$$

$$G_2(x) = \|x\|_1 \qquad \text{Prox}_{\gamma G_2}(x) = \left(\max\left(0, 1 - \frac{\gamma}{|x_i|}\right) x_i \right)_i$$

$\longrightarrow$ efficient if $\Phi\Phi^*$ easy to invert.

Example: Constrainted L1

$$\min_{\Phi x = y} \|x\|_1 \quad \Longleftrightarrow \quad \min_x G_1(x) + G_2(x)$$

$$G_1(x) = i_{\mathcal{C}}(x), \qquad \mathcal{C} = \{x \setminus \Phi x = y\}$$

$$\operatorname{Prox}_{\gamma G_1}(x) = \operatorname{Proj}_{\mathcal{C}}(x) = x + \Phi^*(\Phi\Phi^*)^{-1}(y - \Phi x)$$

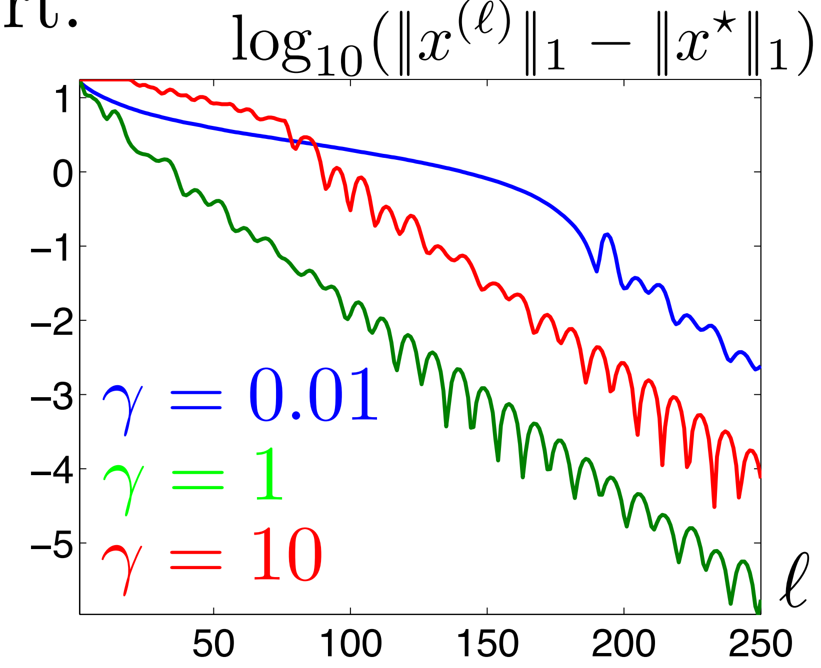
$$G_2(x) = \|x\|_1 \qquad \operatorname{Prox}_{\gamma G_2}(x) = \left(\max\left(0, 1 - \frac{\gamma}{|x_i|}\right) x_i \right)_i$$

$\longrightarrow$ efficient if $\Phi\Phi^*$ easy to invert.

Example: compressed sensing

$\Phi \in \mathbb{R}^{100 \times 400}$ Gaussian matrix

$y = \Phi x_0$ $\quad \|x_0\|_0 = 17$



More than 2 Functionals

$$\min_x G_1(x) + \ldots + G_k(x)$$

each F_i is simple

$$\Longleftrightarrow \quad \min_x G(x_1, \ldots, x_k) + \iota_{\mathcal{C}}(x_1, \ldots, x_k)$$

$$G(x_1, \ldots, x_k) = G_1(x_1) + \ldots + G_k(x_k)$$

$$\mathcal{C} = \big\{ (x_1, \ldots, x_k) \in \mathcal{H}^k \backslash x_1 = \ldots = x_k \big\}$$

More than 2 Functionals

$$\min_x G_1(x) + \ldots + G_k(x) \qquad \text{each } F_i \text{ is simple}$$

$$\Longleftrightarrow \quad \min_x G(x_1, \ldots, x_k) + \iota_{\mathcal{C}}(x_1, \ldots, x_k)$$

$$G(x_1, \ldots, x_k) = G_1(x_1) + \ldots + G_k(x_k)$$

$$\mathcal{C} = \left\{ (x_1, \ldots, x_k) \in \mathcal{H}^k \setminus x_1 = \ldots = x_k \right\}$$

G and $\iota_{\mathcal{C}}$ are simple:

$$\mathrm{Prox}_{\gamma G}(x_1, \ldots, x_k) = \left(\mathrm{Prox}_{\gamma G_i}(x_i) \right)_i$$

$$\mathrm{Prox}_{\gamma \iota_{\mathcal{C}}}(x_1, \ldots, x_k) = (\tilde{x}, \ldots, \tilde{x}) \qquad \text{where} \quad \tilde{x} = \frac{1}{k} \sum_i x_i$$

Overview

- Subdifferential Calculus
- Proximal Calculus
- Forward Backward
- Douglas Rachford
- **Generalized Forward-Backward**

GFB Splitting

$$\min_{x\in\mathbb{R}^N} \boxed{F(x)} + \boxed{\sum_{i=1}^{n} G_i(x)} \qquad (\star)$$

Smooth Simple

$\forall\, i = 1, \ldots, n,$

$$z_i^{(\ell+1)} = z_i^{(\ell)} + \mathrm{Prox}_{n\gamma G_\ell}(2x^{(\ell)} - z_i^{(\ell)} - \gamma\nabla F(x^{(\ell)})) - x^{(\ell)}$$

$$x^{(\ell+1)} = \frac{1}{n}\sum_{i=1}^{n} z_i^{(\ell+1)}$$

GFB Splitting

$$\min_{x \in \mathbb{R}^N} \boxed{F(x)} + \boxed{\sum_{i=1}^{n} G_i(x)} \qquad (\star)$$

Smooth Simple

$\forall\, i = 1, \ldots, n,$

$$z_i^{(\ell+1)} = z_i^{(\ell)} + \text{Prox}_{n\gamma G_\ell}(2x^{(\ell)} - z_i^{(\ell)} - \gamma \nabla F(x^{(\ell)})) - x^{(\ell)}$$

$$x^{(\ell+1)} = \frac{1}{n} \sum_{i=1}^{n} z_i^{(\ell+1)}$$

Theorem: Let ∇F be L-Lipschitz.

If $\gamma < 2/L$, $x^{(\ell)} \rightarrow x^\star$ a solution of $(\star)$

GFB Splitting

$$\min_{x\in\mathbb{R}^N} \boxed{F(x)} + \boxed{\sum_{i=1}^{n} G_i(x)} \qquad (\star)$$

Smooth Simple

$\forall\, i = 1, \ldots, n,$

$$z_i^{(\ell+1)} = z_i^{(\ell)} + \mathrm{Prox}_{n\gamma G_\ell}(2x^{(\ell)} - z_i^{(\ell)} - \gamma \nabla F(x^{(\ell)})) - x^{(\ell)}$$

$$x^{(\ell+1)} = \frac{1}{n}\sum_{i=1}^{n} z_i^{(\ell+1)}$$

Theorem: Let ∇F be L-Lipschitz.

If $\gamma < 2/L$, $x^{(\ell)} \to x^\star$ a solution of $(\star)$

$n = 1 \longrightarrow$ Forward-backward.

$F = 0 \longrightarrow$ Douglas-Rachford.

GFB Fix Point

$$x \in \underset{x \in \mathbb{R}^N}{\text{argmin}}\; F(x) + \textstyle\sum_i G_i(x) \iff 0 \in \nabla F(x^\star) + \textstyle\sum_i \partial G_i(x^\star)$$

$$\iff \quad \exists\, y_i \in \partial G_i(x^\star),\; \nabla F(x^\star) + \textstyle\sum_i y_i = 0$$

GFB Fix Point

$$x \in \underset{x\in\mathbb{R}^N}{\text{argmin}}\ F(x) + \textstyle\sum_i G_i(x) \iff 0 \in \nabla F(x^\star) + \textstyle\sum_i \partial G_i(x^\star)$$

$$\iff \quad \exists\, y_i \in \partial G_i(x^\star),\ \nabla F(x^\star) + \textstyle\sum_i y_i = 0$$

$$\iff \quad \boxed{\exists (z_i)_{i=1}^n,\ \forall\, i,\ \frac{1}{n}\big(x^\star - z_i - \gamma\nabla F(x^\star)\big) \in \gamma\partial G_i(x^\star)}$$

$$\boxed{x^\star = \tfrac{1}{n}\textstyle\sum_i z_i} \qquad (\text{use } z_i = x^\star - \gamma\nabla F(x^\star) - n\gamma y_i)$$

GFB Fix Point

$$x \in \underset{x\in\mathbb{R}^N}{\mathrm{argmin}}\ F(x) + \sum_i G_i(x) \iff 0 \in \nabla F(x^\star) + \sum_i \partial G_i(x^\star)$$

$$\iff \quad \exists\, y_i \in \partial G_i(x^\star),\ \nabla F(x^\star) + \sum_i y_i = 0$$

$$\iff \quad \boxed{\exists (z_i)_{i=1}^n,\ \forall\, i,\ \frac{1}{n}\big(x^\star - z_i - \gamma\nabla F(x^\star)\big) \in \gamma\partial G_i(x^\star)}$$

$$\boxed{x^\star = \tfrac{1}{n}\sum_i z_i} \qquad (\text{use } z_i = x^\star - \gamma\nabla F(x^\star) - n\gamma y_i)$$

🟦 $$\iff \quad (2x^\star - z_i - \gamma\nabla F(x^\star)) - x^\star \in n\gamma\partial G_i(x^\star)$$

$$\iff \quad x^\star = \mathrm{Prox}_{n\gamma G_i}(2x^\star - z_i - \gamma\nabla F(x^\star))$$

$$\iff \quad z_i = z_i + \mathrm{Prox}_{n\gamma G_\ell}(2x^\star - z_i - \gamma\nabla F(x^\star)) - x^\star$$

GFB Fix Point

$$x \in \underset{x\in\mathbb{R}^N}{\text{argmin}}\ F(x) + \textstyle\sum_i G_i(x) \iff 0 \in \nabla F(x^\star) + \textstyle\sum_i \partial G_i(x^\star)$$

$$\iff \quad \exists\, y_i \in \partial G_i(x^\star),\ \nabla F(x^\star) + \textstyle\sum_i y_i = 0$$

$$\iff \quad \boxed{\exists (z_i)_{i=1}^n,\ \forall\, i,\ \frac{1}{n}\big(x^\star - z_i - \gamma \nabla F(x^\star)\big) \in \gamma \partial G_i(x^\star)}$$

$$\boxed{x^\star = \tfrac{1}{n} \textstyle\sum_i z_i} \qquad (\text{use } z_i = x^\star - \gamma \nabla F(x^\star) - n\gamma y_i)$$

$$🟦 \quad \iff \quad (2x^\star - z_i - \gamma \nabla F(x^\star)) - x^\star \in n\gamma \partial G_i(x^\star)$$

$$\iff \quad x^\star = \text{Prox}_{n\gamma G_i}(2x^\star - z_i - \gamma \nabla F(x^\star))$$

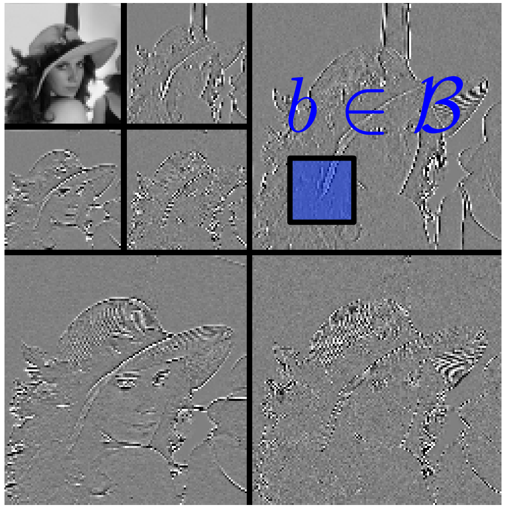
$$\iff \quad z_i = z_i + \text{Prox}_{n\gamma G_\ell}(2x^\star - z_i - \gamma \nabla F(x^\star)) - x^\star$$

🟦 + 🟥 $\longrightarrow$ Fix point equation on $(x^\star, z_1, \ldots, z_n)$.

Block Regularization

$\ell^1 - \ell^2$ block sparsity: $G(x) = \sum_{b \in \mathcal{B}} \|x^{[b]}\|, \qquad \|x^{[b]}\|^2 = \sum_{m \in b} x_m^2$

Image $f = \Psi x$



Coefficients x.

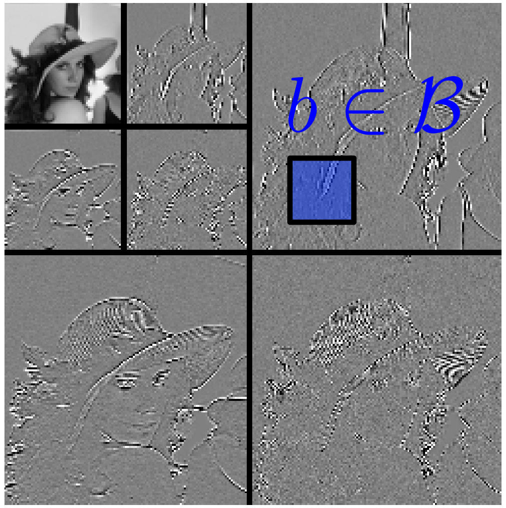
Block Regularization

$\ell^1 - \ell^2$ block sparsity: $G(x) = \sum_{b \in \mathcal{B}} \|x^{[b]}\|, \qquad \|x^{[b]}\|^2 = \sum_{m \in b} x_m^2$

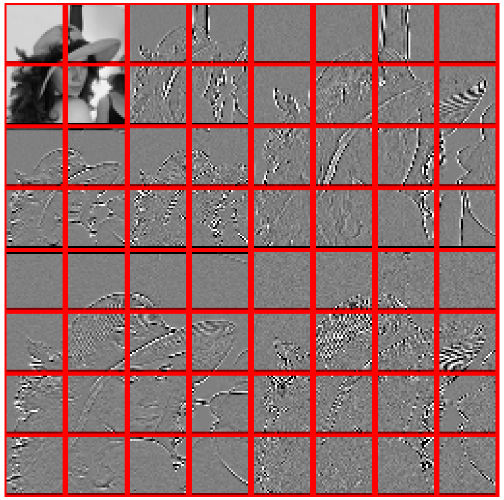
Non-overlapping decomposition: $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_n$

$$G(x) = \sum_{i=1}^{n} G_i(x) \qquad G_i(x) = \sum_{b \in \mathcal{B}_i} \|x^{[b]}\|,$$

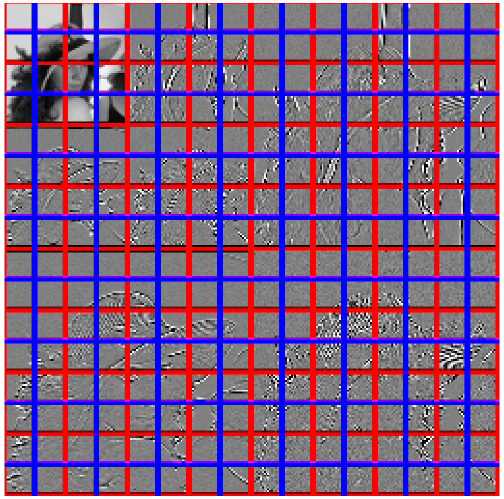
Image $f = \Psi x$



Coefficients x.



Blocks $\mathcal{B}_1$



$\mathcal{B}_1 \cup \mathcal{B}_2$

Block Regularization

$\ell^1 - \ell^2$ block sparsity: $G(x) = \sum_{b \in \mathcal{B}} \|x^{[b]}\|, \qquad \|x^{[b]}\|^2 = \sum_{m \in b} x_m^2$

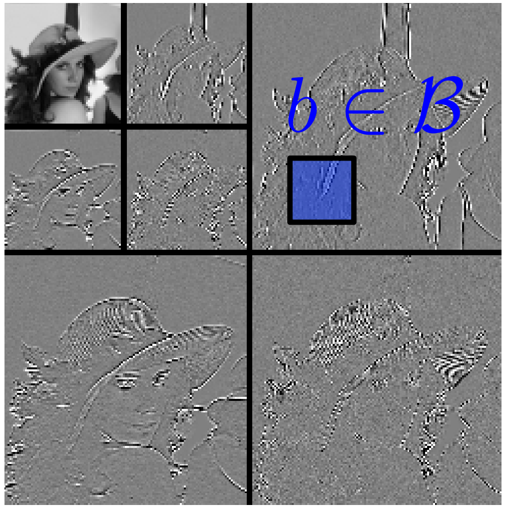
Non-overlapping decomposition: $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_n$

$$G(x) = \sum_{i=1}^{n} G_i(x) \qquad G_i(x) = \sum_{b \in \mathcal{B}_i} \|x^{[b]}\|,$$

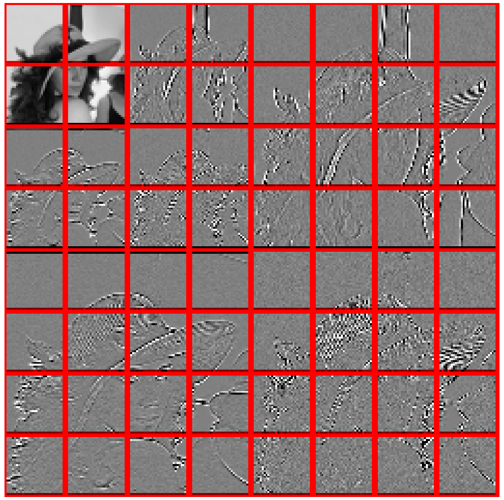
Each G_i is simple:

$$\forall\, m \in b \in \mathcal{B}_i, \quad \mathrm{Prox}_{\gamma G_i}(x)_m = \max\left(0, 1 - \frac{\gamma}{\|x^{[b]}\|}\right) x_m$$

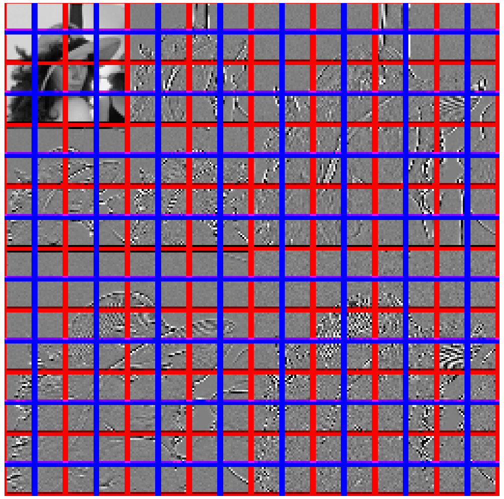
Image $f = \Psi x$



Coefficients x.



Blocks $\mathcal{B}_1$



$\mathcal{B}_1 \cup \mathcal{B}_2$

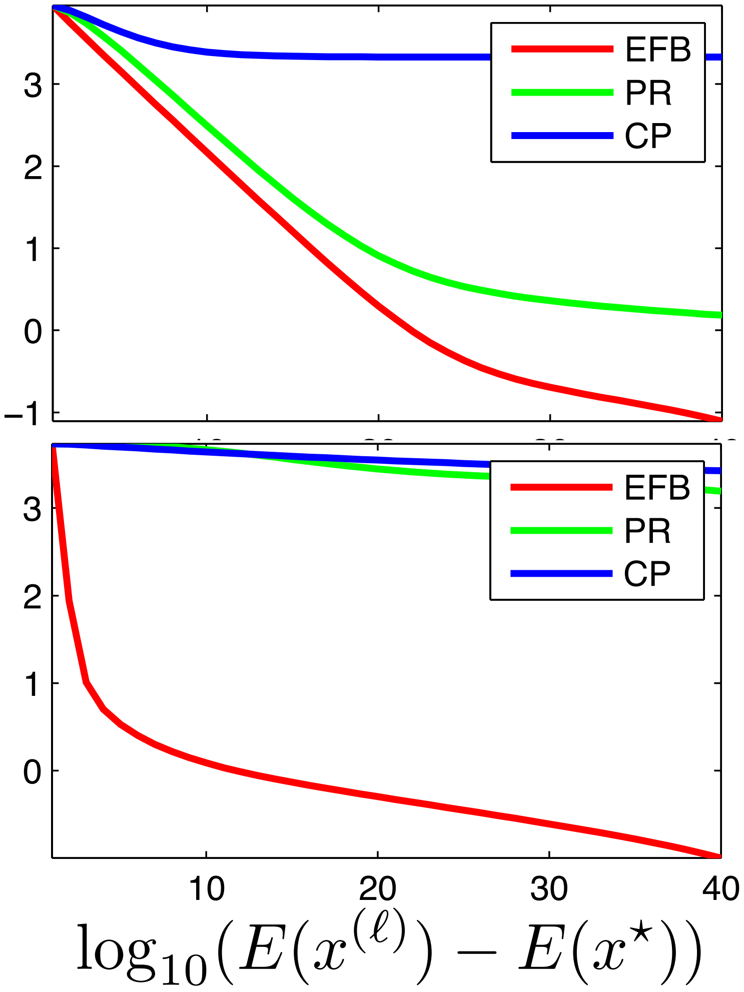
Numerical Illustration

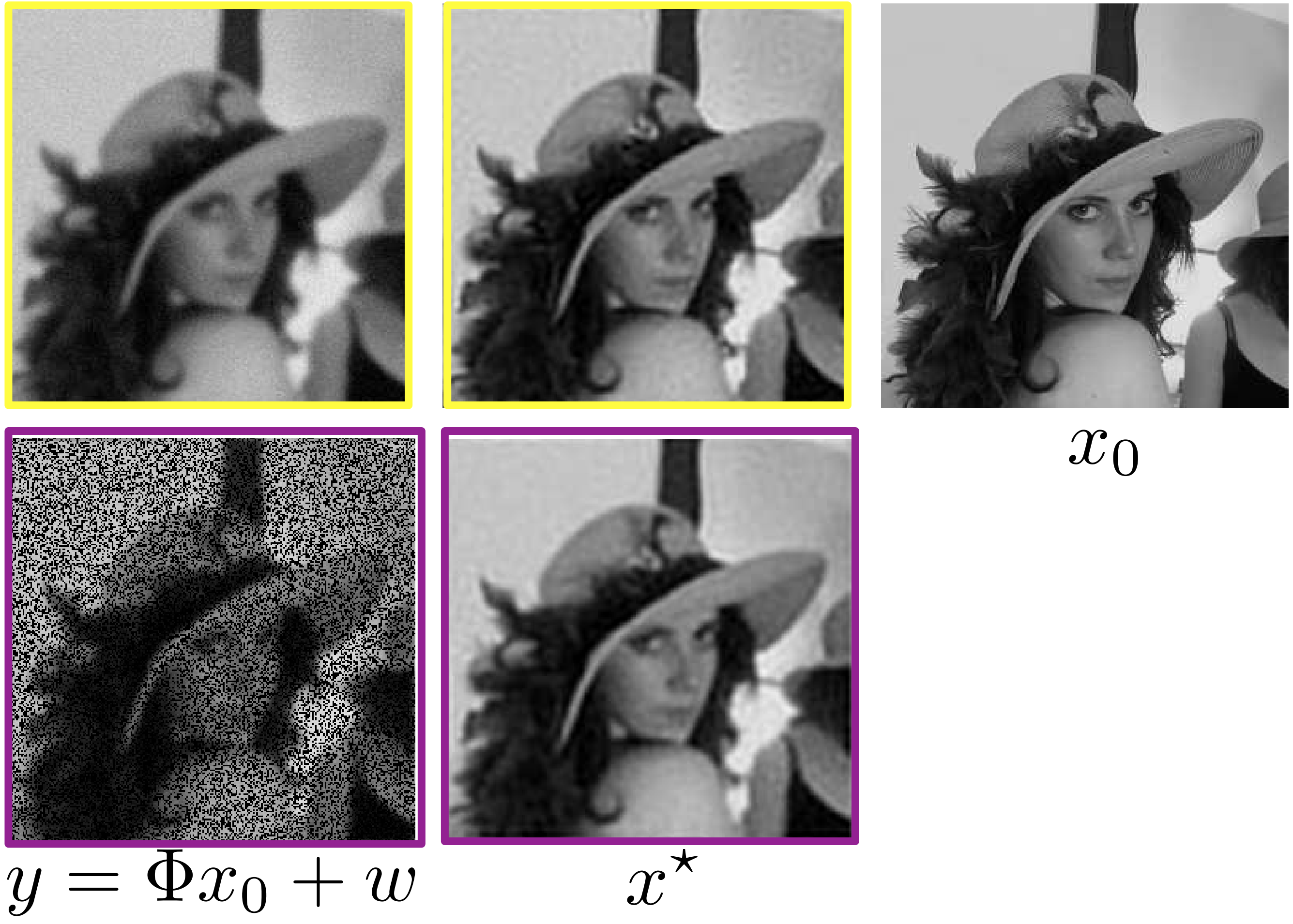
$$\min_x \frac{1}{2}\|y - \Phi\Psi x\|^2 + \lambda \sum_i G_i(x)$$

$\Psi = $ TI wavelets

$\Phi = $ convolution

$\Phi = $ inpainting+convolution





Conclusion

Inverse problems in imaging:

→ Large scale, $N \geqslant 10^6$.

→ Non-smooth (sparsity, TV, . . .)

→ (Sometimes) convex.

→ Highly structured (separability, ℓ^p norms, . . .).

Conclusion

Inverse problems in imaging:

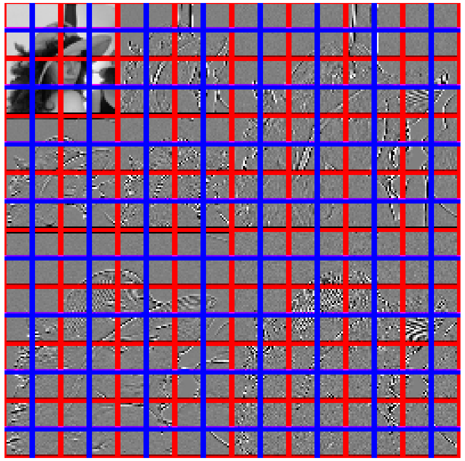
→ Large scale, $N \geqslant 10^6$.

→ Non-smooth (sparsity, TV, . . .)

→ (Sometimes) convex.

→ Highly structured (separability, ℓ^p norms, . . .).

Proximal splitting:



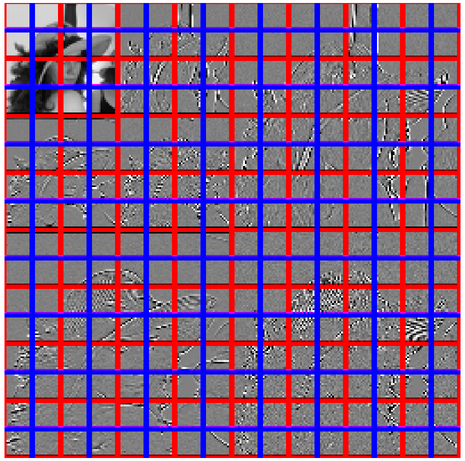
→ Unravel the structure of problems.

→ Parallelizable.

Conclusion

Inverse problems in imaging:

- → Large scale, $N \geqslant 10^6$.
- → Non-smooth (sparsity, TV, …)
- → (Sometimes) convex.
- → Highly structured (separability, ℓ^p norms, …).



Proximal splitting:

- → Unravel the structure of problems.
- → Parallelizable.

Open problems:

- → Less structured problems without smoothness.
- → Non-convex optimization.