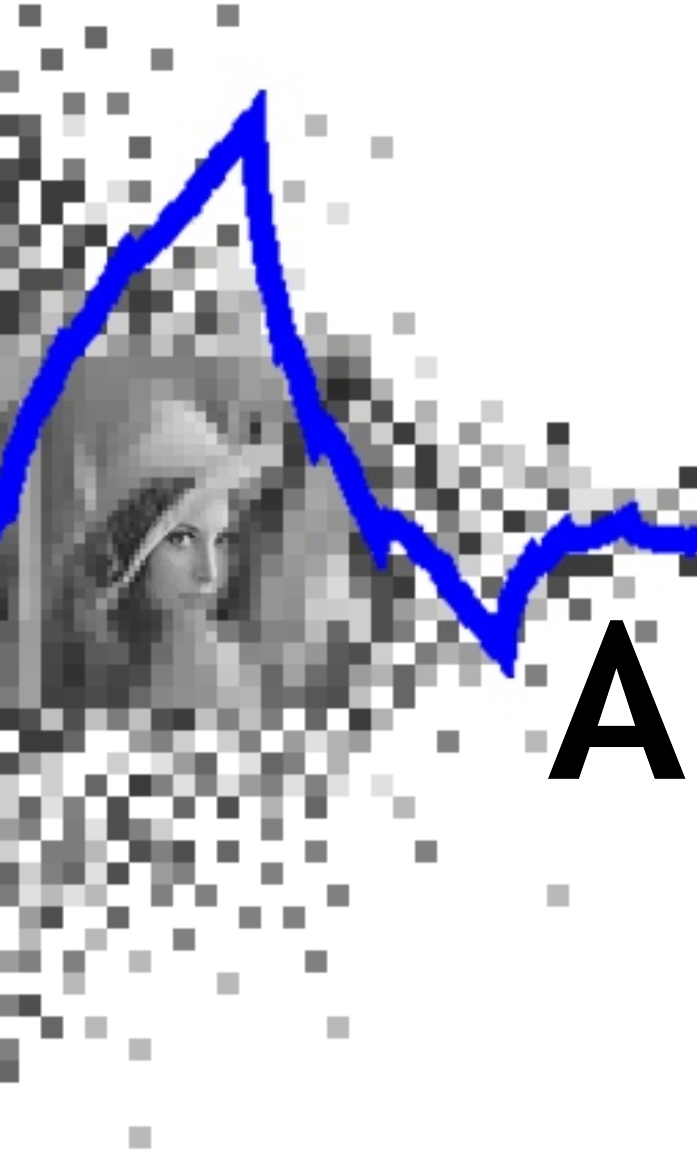


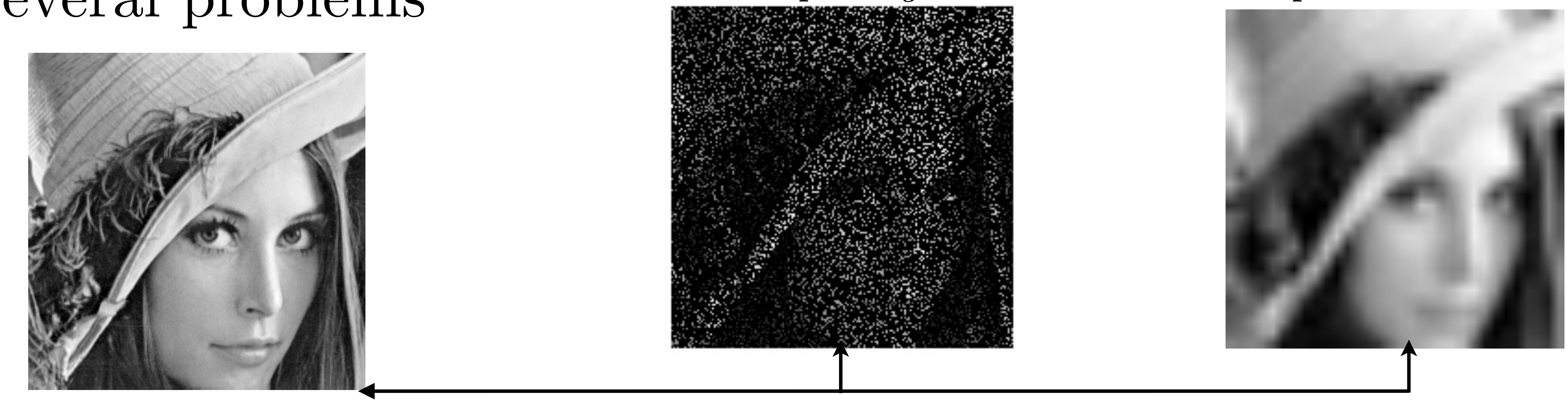
Robust Sparse Analysis Recovery

Samuel Vaiter

Inverse Problems

Inverse Problems

Several problems



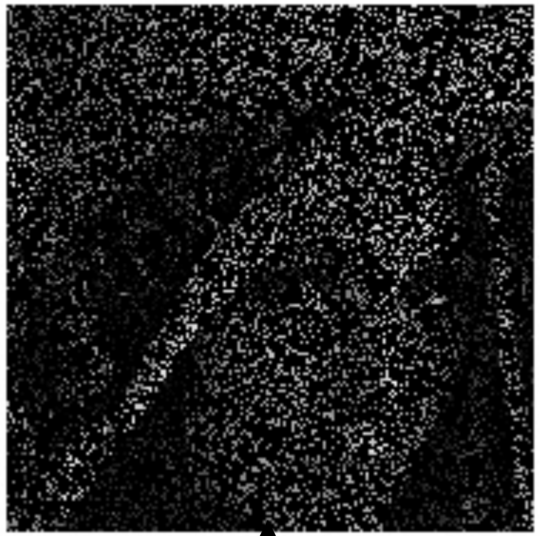
Inpaiting

Super-resolution

Inverse Problems

Several problems

Inpaiting



Super-resolution

One model

Observations ⟶ $y = \Phi x_0 + w$ ⟵ Noise

Operator ↗ ↖ Unknown signal

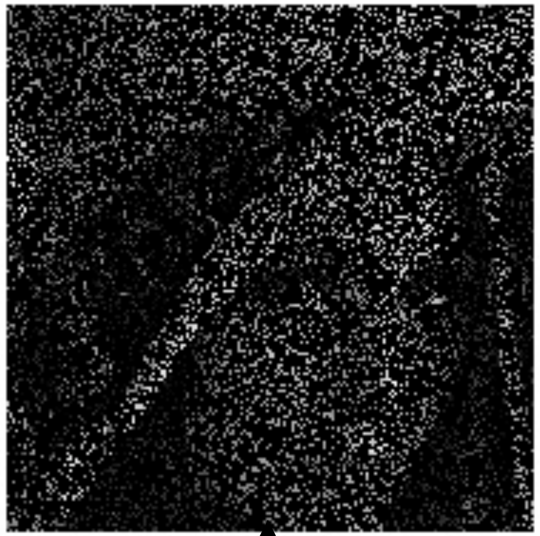
Linear hypothesis

Φ ill-posed

Inverse Problems

Several problems

Inpaiting



Super-resolution

One model

$$\text{Observations} \longrightarrow y = \Phi x_0 + w \longleftarrow \text{Noise}$$

Operator ↗ ↖ Unknown signal

Linear hypothesis

Φ ill-posed

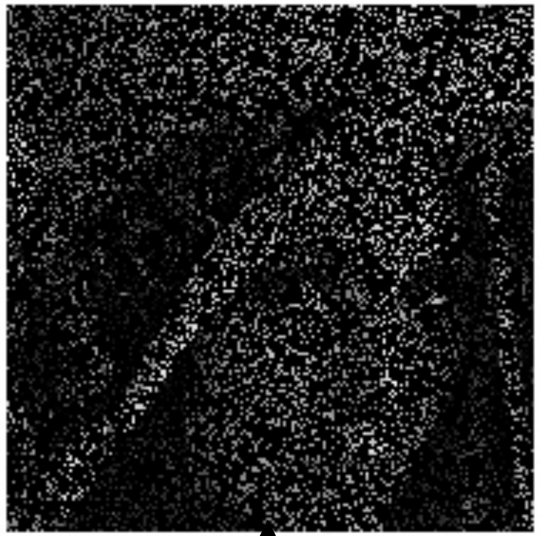
Regularization

$$x^\star \in \underset{x\in\mathbb{R}^N}{\operatorname{argmin}} \underbrace{\boxed{\frac{1}{2}\|y-\Phi x\|_2^2}}_{\text{Data fitting}} + \underbrace{\boxed{\lambda J(x)}}_{\text{Regularity}}$$

Inverse Problems

Several problems

Inpaiting



Super-resolution

One model

$$\text{Observations} \longrightarrow y = \Phi x_0 + w \longleftarrow \text{Noise}$$

Operator   Unknown signal

Linear hypothesis Φ ill-posed

Regularization

$$x^\star \in \underset{x \in \mathbb{R}^N}{\text{argmin}} \underbrace{\boxed{\frac{1}{2}\|y - \Phi x\|_2^2}}_{\text{Data fitting}} + \underbrace{\boxed{\lambda J(x)}}_{\text{Regularity}} \xrightarrow[\lambda \to 0]{\text{Noiseless}} x^\star \in \underset{\Phi x = y}{\text{argmin}}\, J(x)$$

Image Priors

Sobolev

$$J(x) = \frac{1}{2}\int \|\nabla x\|^2$$

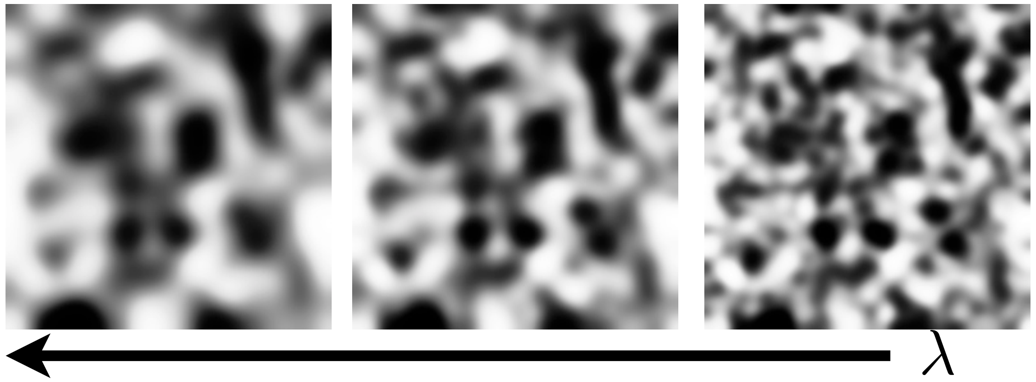
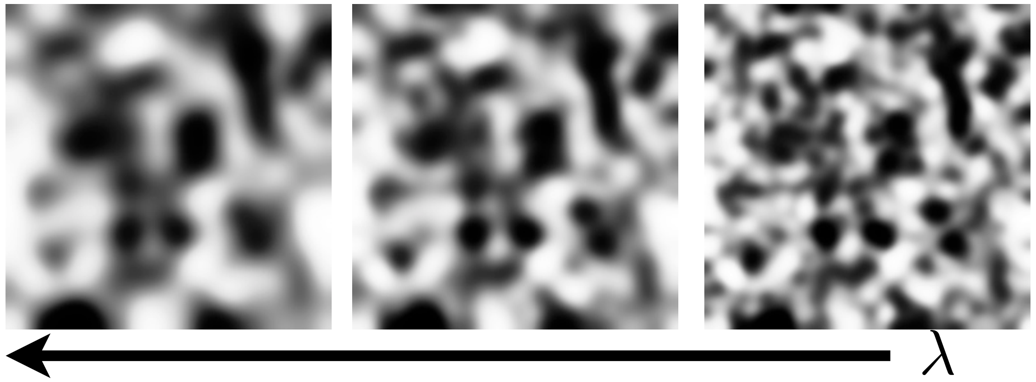


Image Priors

Sobolev

$$J(x) = \frac{1}{2}\int \|\nabla x\|^2$$



Total variation

$$J(x) = \int \|\nabla x\|$$

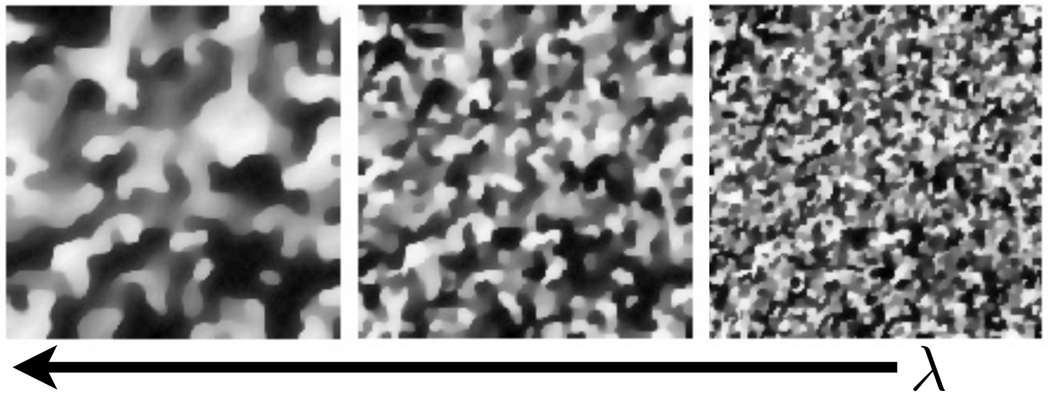
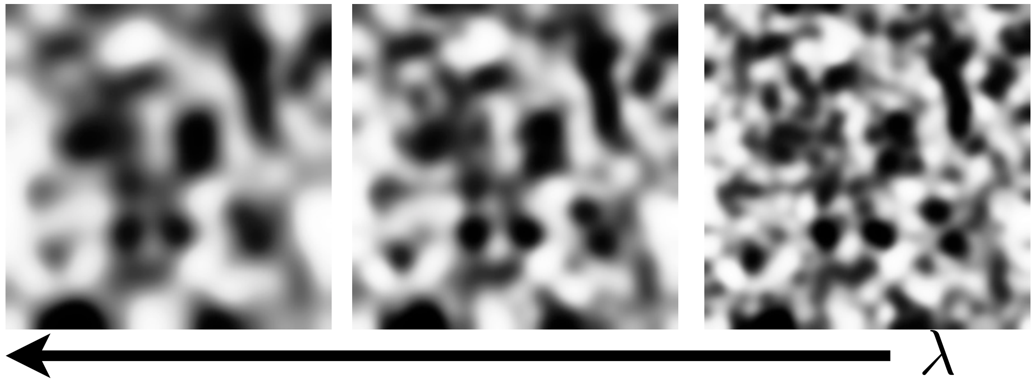


Image Priors

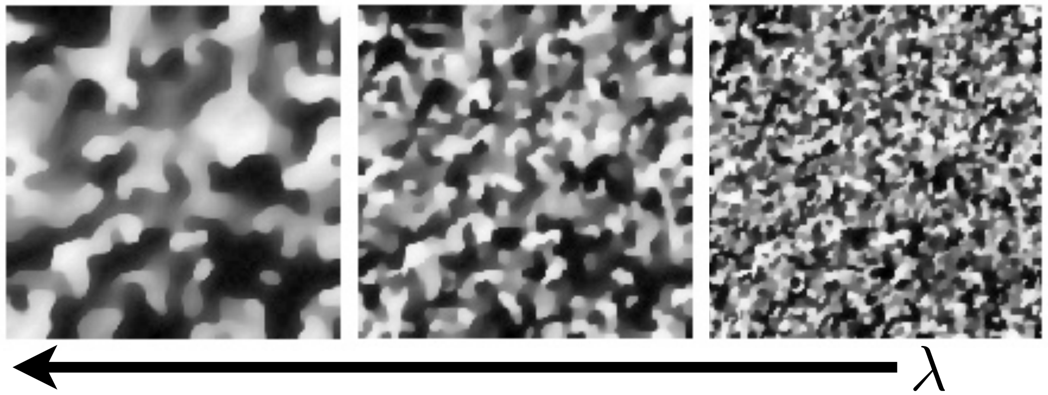
Sobolev

$$J(x) = \frac{1}{2}\int \|\nabla x\|^2$$



Total variation

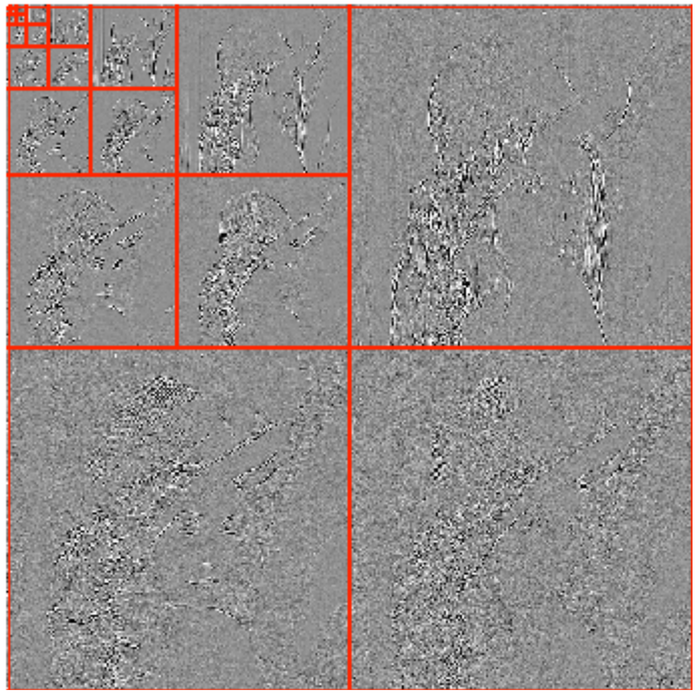
$$J(x) = \int \|\nabla x\|$$



Wavelet sparsity

$$J(x) = |\,\{i \setminus \langle x,\, \psi_i \rangle \neq 0\}\,|$$

(ideal prior)



Overview

- Analysis vs. Synthesis Regularization
- Local Parameterization of Analysis Regularization
- Identifiability and Stability
- Numerical Evaluation
- Perspectives

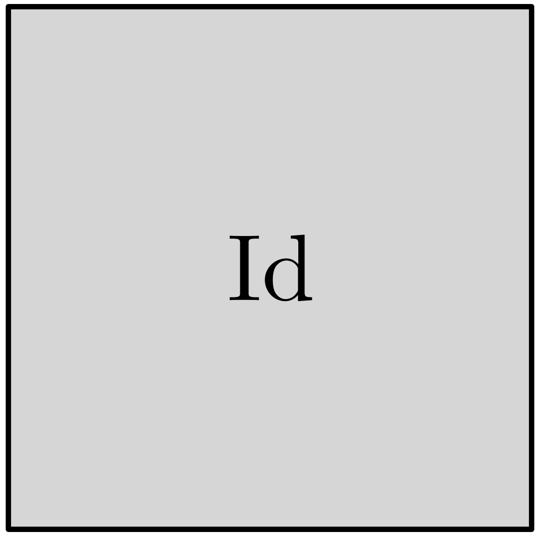
Dictionary

Redundant dictionary of $\mathbb{R}^N : \{d_i\}_{i=0}^{P-1}, P \geqslant N$

Dictionary

Redundant dictionary of $\mathbb{R}^N$: $\{d_i\}_{i=0}^{P-1}, P \geqslant N$

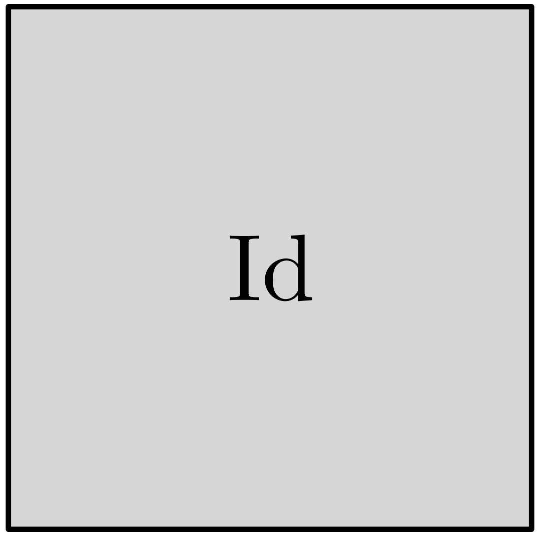
Identity



Dictionary

Redundant dictionary of $\mathbb{R}^N$: $\{d_i\}_{i=0}^{P-1}, P \geqslant N$

Identity

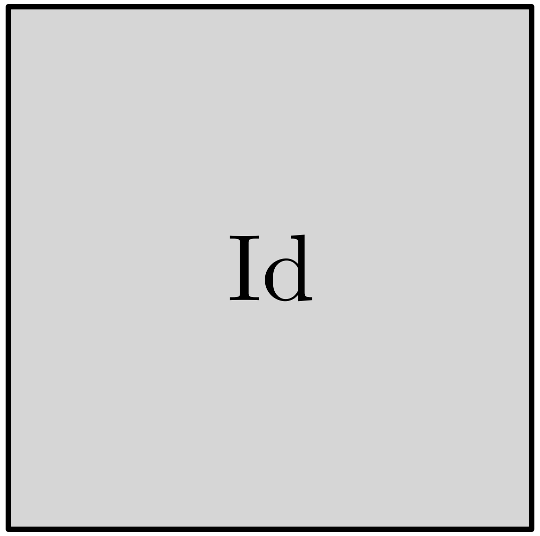


shift invariant wavelet frame

Dictionary

Redundant dictionary of $\mathbb{R}^N$: $\{d_i\}_{i=0}^{P-1}, P \geqslant N$

Identity



shift invariant wavelet frame

finite difference operator

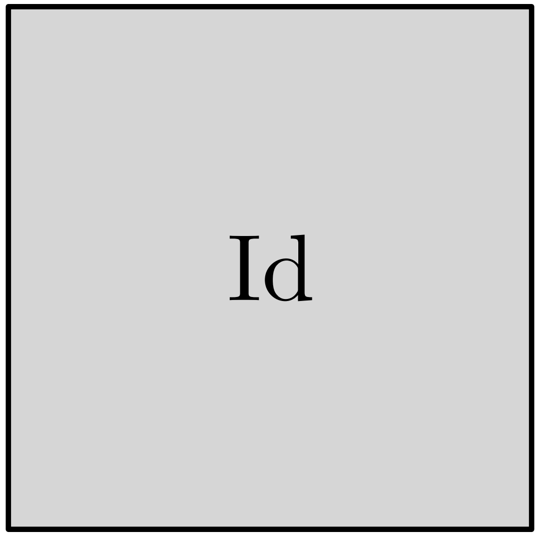
$$\Omega_{\mathrm{DIF}}^{n}$$

$$\begin{pmatrix} -1 & & & 0 \\ +1 & -1 & & \\ & +1 & \ddots & \\ & & \ddots & -1 \\ 0 & & & +1 \end{pmatrix}$$

Dictionary

Redundant dictionary of $\mathbb{R}^N$: $\{d_i\}_{i=0}^{P-1}, P \geqslant N$

Identity



Id

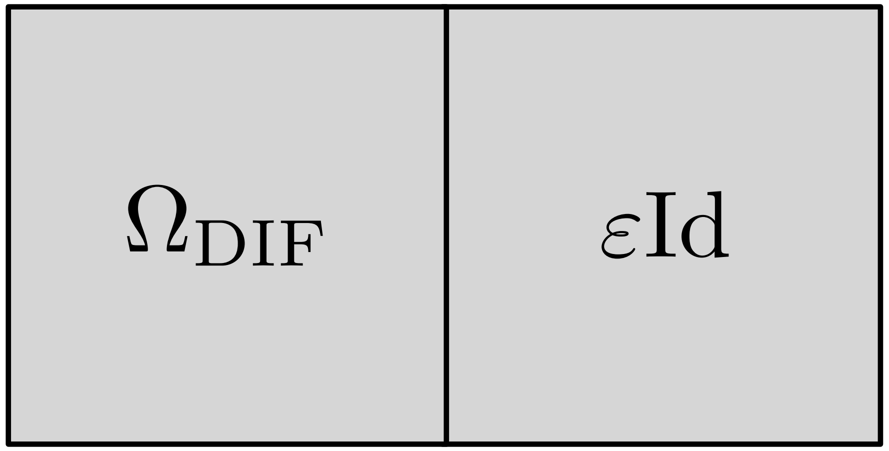
shift invariant wavelet frame

finite difference operator

$$\Omega_{\text{DIF}}^n$$

$$\begin{pmatrix} -1 & & & 0 \\ +1 & -1 & & \\ & +1 & \ddots & \\ & & \ddots & -1 \\ 0 & & & +1 \end{pmatrix}$$

fussed lasso



Ω_{DIF}

εId

Analysis versus Synthesis

Two point of view

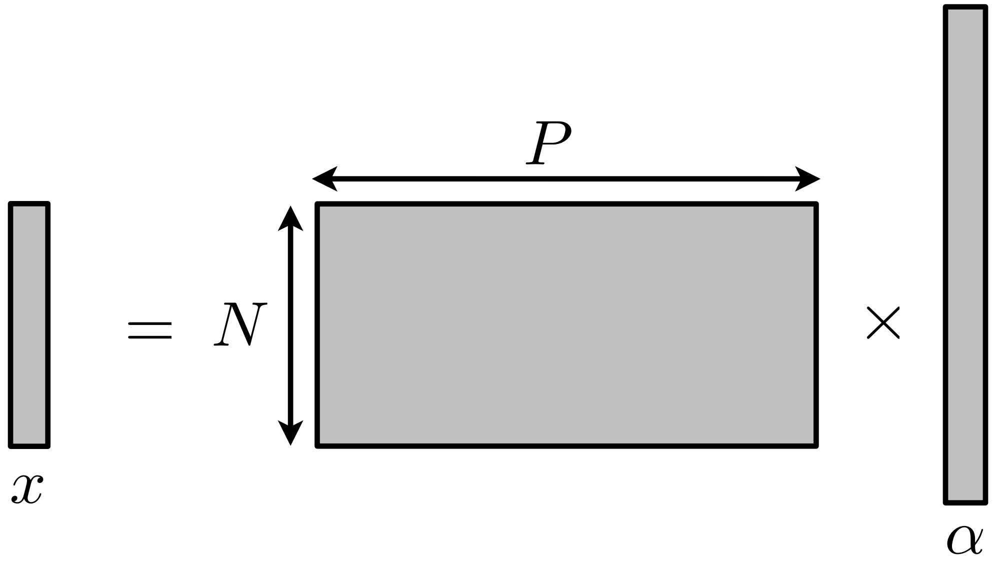
Analysis versus Synthesis

Two point of view

$$x = D\alpha$$

Synthesis

$\longrightarrow$ non-unique if $P > N$



"Generate" x

Analysis versus Synthesis

Two point of view

$x = D\alpha$

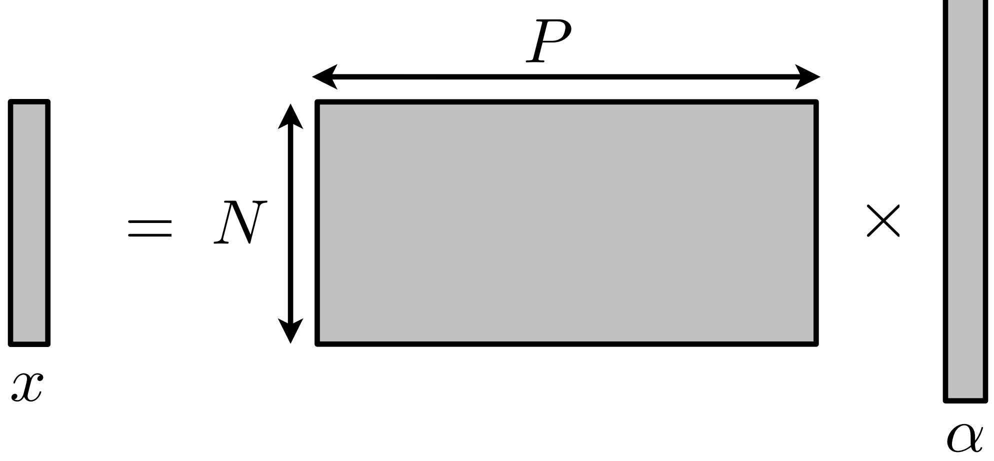
Synthesis

OR ?

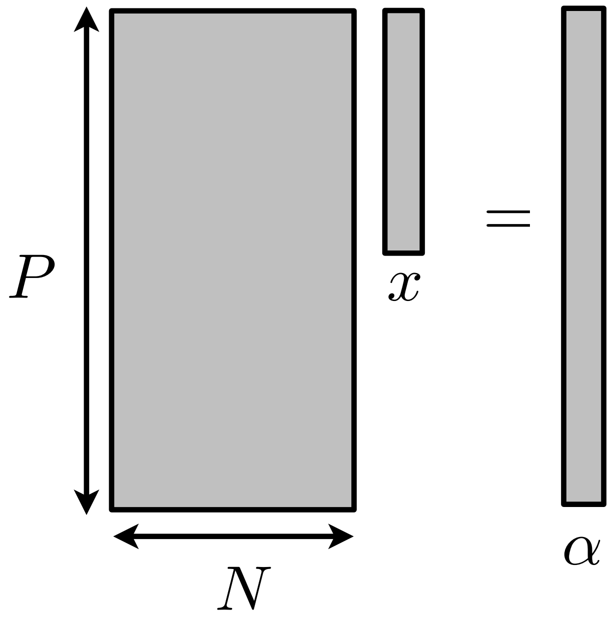
$D^* x = \alpha$

Analysis

$\longrightarrow$ non-unique if $P > N$



“Generate” x



“Analyze” x

A Bird's Eye View of Sparsity

"Ideal" sparsity prior: $J_0(\alpha) = |\,\{i \setminus \alpha_i \neq 0\}\,|$

A Bird's Eye View of Sparsity

"Ideal" sparsity prior: $J_0(\alpha) = |\,\{i \setminus \alpha_i \neq 0\}\,|$

ℓ^0 minimization is NP-hard

A Bird's Eye View of Sparsity

"Ideal" sparsity prior: $J_0(\alpha) = |\{i \setminus \alpha_i \neq 0\}|$

ℓ^0 minimization is NP-hard

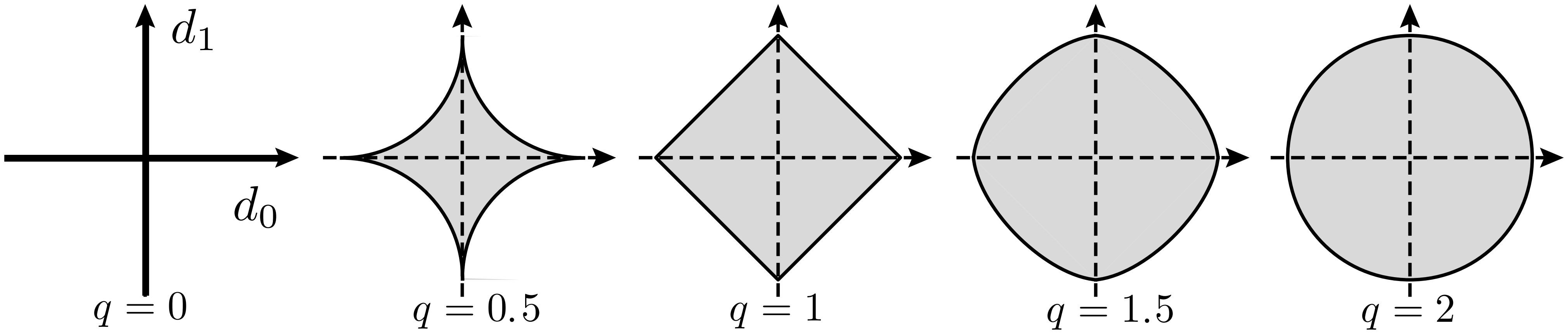
ℓ^q prior: $J_q(\alpha) = \sum_i |\alpha_i|^q$ convex (norms) for $q \geqslant 1$

A Bird's Eye View of Sparsity

"Ideal" sparsity prior: $J_0(\alpha) = |\{i \setminus \alpha_i \neq 0\}|$

ℓ^0 minimization is NP-hard

ℓ^q prior: $J_q(\alpha) = \sum_i |\alpha_i|^q$ convex (norms) for $q \geqslant 1$

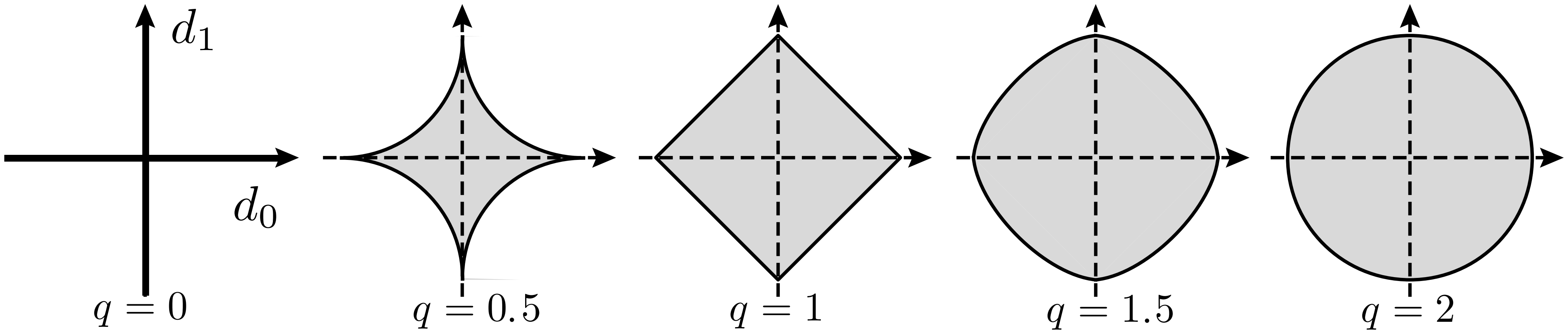


A Bird's Eye View of Sparsity

"Ideal" sparsity prior: $J_0(\alpha) = |\,\{i \setminus \alpha_i \neq 0\}\,|$

ℓ^0 minimization is NP-hard

ℓ^q prior: $J_q(\alpha) = \sum_i |\alpha_i|^q$ convex (norms) for $q \geqslant 1$



ℓ^1 norm: *convexification* of ℓ^0 prior

Sparse Regularizations

Synthesis

$$\underset{\alpha \in \mathbb{R}^Q}{\operatorname{argmin}} \ \frac{1}{2}\|y - \Psi\alpha\|_2^2 + \lambda\|\alpha\|_1$$

$$\Psi = \Phi D \qquad x = D\alpha$$

Sparse Regularizations

Synthesis

$$\underset{\alpha \in \mathbb{R}^Q}{\operatorname{argmin}} \frac{1}{2}\|y - \Psi\alpha\|_2^2 + \lambda\|\alpha\|_1$$

$$\Psi = \Phi D \qquad x = D\alpha$$

Analysis

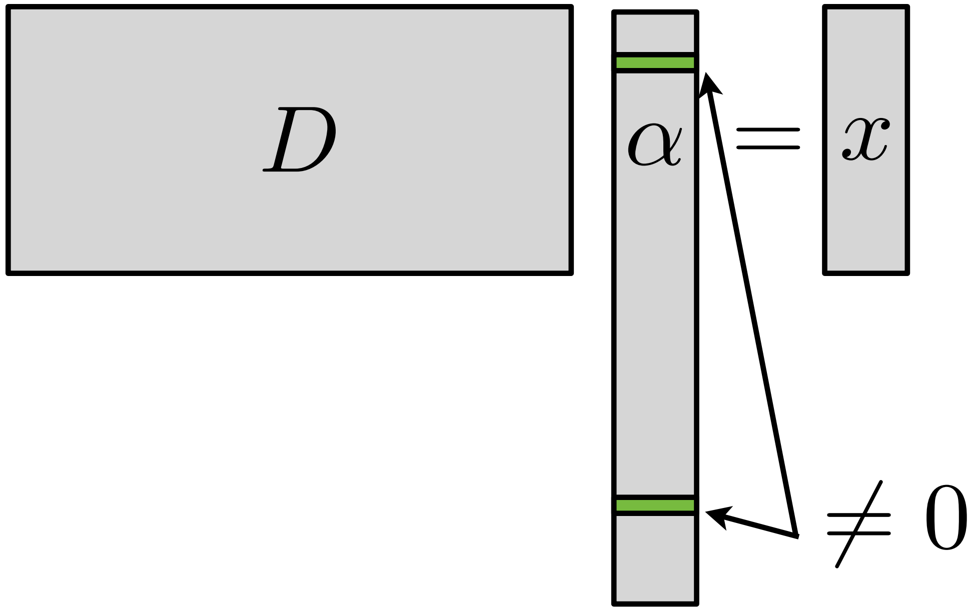
$$\underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda\|D^* x\|_1$$

Sparse Regularizations

Synthesis

$$\underset{\alpha \in \mathbb{R}^Q}{\operatorname{argmin}} \frac{1}{2}\|y - \Psi\alpha\|_2^2 + \lambda\|\alpha\|_1$$

$$\Psi = \Phi D \qquad x = D\alpha$$



Analysis

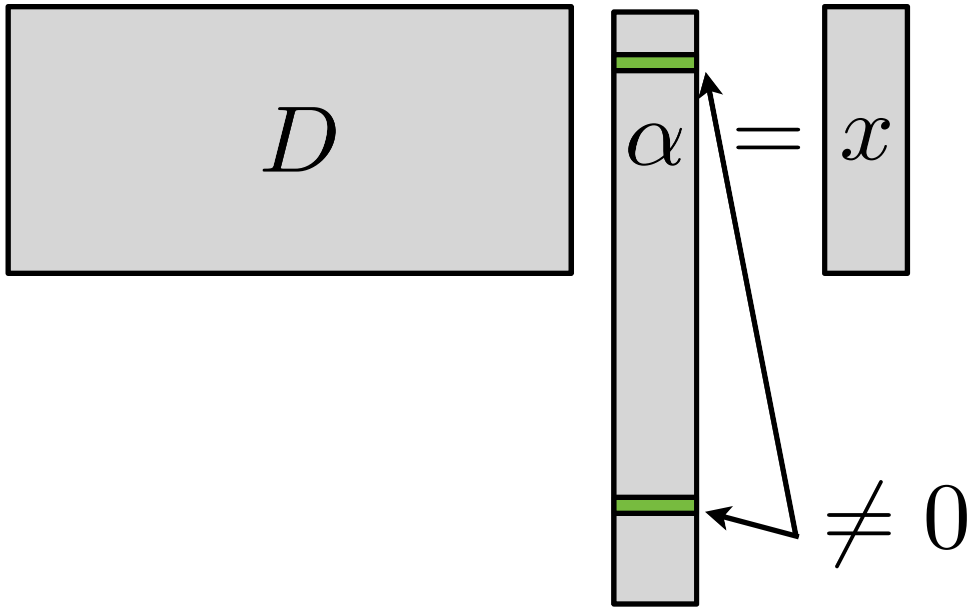
$$\underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda\|D^* x\|_1$$

Sparse Regularizations

Synthesis

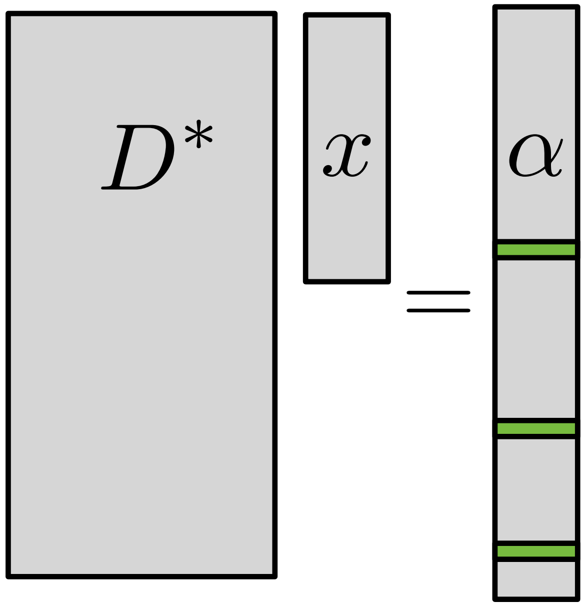
$$\underset{\alpha \in \mathbb{R}^Q}{\operatorname{argmin}} \frac{1}{2}\|y - \Psi\alpha\|_2^2 + \lambda\|\alpha\|_1$$

$$\Psi = \Phi D \qquad x = D\alpha$$



Analysis

$$\underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda\|D^* x\|_1$$

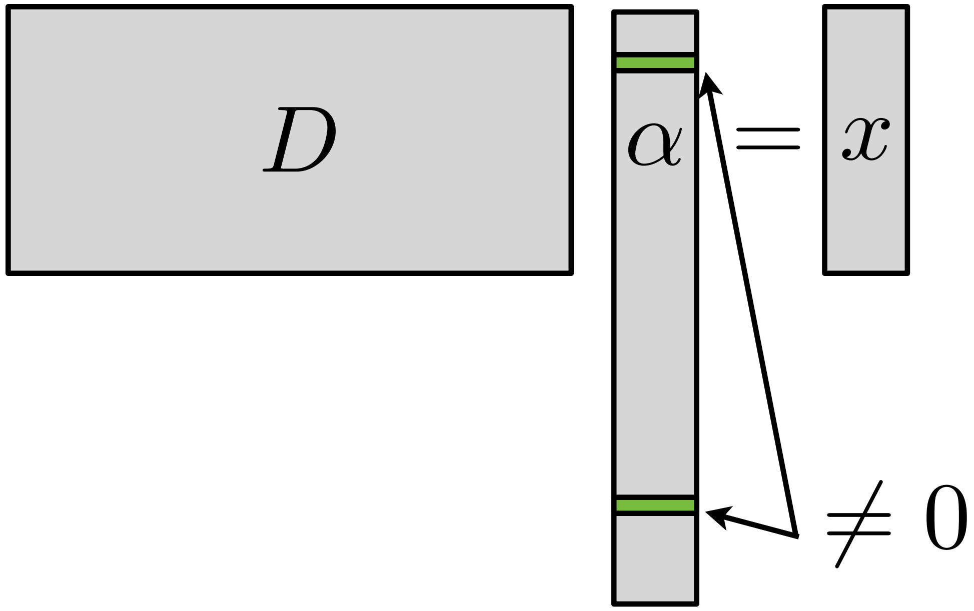


Sparse Regularizations

Synthesis

$$\underset{\alpha \in \mathbb{R}^Q}{\operatorname{argmin}} \ \frac{1}{2}\|y - \Psi\alpha\|_2^2 + \lambda\|\alpha\|_1$$

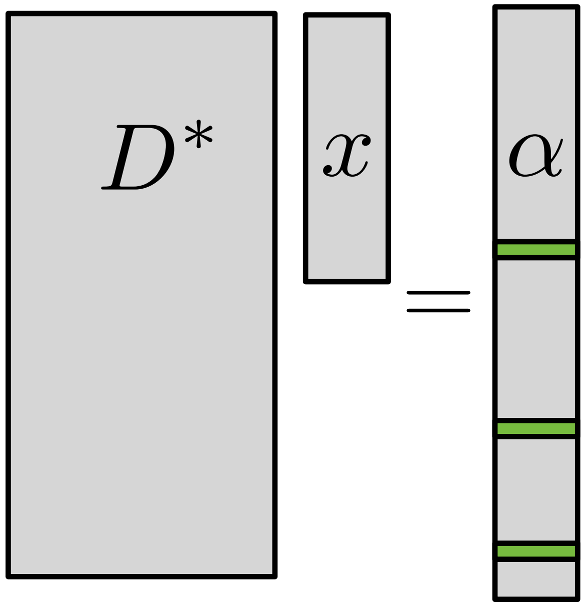
$$\Psi = \Phi D \qquad x = D\alpha$$



Sparse approx. of $x^\star$ in D

Analysis

$$\underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda\|D^* x\|_1$$

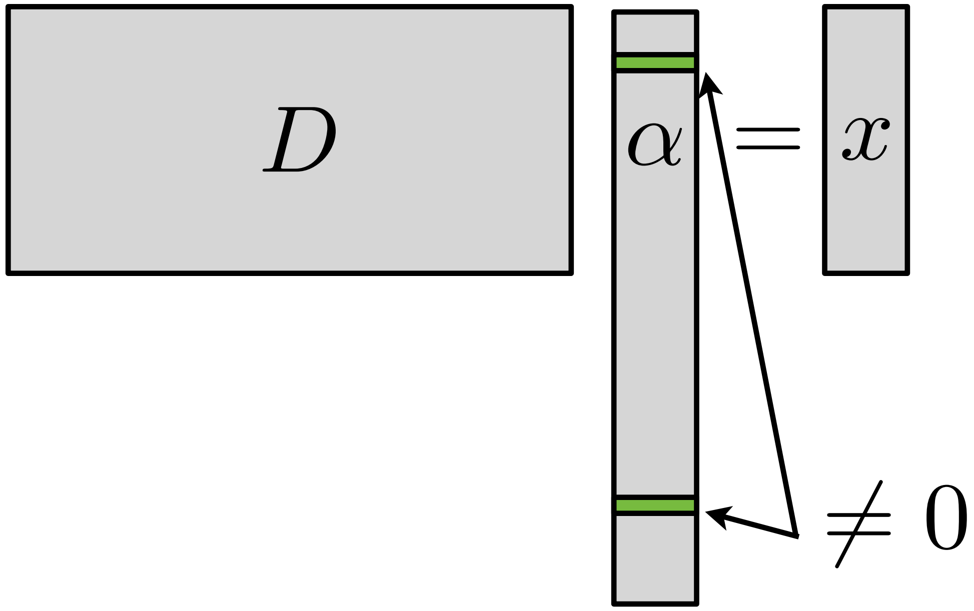


Sparse Regularizations

Synthesis

$$\underset{\alpha \in \mathbb{R}^Q}{\operatorname{argmin}} \frac{1}{2}\|y - \Psi\alpha\|_2^2 + \lambda\|\alpha\|_1$$

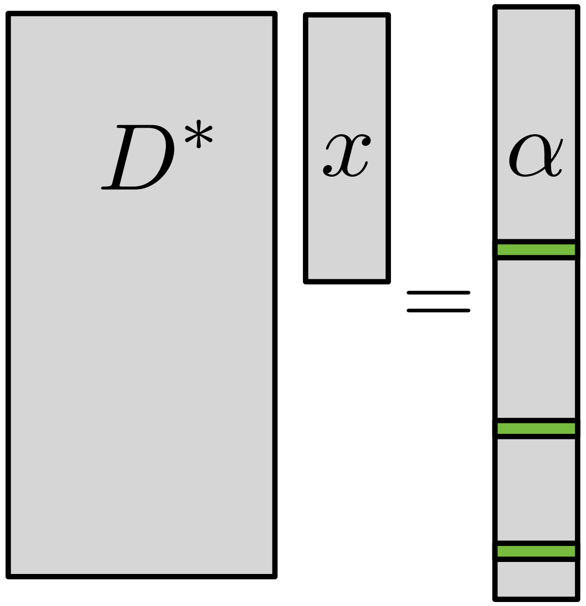
$$\Psi = \Phi D \qquad x = D\alpha$$



Sparse approx. of $x^\star$ in D

Analysis

$$\underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda\|D^* x\|_1$$



Correlation of $x^\star$ and D sparse

Support and Signal Model

$$x^\star \in \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1 \qquad \mathcal{P}(y, \lambda)$$

$$I = \operatorname{supp}(D^* x^\star), J = I^c$$

Support and Signal Model

$$x^\star \in \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1 \qquad \mathcal{P}(y, \lambda)$$

$$I = \operatorname{supp}(D^* x^\star), J = I^c$$

Definition

$$\operatorname{Ker} D_J^* = \mathcal{G}_J$$

Support and Signal Model

$$x^\star \in \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1 \qquad \mathcal{P}(y, \lambda)$$

$$I = \operatorname{supp}(D^* x^\star), J = I^c$$

Definition

$$\operatorname{Ker} D_J^* = \mathcal{G}_J$$

Signal model : *"Union of subspace"*

$$\Theta = \bigcup_{k \in \{1 \dots P\}} \Theta_k \quad \text{where} \quad \Theta_k = \{\mathcal{G}_J \setminus \dim \mathcal{G}_J = k\}$$

$$x^\star \in \mathcal{G}_J$$

Support and Signal Model

$$x^\star \in \operatorname*{argmin}_{x \in \mathbb{R}^N} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1 \qquad \mathcal{P}(y,\lambda)$$

$$I = \operatorname{supp}(D^* x^\star), J = I^c$$

Definition

$$\operatorname{Ker} D_J^* = \mathcal{G}_J$$

Signal model : *"Union of subspace"*

$$\Theta = \bigcup_{k \in \{1 \ldots P\}} \Theta_k \quad \text{where} \quad \Theta_k = \{\mathcal{G}_J \setminus \dim \mathcal{G}_J = k\}$$

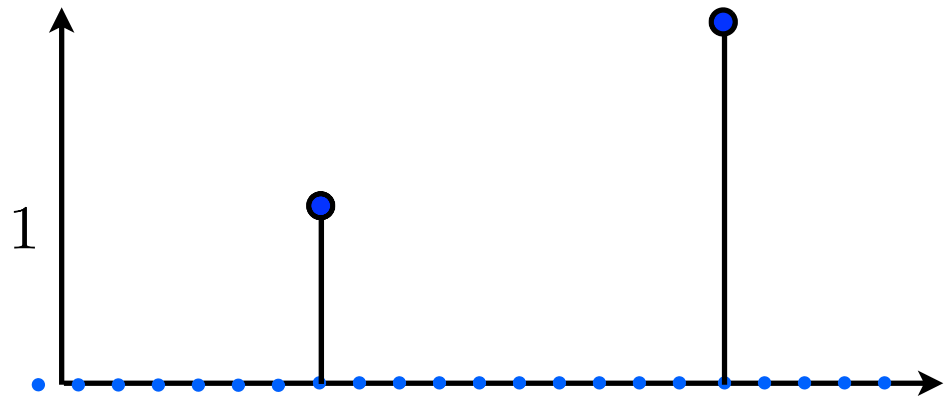
$$x^\star \in \mathcal{G}_J$$

Hypothesis: $\operatorname{Ker} \Phi \cap \operatorname{Ker} D^* = \{0\}$

Examples of Signal Model

Identity

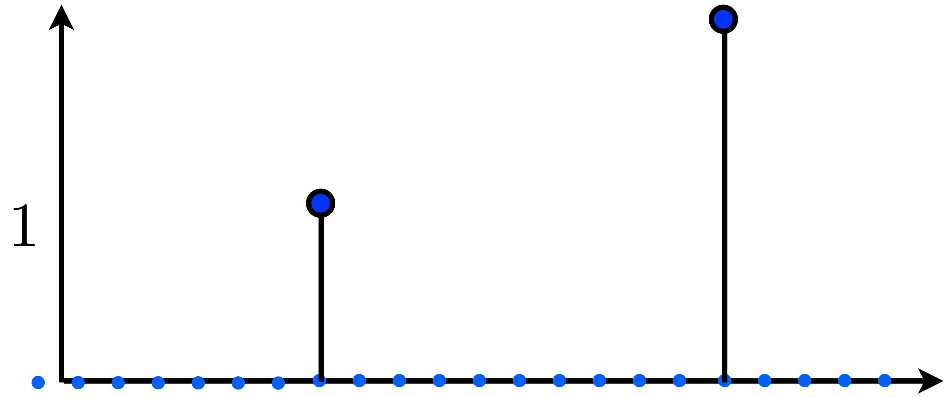
Θ_k : k-sparse signals



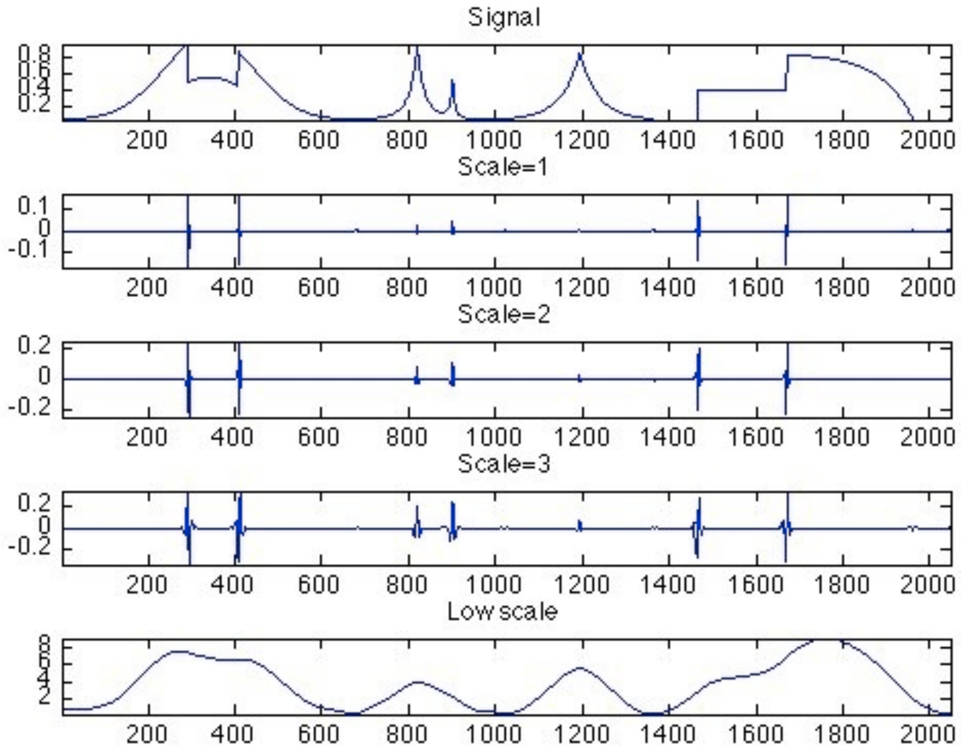
Examples of Signal Model

Identity

Θ_k : k-sparse signals



shift invariant wavelet frame

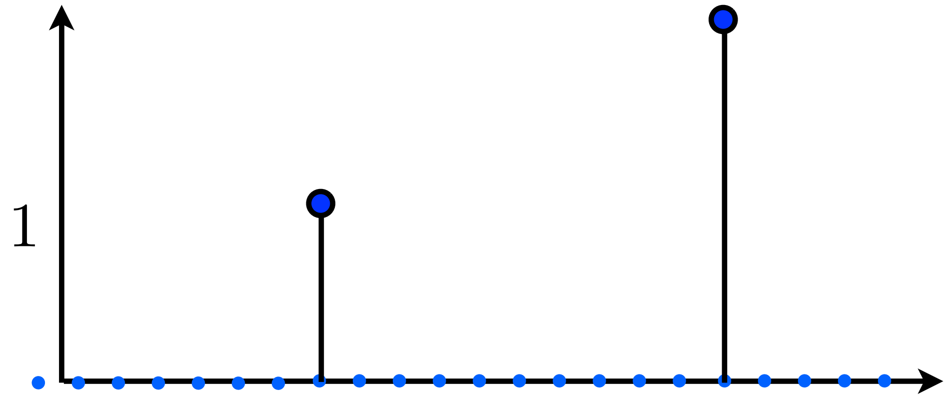


Credit to G. Peyré

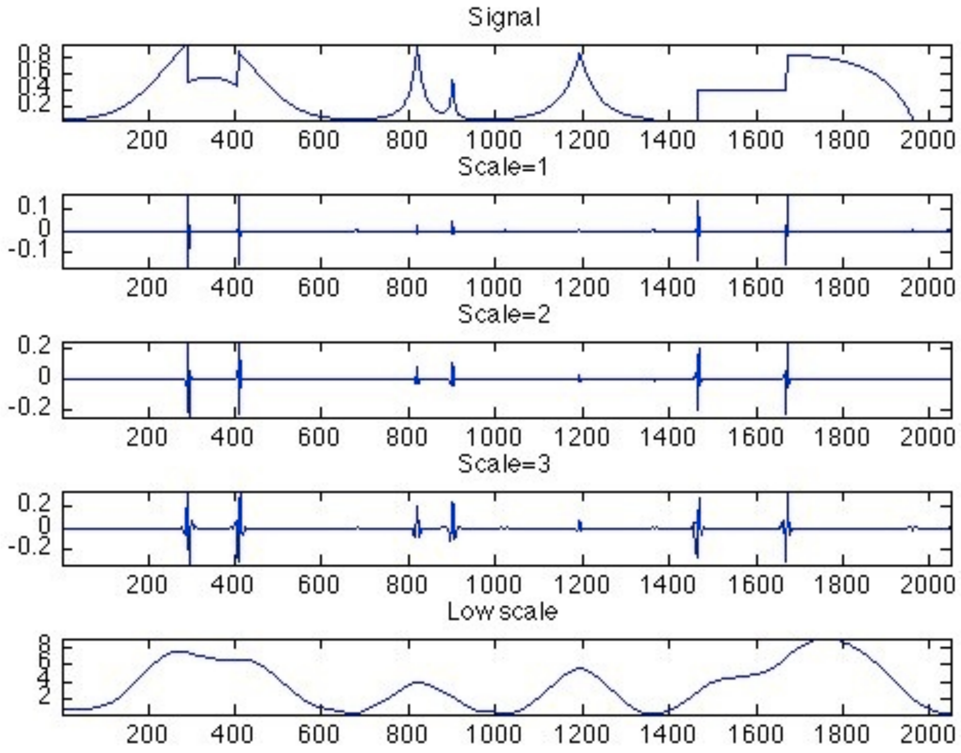
Examples of Signal Model

Identity

Θ_k : k-sparse signals



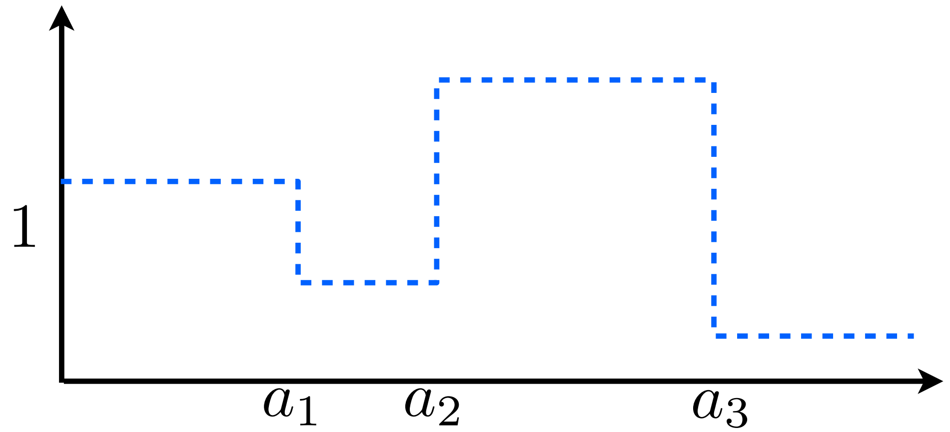
shift invariant wavelet frame



Credit to G. Peyré

finite difference operator

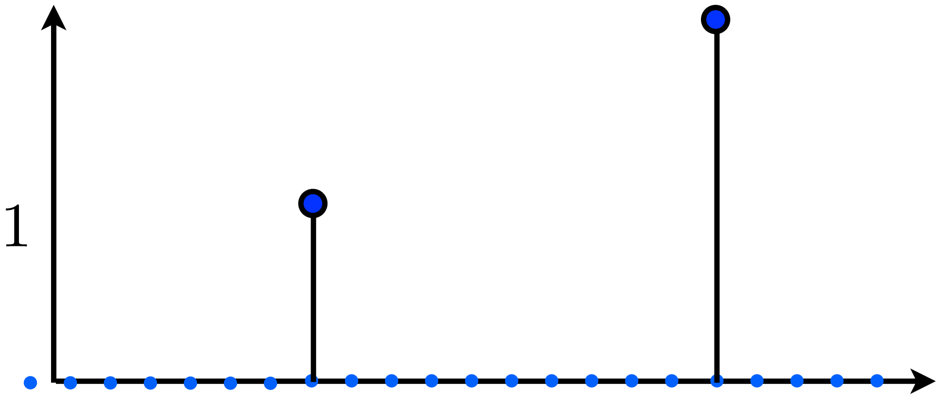
Θ_k : piecewise constant signals with $k-1$ steps



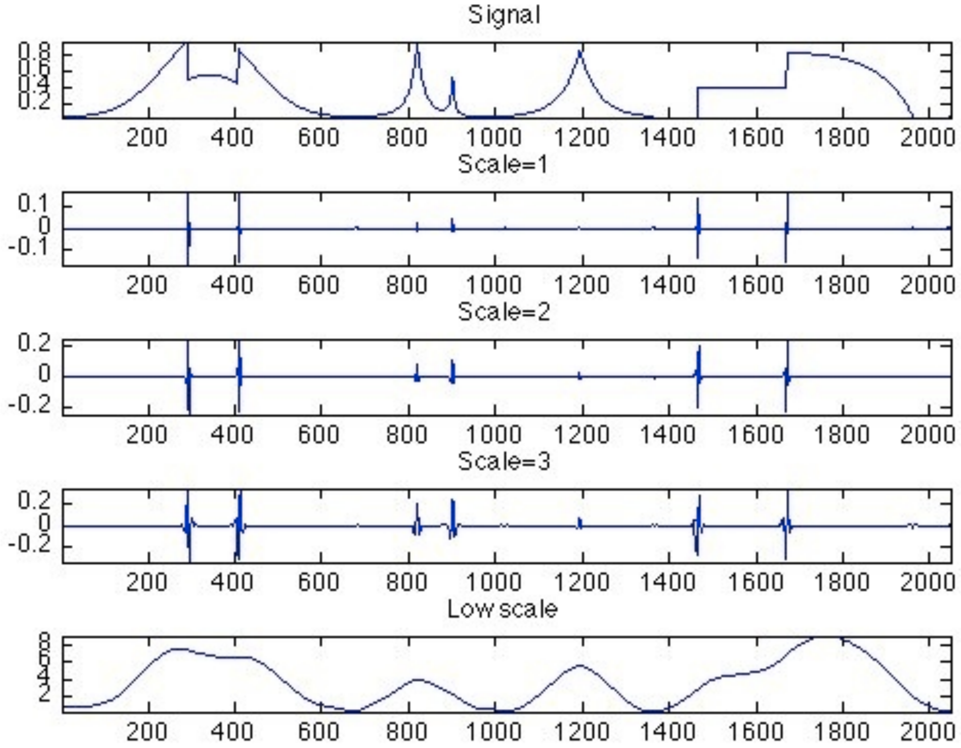
Examples of Signal Model

Identity

Θ_k : k-sparse signals



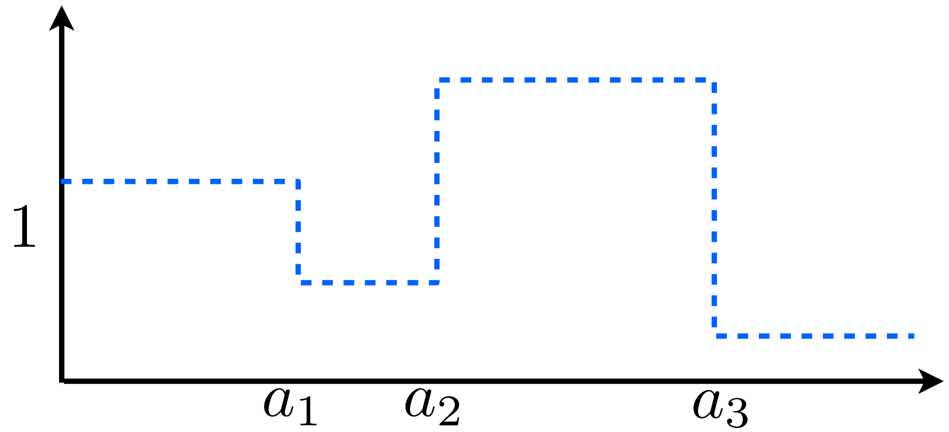
shift invariant wavelet frame



Credit to G. Peyré

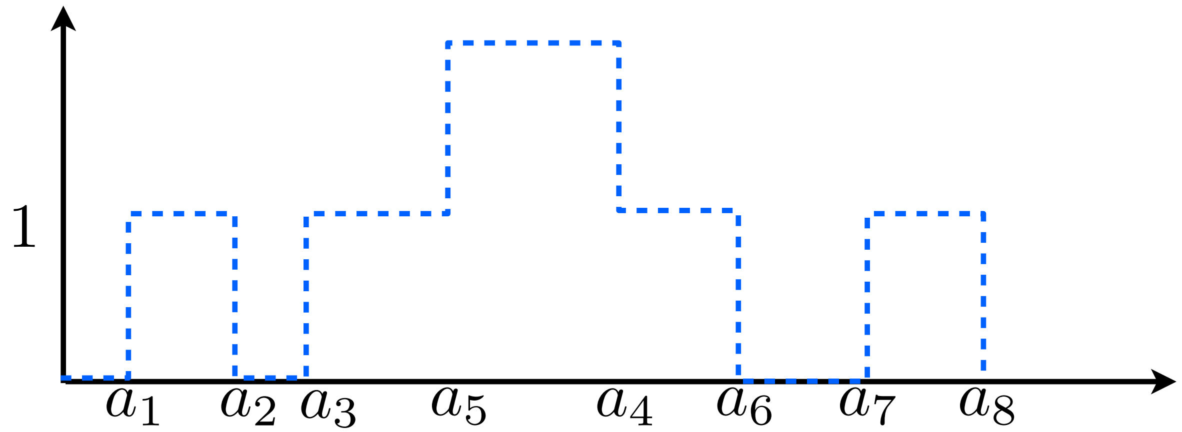
finite difference operator

Θ_k : piecewise constant signals with $k-1$ steps



fussed lasso

Θ_k : sum of k interval characteristic functions



Remember !

Synthesis

$$\alpha^\star = \underset{\alpha \in \mathbb{R}^Q}{\operatorname{argmin}} \frac{1}{2}\|y - \Psi\alpha\|_2^2 + \lambda\|\alpha\|_1$$

Analysis

$$x^\star = \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda\|D^* x\|_1 \qquad \mathcal{P}(y,\lambda)$$

$$\Big\downarrow \; \lambda \to 0$$

$$x^\star = \underset{\Phi x = y}{\operatorname{argmin}} \|D^* x\|_1 \qquad \mathcal{P}(y,0)$$

Toward a Better Understanding

Local behavior ?

Properties of $x^\star$ solution of $\mathcal{P}(y,\lambda)$ as a function of y

Toward a Better Understanding

Local behavior ?

$$\text{Properties of } x^\star \text{ solution of } \mathcal{P}(y, \lambda) \text{ as a function of } y$$

Noiseless identifiability ?

$$\text{Is } x_0 \text{ the unique solution of } \mathcal{P}(\Phi x_0, 0) \text{ ?}$$

Toward a Better Understanding

Local behavior ?

Properties of $x^\star$ solution of $\mathcal{P}(y,\lambda)$ as a function of y

Noiseless identifiability ?

Is x_0 the unique solution of $\mathcal{P}(\Phi x_0, 0)$?

Noise robustness ?

What can we say about $\|x^\star - x_0\|$ for noisy observations ?

From Synthesis to Analysis Results

— Previous works in synthesis

[Fuchs, Tropp, Dossal]: address these questions

From Synthesis to Analysis Results

— Previous works in synthesis

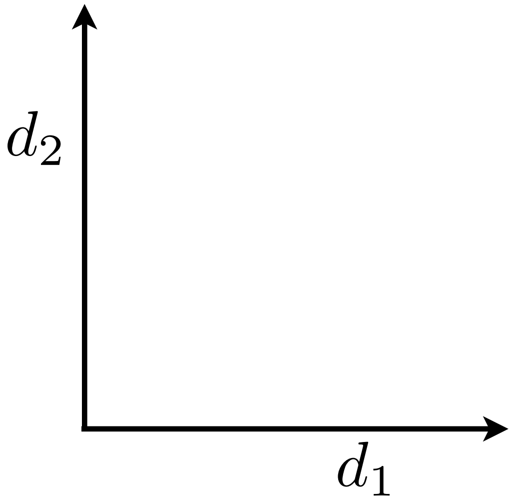
[Fuchs, Tropp, Dossal]: address these questions

— Similar problem but much more difficulties in analysis

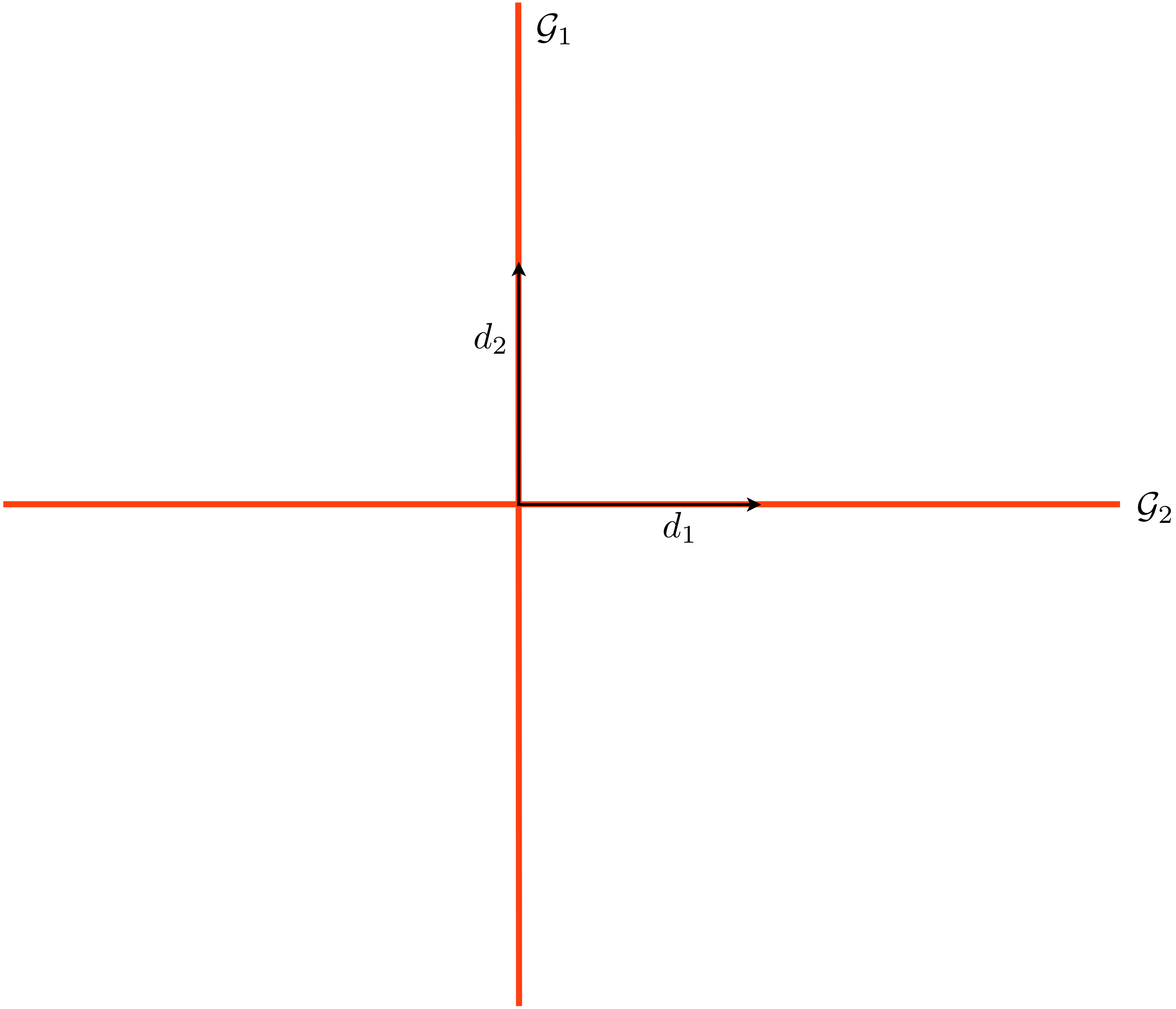
Geometry of the problem ?

From Synthesis to Analysis Results

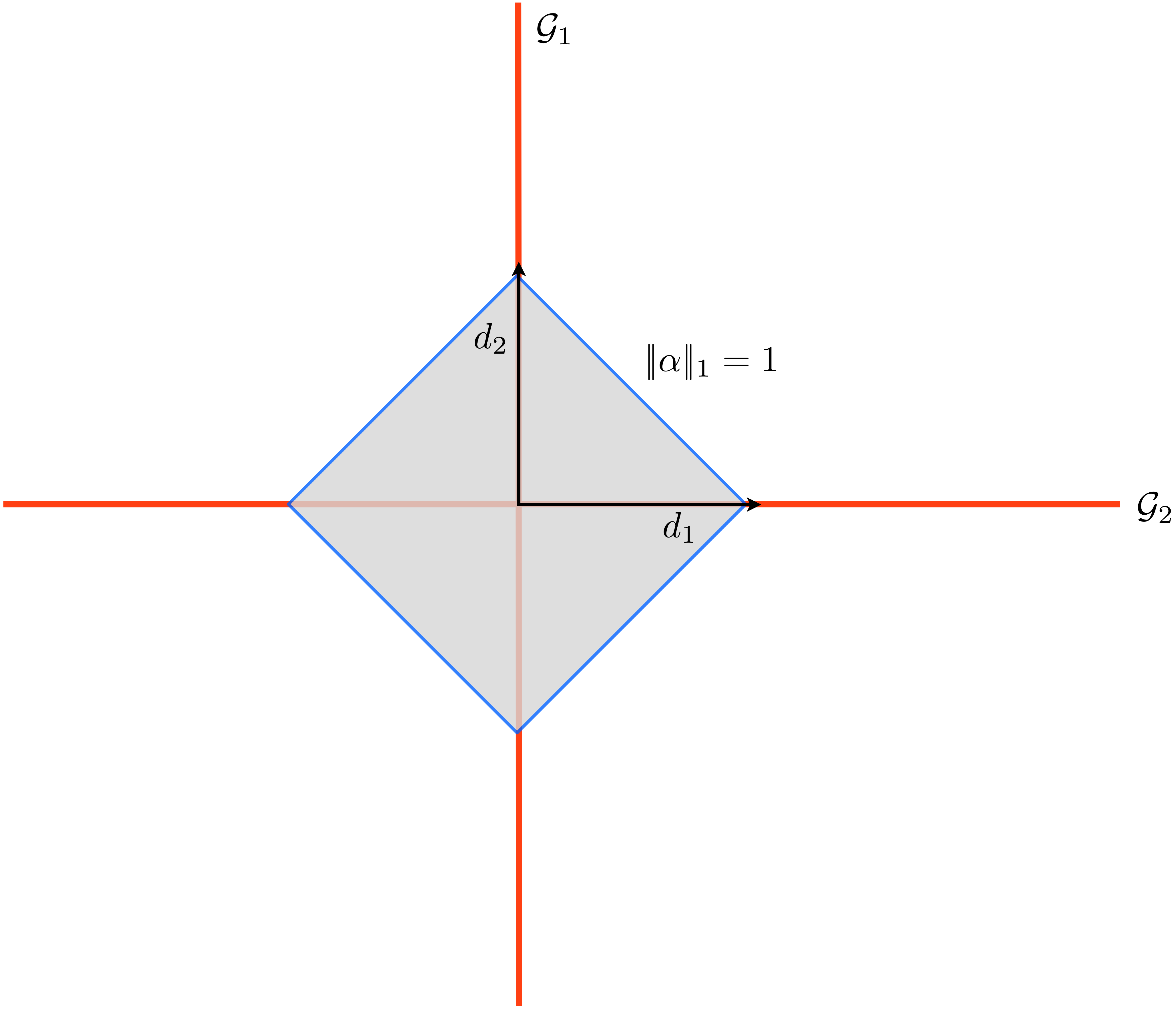
From Synthesis to Analysis Results



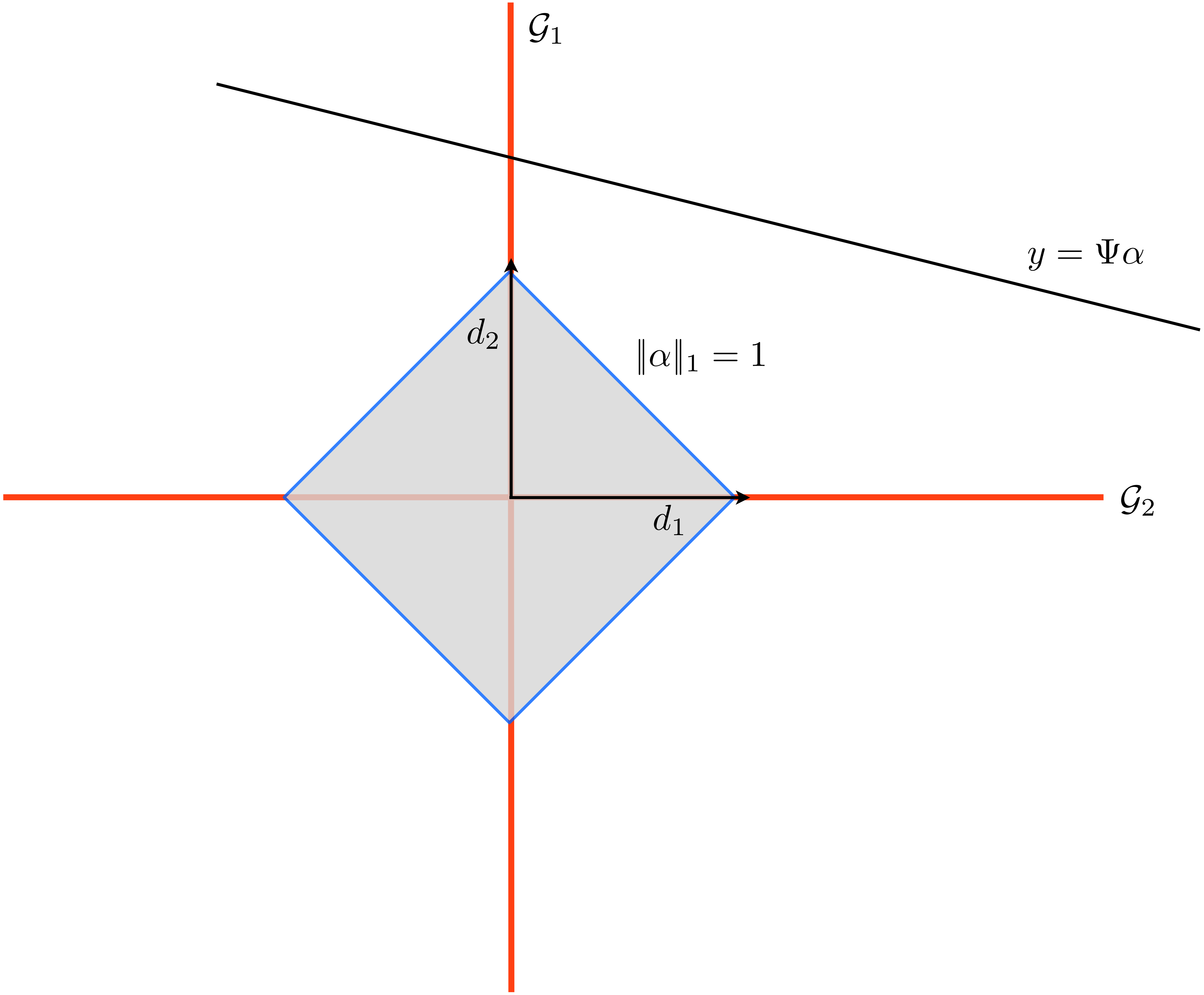
From Synthesis to Analysis Results



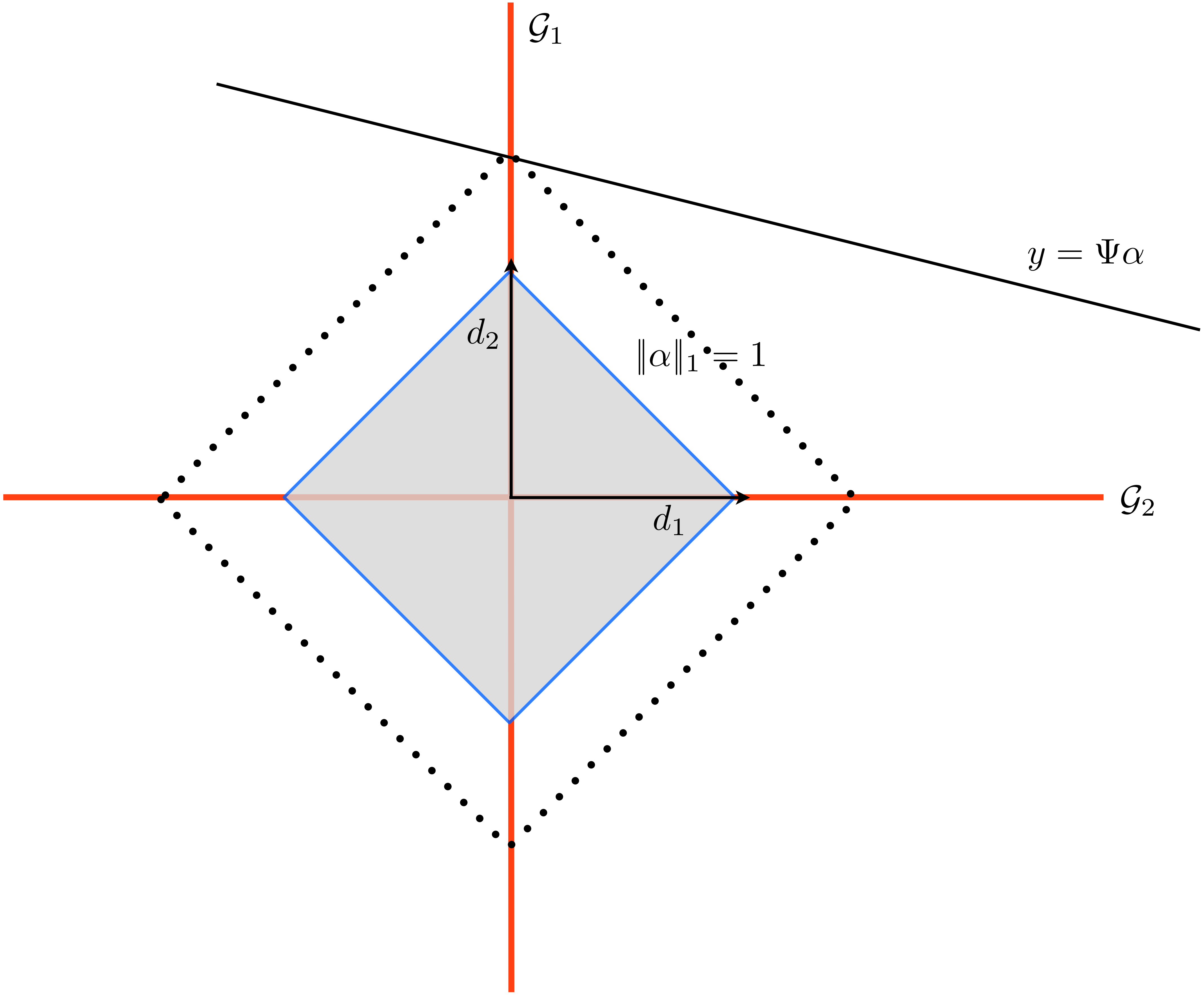
From Synthesis to Analysis Results



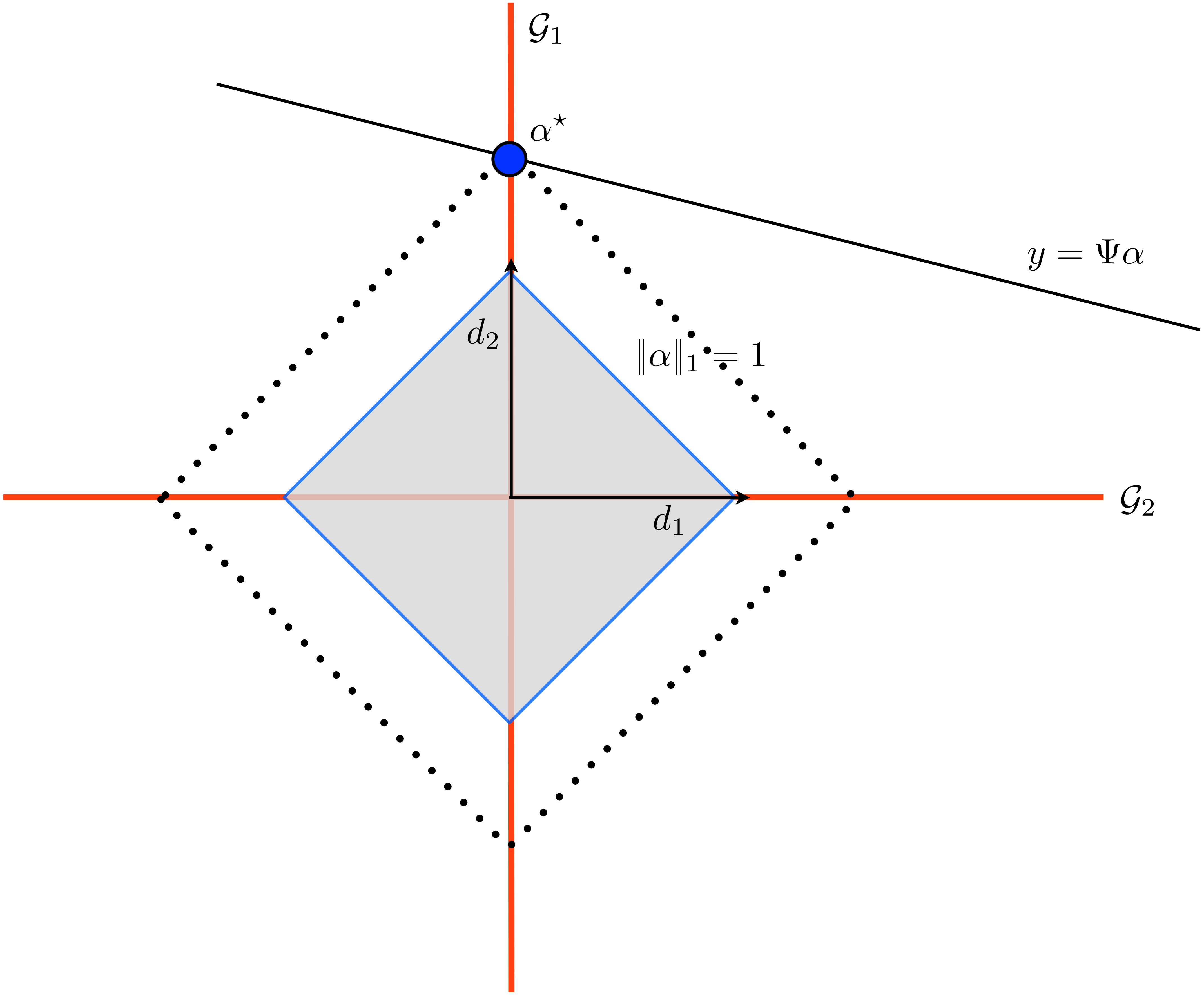
From Synthesis to Analysis Results



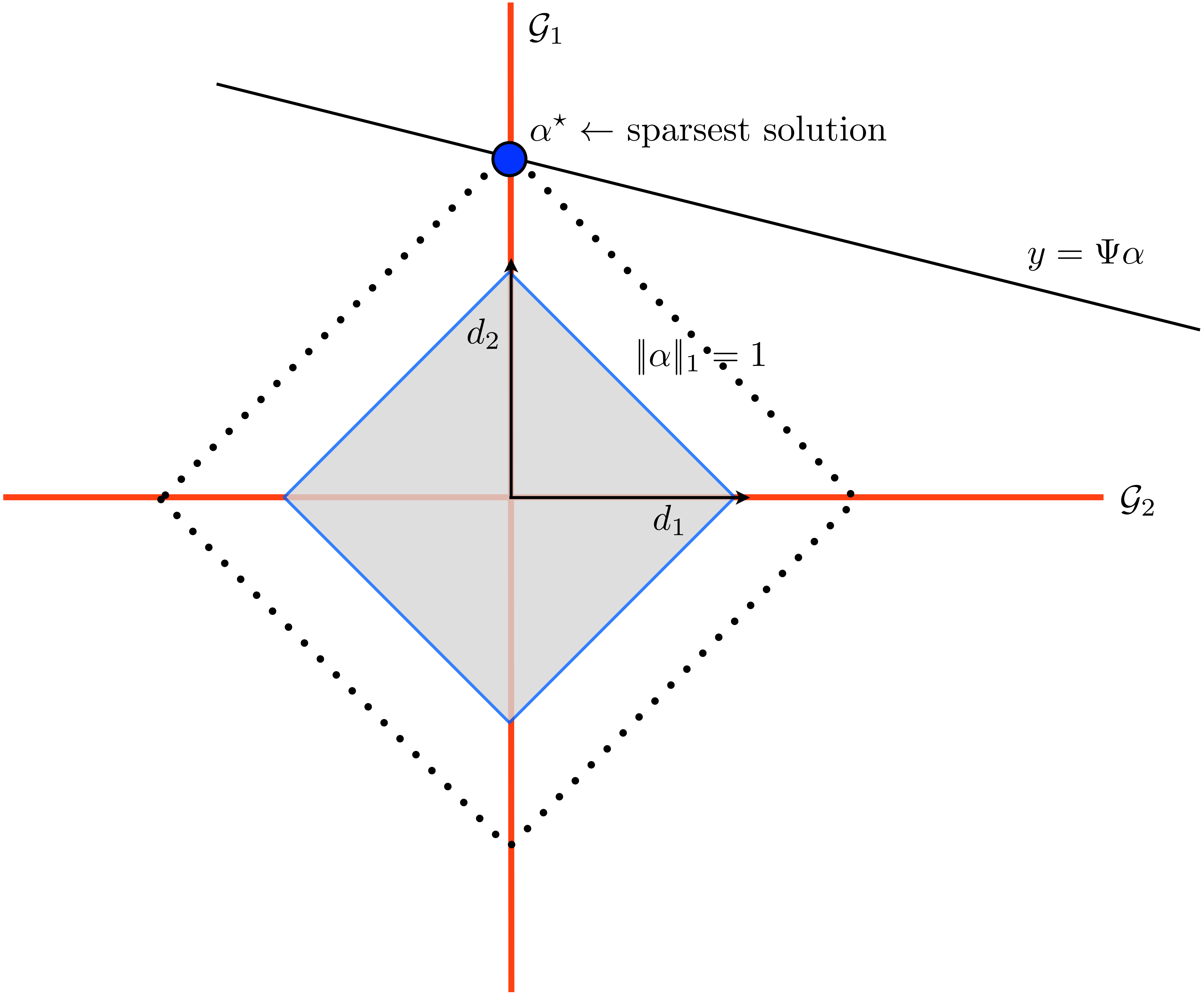
From Synthesis to Analysis Results



From Synthesis to Analysis Results

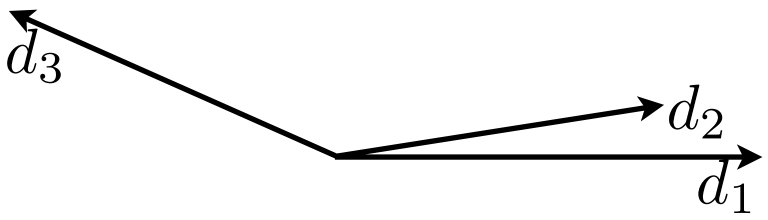


From Synthesis to Analysis Results

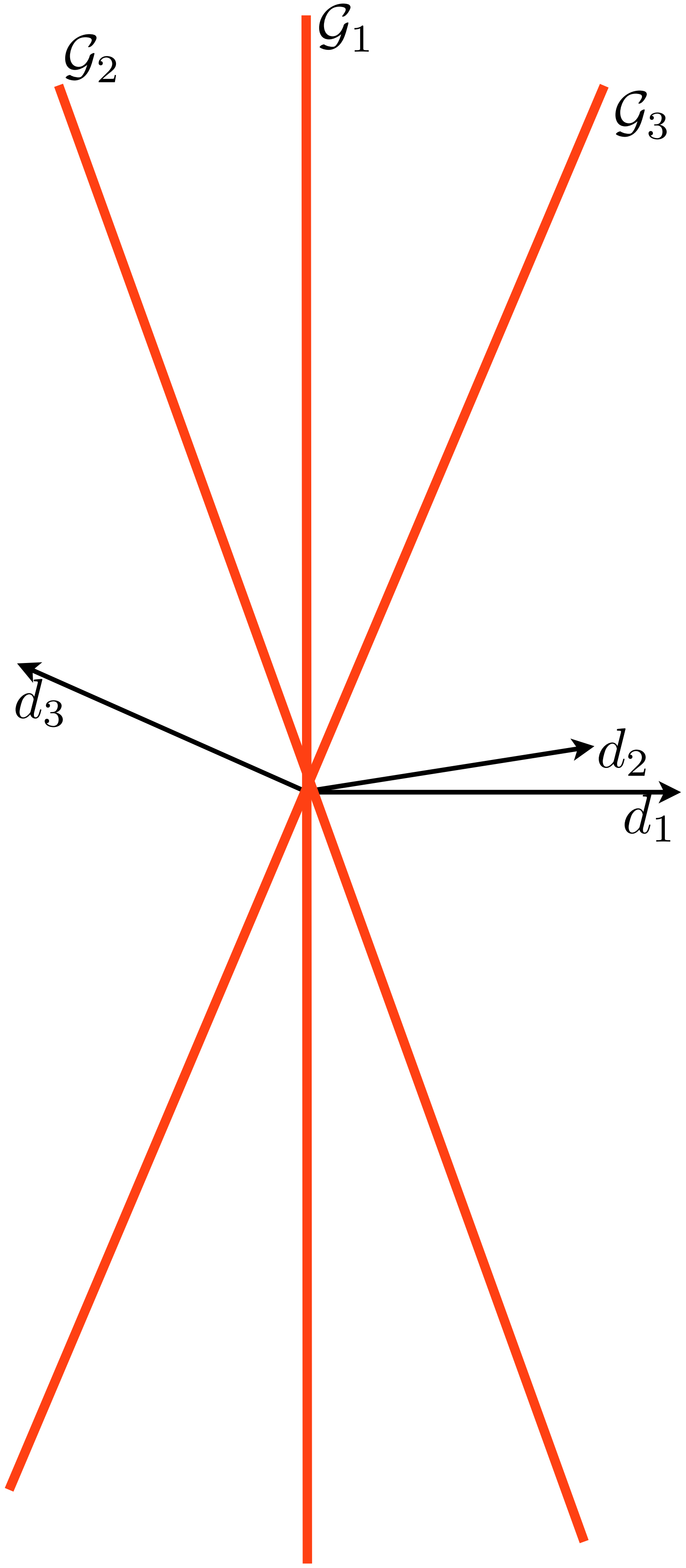


From Synthesis to Analysis Results

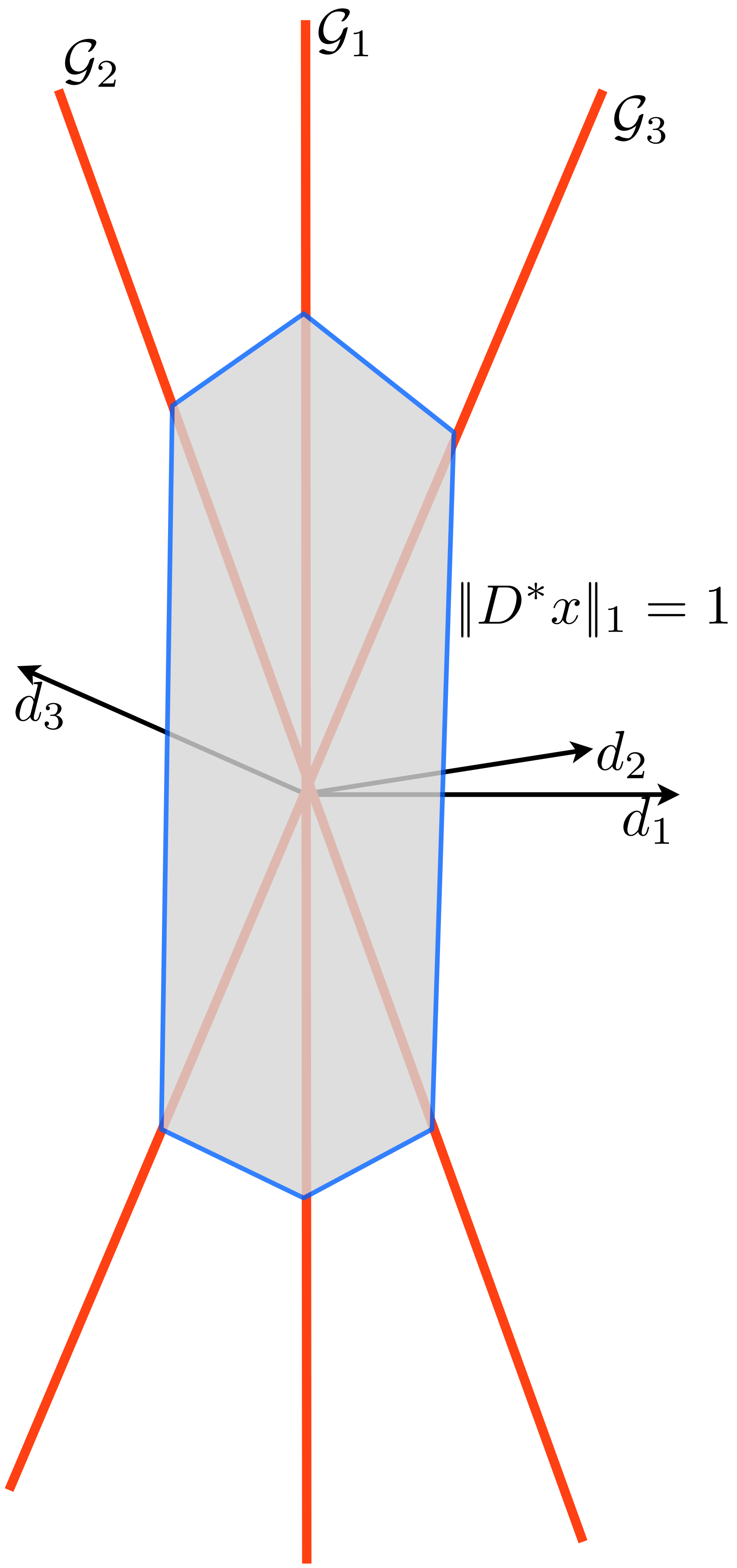
From Synthesis to Analysis Results



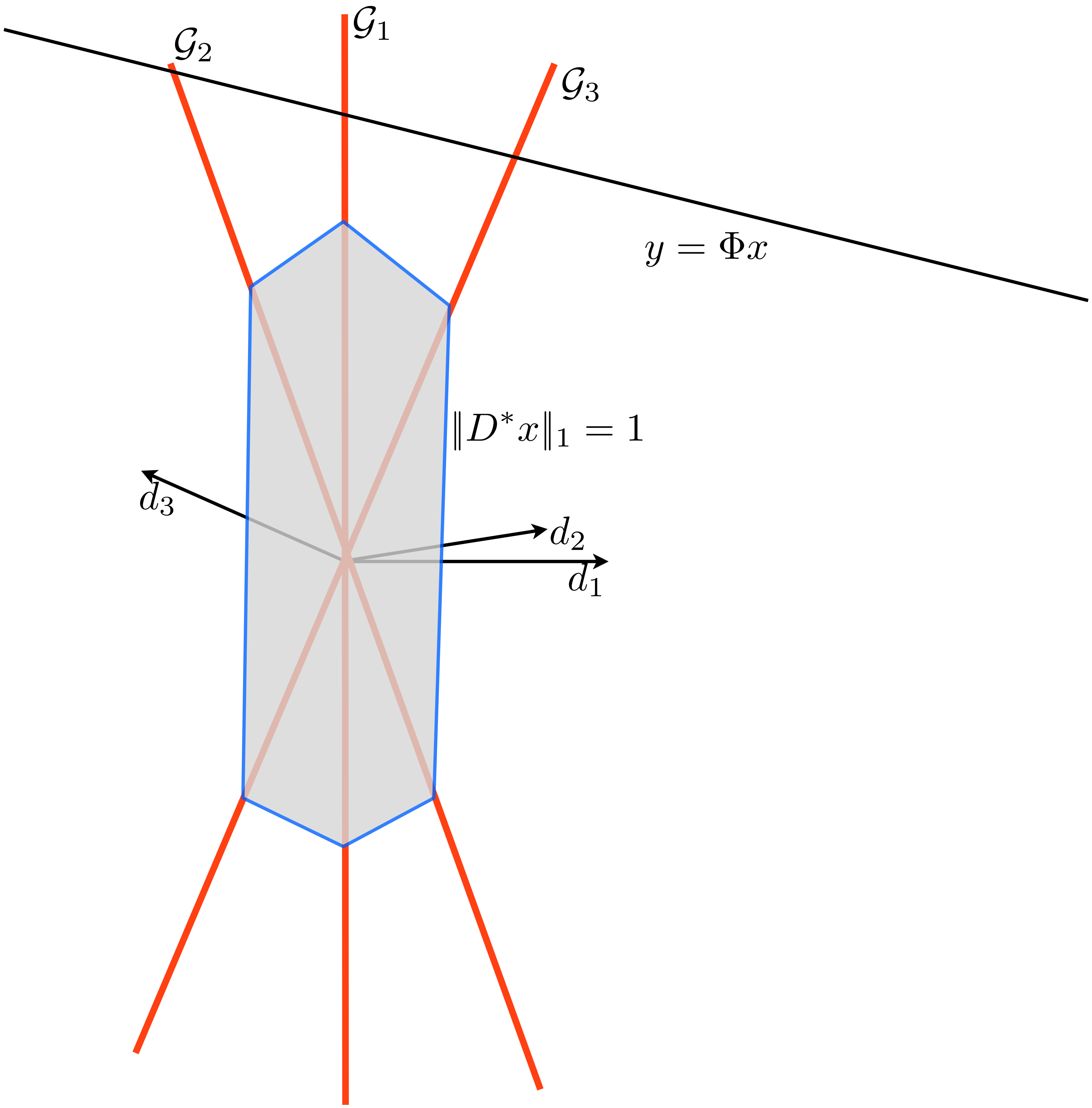
From Synthesis to Analysis Results



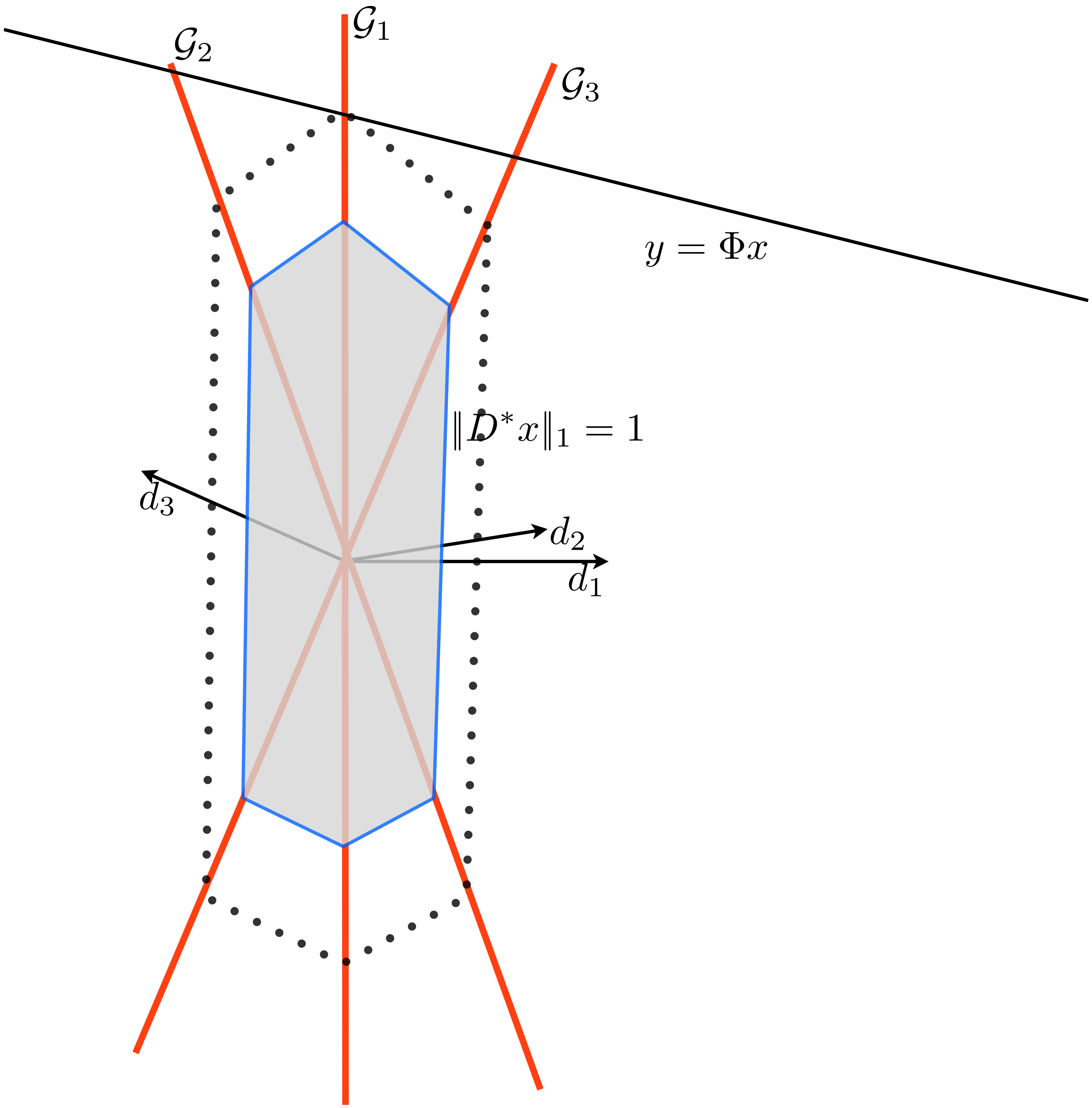
From Synthesis to Analysis Results



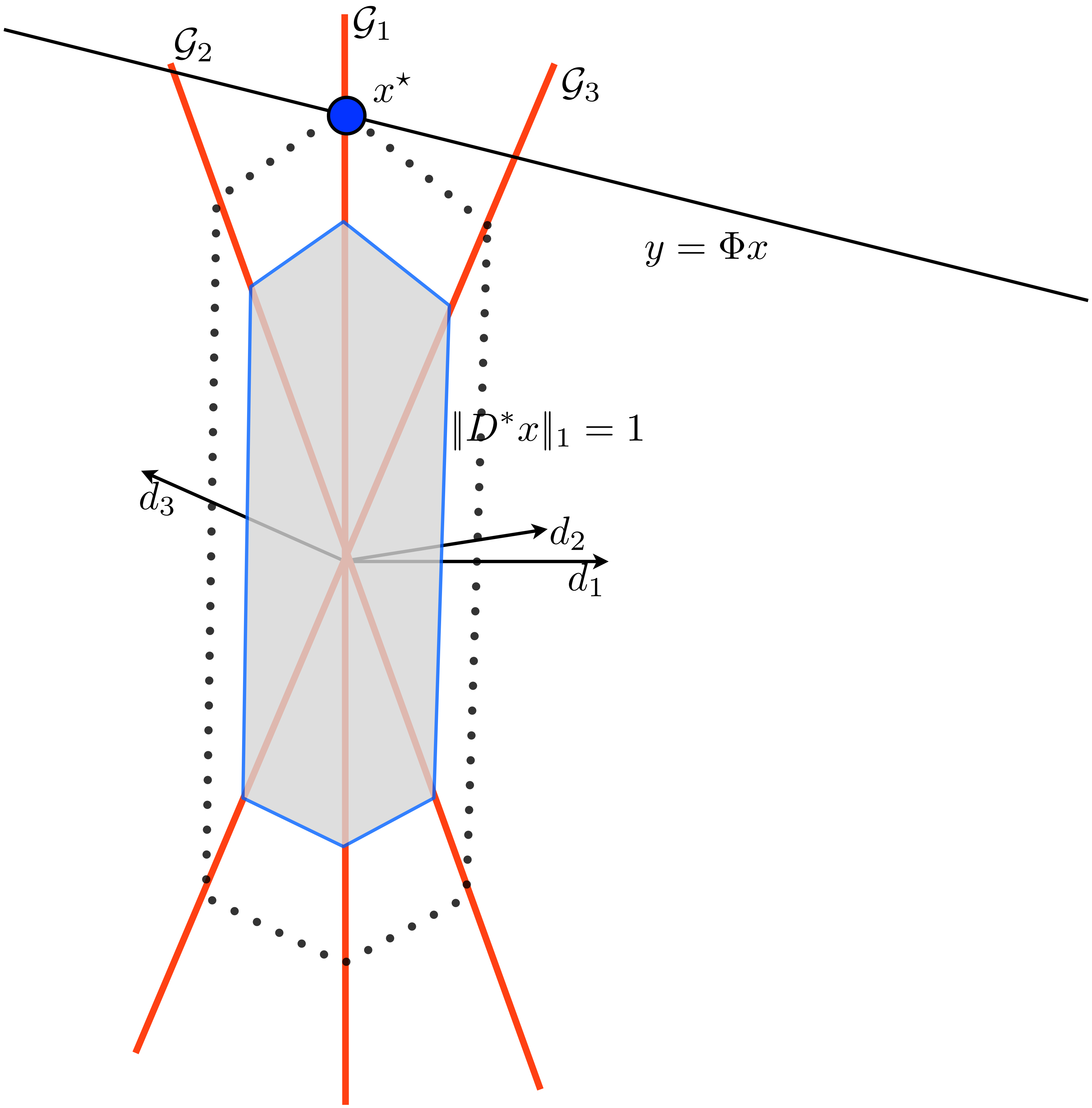
From Synthesis to Analysis Results



From Synthesis to Analysis Results



From Synthesis to Analysis Results



Overview

- Analysis vs. Synthesis Regularization
- Local Parameterization of Analysis Regularization
- Identifiability and Stability
- Numerical Evaluation
- Perspectives

Analysis is Piecewise Affine

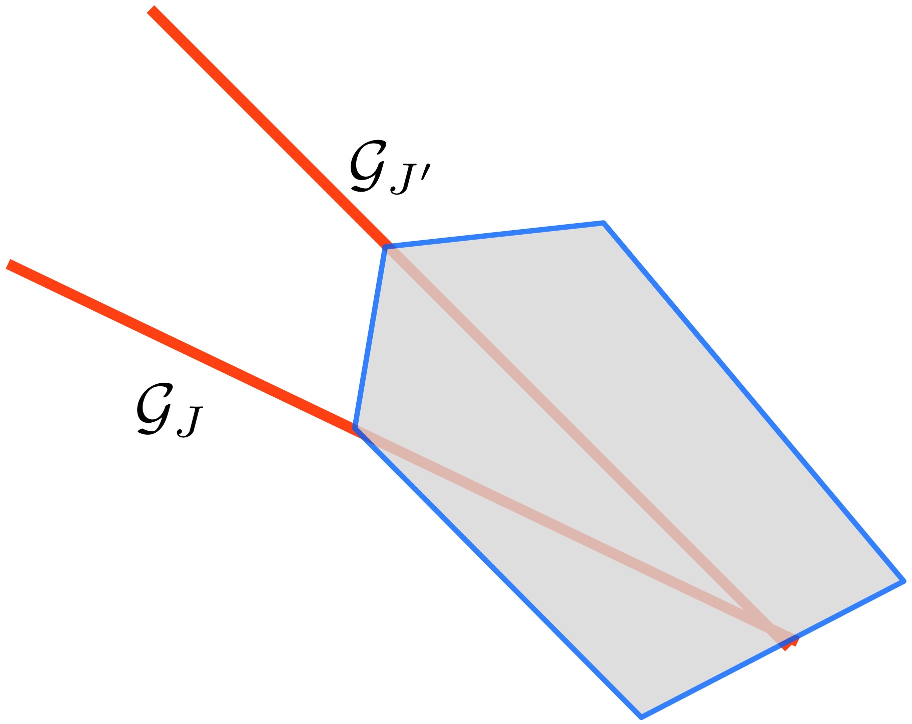
Main idea: $\mathcal{G}_J$ is stable,

i.e solutions of $\mathcal{P}(y,\lambda)$ and $\mathcal{P}(y+\varepsilon,\lambda)$ lives in the same $\mathcal{G}_J$.

Analysis is Piecewise Affine

Main idea: $\mathcal{G}_J$ is stable,

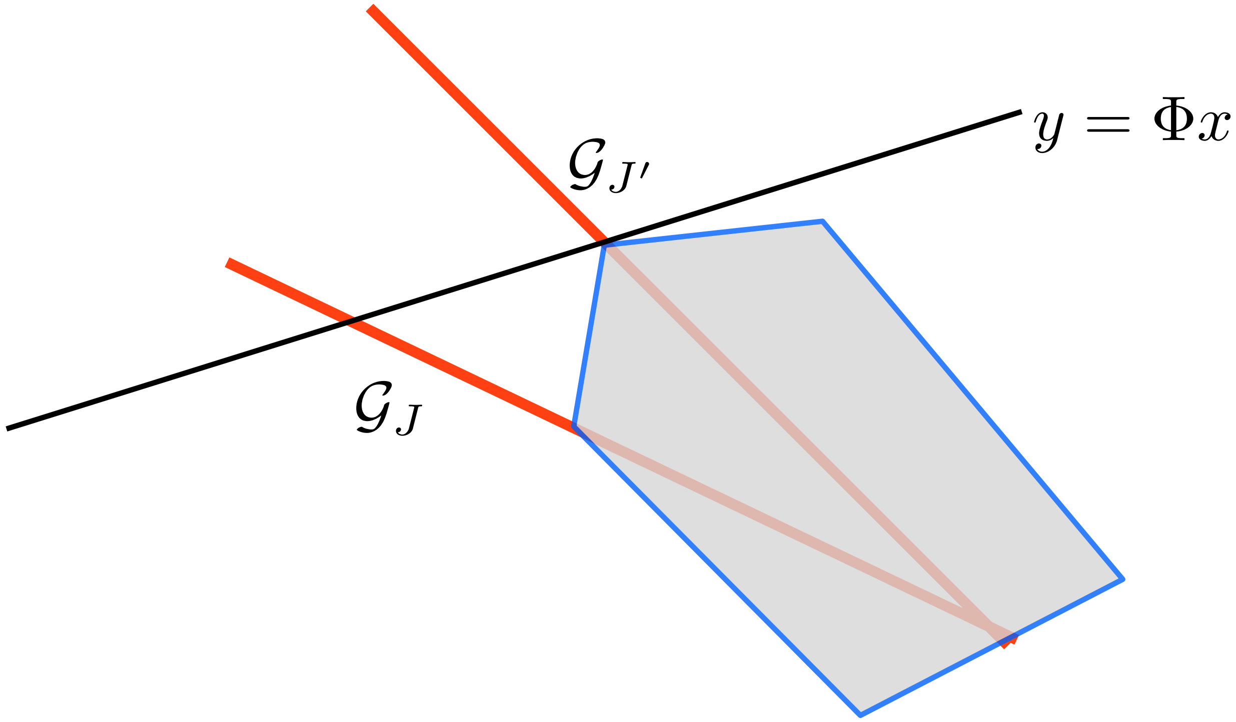
i.e solutions of $\mathcal{P}(y, \lambda)$ and $\mathcal{P}(y+\varepsilon, \lambda)$ lives in the same $\mathcal{G}_J$.



Analysis is Piecewise Affine

Main idea: $\mathcal{G}_J$ is stable,

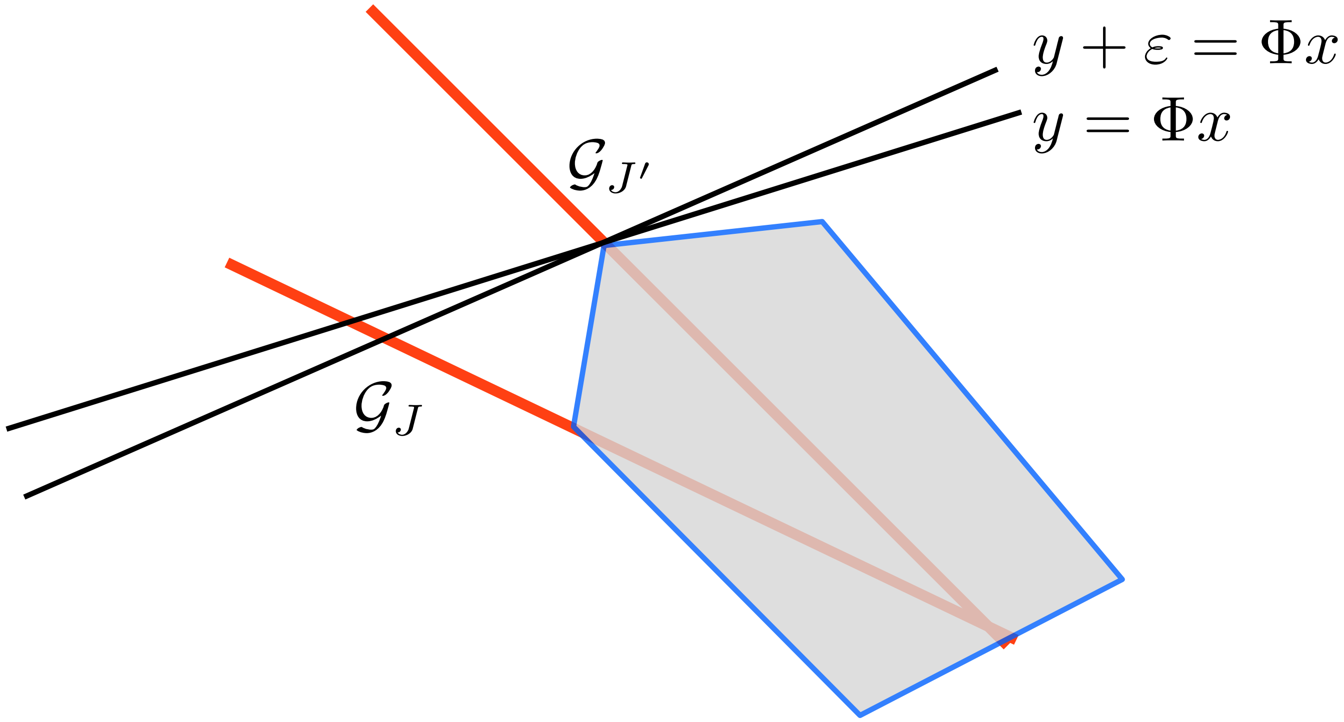
i.e solutions of $\mathcal{P}(y,\lambda)$ and $\mathcal{P}(y+\varepsilon,\lambda)$ lives in the same $\mathcal{G}_J$.



Analysis is Piecewise Affine

Main idea: $\mathcal{G}_J$ is stable,

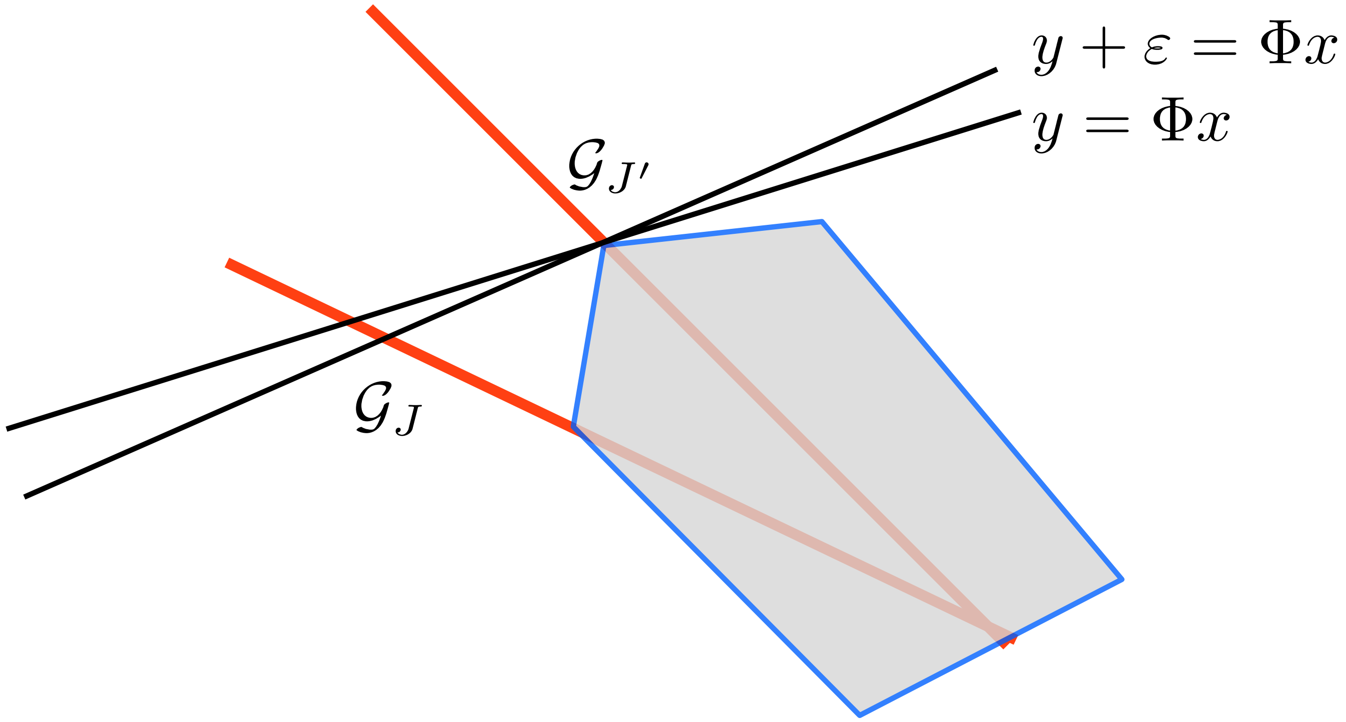
i.e solutions of $\mathcal{P}(y,\lambda)$ and $\mathcal{P}(y+\varepsilon,\lambda)$ lives in the same $\mathcal{G}_J$.



Analysis is Piecewise Affine

Main idea: $\mathcal{G}_J$ is stable,

i.e solutions of $\mathcal{P}(y,\lambda)$ and $\mathcal{P}(y+\varepsilon,\lambda)$ lives in the same $\mathcal{G}_J$.

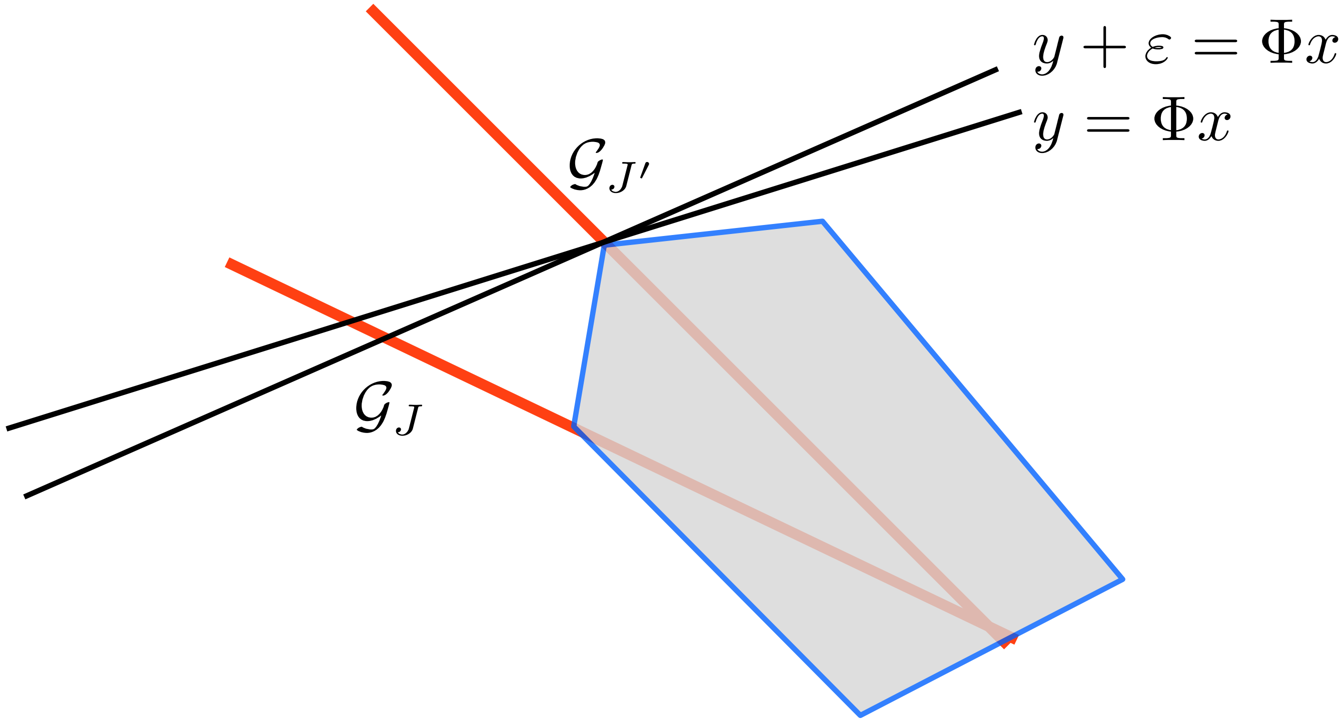


Affine function: $\bar{y} \mapsto x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$

Analysis is Piecewise Affine

Main idea: $\mathcal{G}_J$ is stable,

i.e solutions of $\mathcal{P}(y,\lambda)$ and $\mathcal{P}(y+\varepsilon,\lambda)$ lives in the same $\mathcal{G}_J$.



Affine function: $\bar{y}\mapsto x(\bar{y}) = A\Phi^*\bar{y} - \lambda A D_I s$

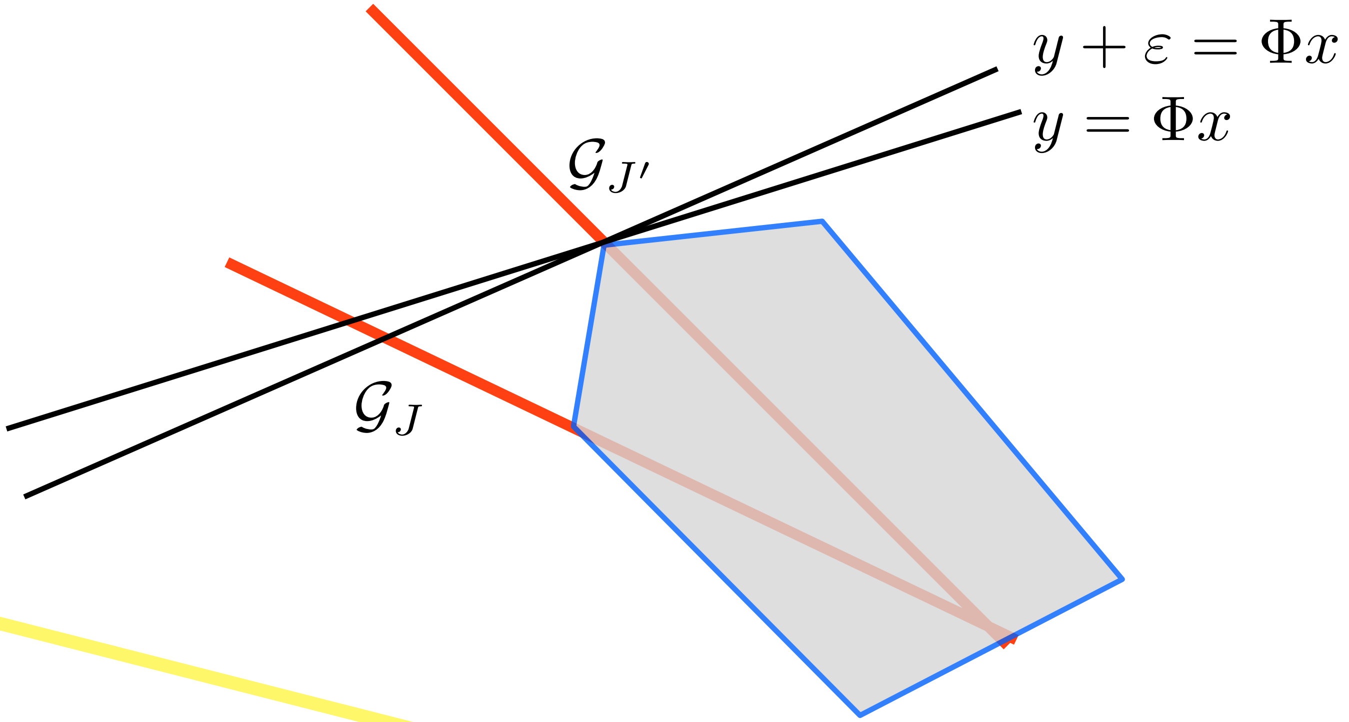
Theorem 1

Except for $y\in\mathcal{H}$, if $\bar{y}$ is close enough from y, then $x(\bar{y})$ is a solution of $\mathcal{P}(\bar{y},\lambda)$.

Analysis is Piecewise Affine

Main idea: $\mathcal{G}_J$ is stable,

i.e solutions of $\mathcal{P}(y, \lambda)$ and $\mathcal{P}(y + \varepsilon, \lambda)$ lives in the same $\mathcal{G}_J$.



definition in few minutes

Affine function: $\bar{y} \mapsto x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$

Theorem 1

Except for $y \in \mathcal{H}$, if $\bar{y}$ is close enough from y,
then $x(\bar{y})$ is a solution of $\mathcal{P}(\bar{y}, \lambda)$.

Sketch of the Proof

Problem : Lasso

$$x^\star \in \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \; \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1 \qquad \mathcal{P}(y, \lambda)$$

Sketch of the Proof

Problem : Lasso

$$x^\star \in \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda\|D^* x\|_1 \qquad \mathcal{P}(y,\lambda)$$

Support

$$I = \operatorname{supp}(D^* x^\star), J = I^c$$

Sketch of the Proof

Problem : Lasso

$$x^\star \in \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1 \qquad \mathcal{P}(y, \lambda)$$

Support

$$I = \operatorname{supp}(D^* x^\star), J = I^c$$

Subspace of analysis

$$\operatorname{Ker} D_J^* = \mathcal{G}_J$$

Sketch of the Proof

Problem : Lasso

$$x^\star \in \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1 \qquad \mathcal{P}(y, \lambda)$$

Support

$$I = \operatorname{supp}(D^* x^\star), J = I^c$$

Subspace of analysis

$$\operatorname{Ker} D_J^* = \mathcal{G}_J$$

Hypothesis

$$\operatorname{Ker} \Phi \cap \mathcal{G}_J = \{0\}$$

Sketch of the Proof

Problem : Lasso

$$x^\star \in \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1 \qquad \mathcal{P}(y, \lambda)$$

Support

$$I = \operatorname{supp}(D^* x^\star), J = I^c$$

Subspace of analysis

$$\operatorname{Ker} D_J^* = \mathcal{G}_J$$

Hypothesis

$$\operatorname{Ker} \Phi \cap \mathcal{G}_J = \{0\}$$

— We fix observations y

— $I, J, s = \operatorname{sign}(D^* x^\star)$ are fixed by $x^\star$

First Order Conditions

$$x^\star \in \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1 \qquad \mathcal{P}(y, \lambda)$$

First Order Conditions

$$x^\star \in \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1 \qquad \mathcal{P}(y, \lambda)$$

Non differentiable problem

$$x^\star \text{ is a minimum of } \mathcal{P}(y,\lambda) \text{ if, and only if, } 0 \in \partial f(x^\star)$$

First Order Conditions

$$x^\star \in \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1 \qquad \mathcal{P}(y, \lambda)$$

Non differentiable problem

$$x^\star \text{ is a minimum of } \mathcal{P}(y,\lambda) \text{ if, and only if, } 0 \in \partial f(x^\star)$$

First-order conditions of Lasso

$$x^\star \text{ solution of } \mathcal{P}(y,\lambda) \Leftrightarrow \exists \sigma \in \Sigma_y(x^\star), \|\sigma\|_\infty \leqslant 1$$

$$\Sigma_y(x) = \left\{ \sigma \in \mathbb{R}^{|J|} \setminus \underbrace{\Phi^*(\Phi x - y)}_{\text{Gradient}} + \underbrace{\lambda D_I s + \lambda D_J \sigma}_{\text{Subdifferential}} = 0 \right\}$$

A Solution of Lasso

$$x(y) \in \underset{x \in \mathcal{G}_J}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

A Solution of Lasso

$$x(y) \in \operatorname*{argmin}_{x \in \mathcal{G}_J} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

How to implicit a solution ?

A Solution of Lasso

$$x(y) \in \operatorname*{argmin}_{x \in \mathcal{G}_J} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

How to implicit a solution ?

$$\Phi^* \Phi x(y) = \Phi^* y - \lambda D_I s - \lambda D_J \sigma$$

A Solution of Lasso

$$x(y) \in \underset{\color{red}{x\in\mathcal{G}_J}}{\operatorname{argmin}} \frac{1}{2}\|y-\Phi x\|_2^2 + \lambda\|D^* x\|_1$$

How to implicit a solution ?

$$\underbrace{\Phi^*\Phi}_{\text{Non-inversible}} x(y) = \Phi^* y - \lambda D_I s - \lambda D_J \sigma$$

A Solution of Lasso

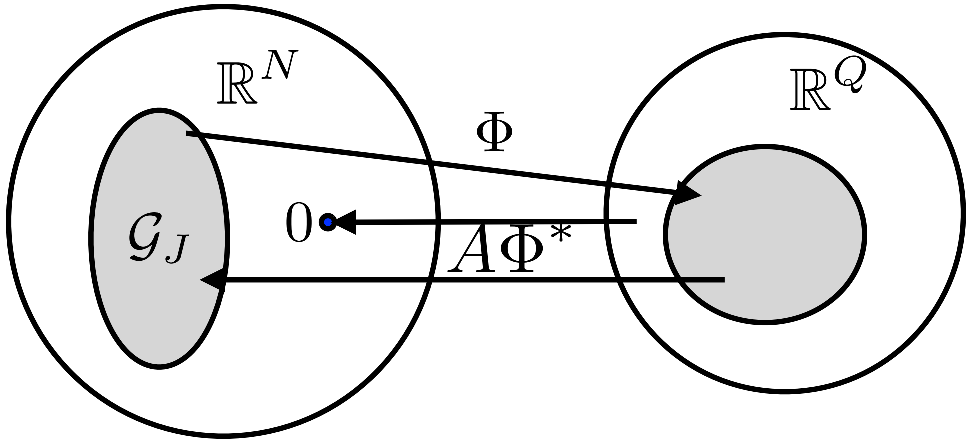
$$x(y) \in \underset{\color{red}{x \in \mathcal{G}_J}}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

How to implicit a solution ?

$$\underbrace{\Phi^* \Phi}_{\text{Non-inversible}} x(y) = \Phi^* y - \lambda D_I s - \lambda D_J \sigma$$

$A\Phi^*$ inverse of Φ on $\mathcal{G}_J$

$$A\Phi^* : \left\{ \begin{array}{l} A\Phi^*{}_{|\Phi(\mathcal{G}_J)} = \left(\Phi_{|\mathcal{G}_J}\right)^{-1} \\ A\Phi^*{}_{|\Phi(\mathcal{G}_J)^\perp} = 0 \end{array} \right.$$



A Solution of Lasso

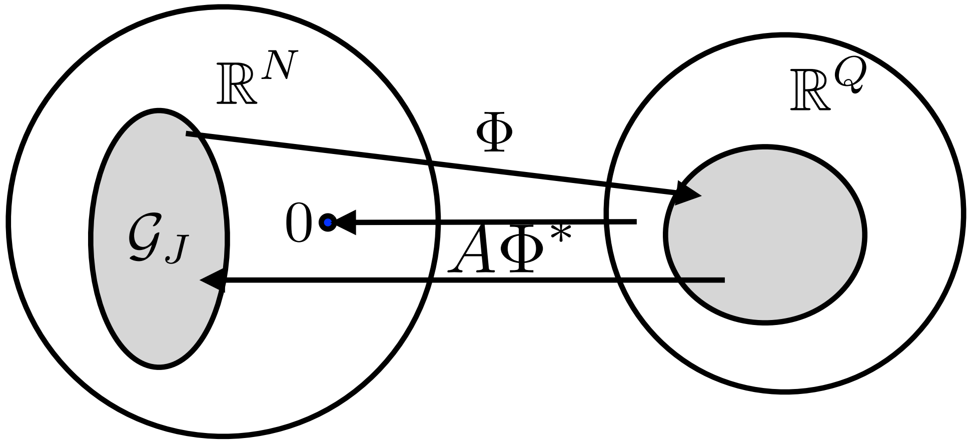
$$x(y) \in \underset{x \in \mathcal{G}_J}{\text{argmin}} \; \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

How to implicit a solution ?

$$\underbrace{\Phi^* \Phi}_{\text{Non-inversible}} x(y) = \Phi^* y - \lambda D_I s - \lambda D_J \sigma$$

$A\Phi^*$ inverse of Φ on $\mathcal{G}_J$

$$A\Phi^* : \begin{cases} A\Phi^*{}_{|\Phi(\mathcal{G}_J)} = \left(\Phi_{|\mathcal{G}_J}\right)^{-1} \\ A\Phi^*{}_{|\Phi(\mathcal{G}_J)^\perp} = 0 \end{cases}$$



$$x(y) = A\Phi^* y - \lambda A D_I s - \lambda A D_J \sigma$$

A Solution of Lasso

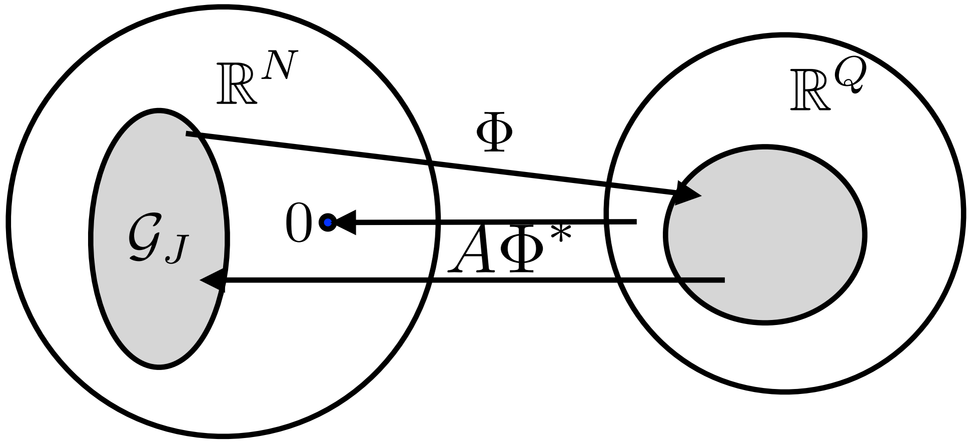
$$x(y) \in \underset{x \in \mathcal{G}_J}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda \|D^* x\|_1$$

How to implicit a solution ?

$$\underbrace{\Phi^* \Phi}_{\text{Non-inversible}} x(y) = \Phi^* y - \lambda D_I s - \lambda D_J \sigma$$

$A\Phi^*$ inverse of Φ on $\mathcal{G}_J$

$$A\Phi^* : \begin{cases} A\Phi^*{}_{|\Phi(\mathcal{G}_J)} = \left(\Phi_{|\mathcal{G}_J}\right)^{-1} \\ A\Phi^*{}_{|\Phi(\mathcal{G}_J)^\perp} = 0 \end{cases}$$



$$x(y) = A\Phi^* y - \lambda A D_I s - \underbrace{\lambda A D_J \sigma}_{= 0} \quad (x(y) \in \mathcal{G}_J)$$

Transition Space

$$\mathcal{H} = \left\{ y \in \mathbb{R}^Q \setminus \exists x \in \mathbb{R}^N : \min_{\sigma \in \Sigma_y(x)} \|\sigma\|_\infty = 1 \right\}$$

Transition Space

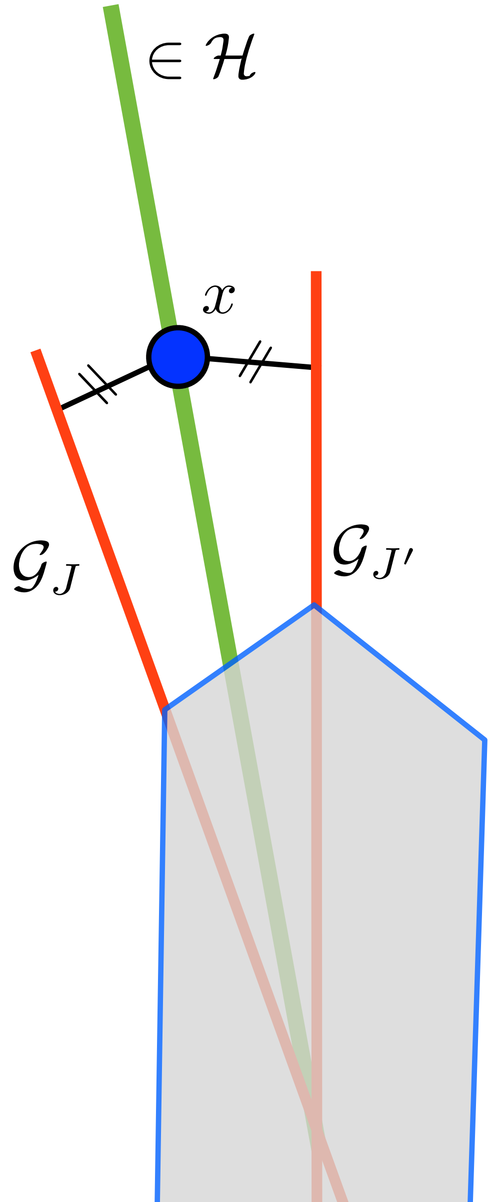
$$\mathcal{H} = \left\{ y \in \mathbb{R}^Q \setminus \exists x \in \mathbb{R}^N : \min_{\sigma \in \Sigma_y(x)} \|\sigma\|_\infty = 1 \right\}$$

$\mathcal{H}$: first order conditions saturation → “jump” from $\mathcal{G}_J$ to $\mathcal{G}_{J'}$

Transition Space

$$\mathcal{H} = \left\{ y \in \mathbb{R}^Q \setminus \exists x \in \mathbb{R}^N : \min_{\sigma \in \Sigma_y(x)} \|\sigma\|_\infty = 1 \right\}$$

$\mathcal{H}$: first order conditions saturation → "jump" from $\mathcal{G}_J$ to $\mathcal{G}_{J'}$



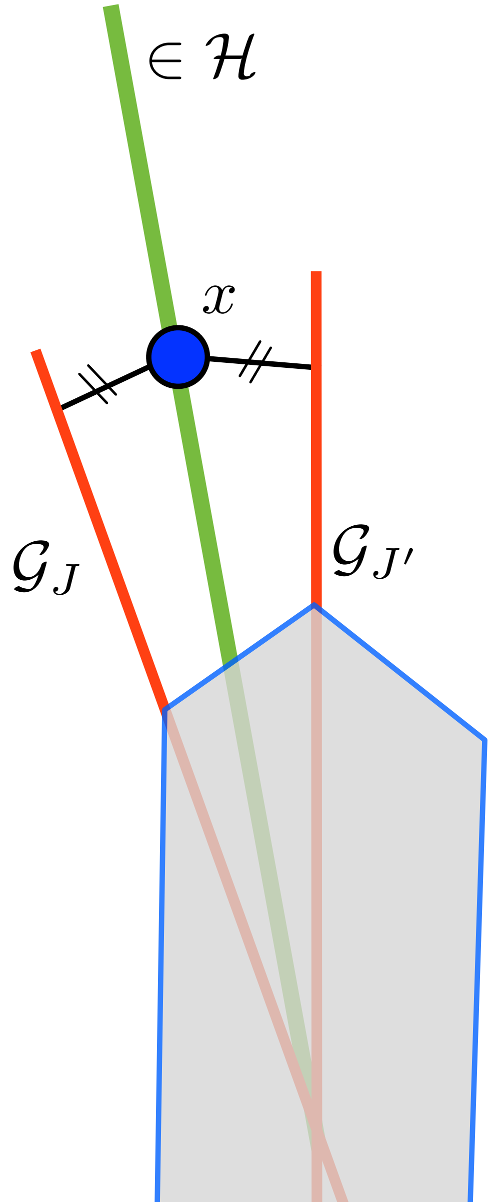
Transition Space

$$\mathcal{H} = \left\{ y \in \mathbb{R}^Q \setminus \exists x \in \mathbb{R}^N : \min_{\sigma \in \Sigma_y(x)} \|\sigma\|_\infty = 1 \right\}$$

$\mathcal{H}$: first order conditions saturation → "jump" from $\mathcal{G}_J$ to $\mathcal{G}_{J'}$

Open question

Smallest union of subspace containing $\mathcal{H}$?



End of the Proof

— Consider $x(y)$ as a mapping of observations $\bar{y} \mapsto x(\bar{y})$

$$\bar{y} \mapsto x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

End of the Proof

— Consider $x(y)$ as a mapping of observations $\bar{y} \mapsto x(\bar{y})$

$$\bar{y} \mapsto x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

— Fix $\bar{y}$ close enough to have $\text{sign}(D^* x(y)) = \text{sign}(D^* x(\bar{y}))$

Sign stability

End of the Proof

- Consider $x(y)$ as a mapping of observations $\bar{y} \mapsto x(\bar{y})$

$$\bar{y} \mapsto x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

- Fix $\bar{y}$ close enough to have $\operatorname{sign}(D^* x(y)) = \operatorname{sign}(D^* x(\bar{y}))$

Sign stability

- Check that $x(\bar{y})$ is indeed solution of $\mathcal{P}(\bar{y}, \lambda)$

Use of first order conditions

Remember !

$$x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

Remember !

$$x(\bar{y}) = A\Phi^*\bar{y} - \lambda A D_I s$$

Inverse of Φ on $\mathcal{G}_J$

Remember !

$$x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

Inverse of Φ on $\mathcal{G}_J$

$y \mapsto x^\star(y)$ is :

— continuous

Remember !

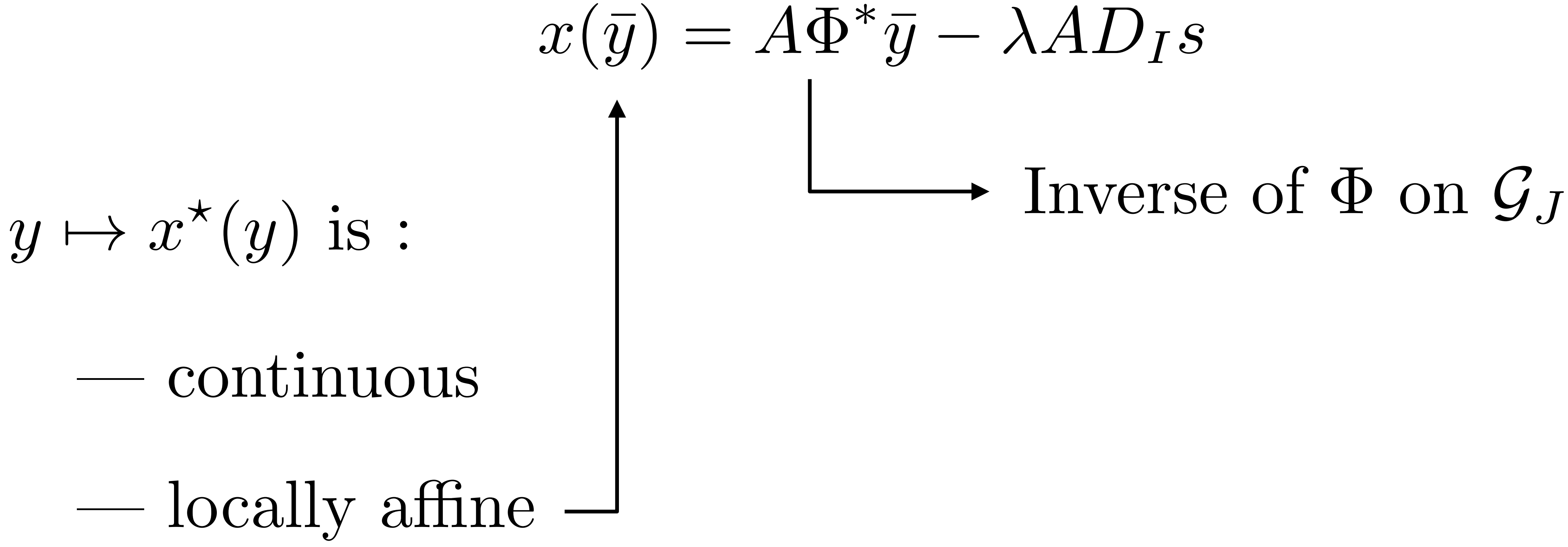
$$x(\bar{y}) = A\Phi^*\bar{y} - \lambda A D_I s$$

Inverse of Φ on $\mathcal{G}_J$

$y \mapsto x^\star(y)$ is :

- — continuous
- — locally affine

Remember !



$$x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

Inverse of Φ on $\mathcal{G}_J$

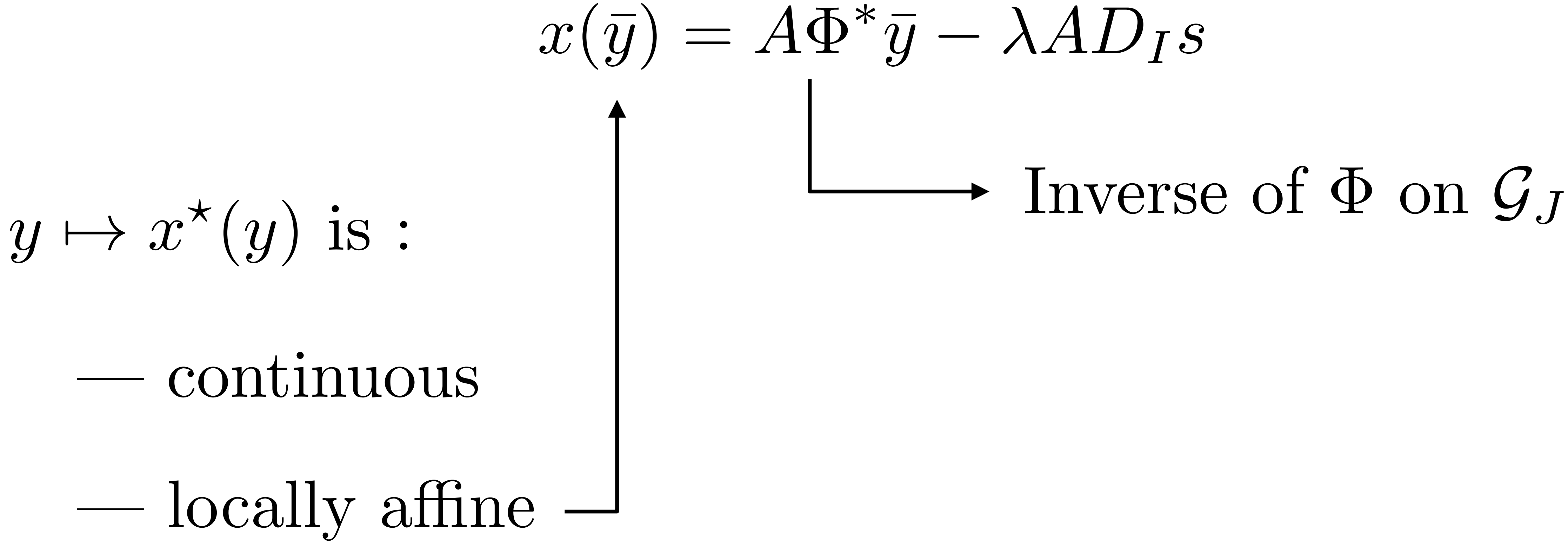
$y \mapsto x^\star(y)$ is :

— continuous

— locally affine

Property given by sign stability

Remember !



$$x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

Inverse of Φ on $\mathcal{G}_J$

$y \mapsto x^\star(y)$ is :

— continuous

— locally affine

Property given by sign stability

Useful for :

— Robustness study

— SURE denoising risk estimation

— Inverse problem on Φx

Overview

- Analysis vs. Synthesis Regularization
- Local Parameterization of Analysis Regularization
- Identifiability and Stability
- Numerical Evaluation
- Perspectives

Identifiability

Identifiability: x_0 unique solution of $\mathcal{P}(\Phi x_0, 0)$

$$\{x_0\} \overset{?}{=} \underset{\Phi x = \Phi x_0}{\operatorname{argmin}} \|D^* x\|_1$$

Identifiability

Identifiability: x_0 unique solution of $\mathcal{P}(\Phi x_0, 0)$

$$\{x_0\} \overset{?}{=} \underset{\Phi x = \Phi x_0}{\text{argmin}} \, \|D^* x\|_1$$

Strategy: $\mathcal{P}(y, \lambda)$ is almost $\mathcal{P}(y, 0)$ for small values of λ

Identifiability

Identifiability: x_0 unique solution of $\mathcal{P}(\Phi x_0, 0)$

$$\{x_0\} \overset{?}{=} \underset{\Phi x = \Phi x_0}{\text{argmin}} \; \|D^* x\|_1$$

Strategy: $\mathcal{P}(y, \lambda)$ is almost $\mathcal{P}(y, 0)$ for small values of λ

Assumption: $\mathcal{G}_J$ must be stable for small values of λ

$\longrightarrow$ Restrictive condition !

But gives a stability results for small noise.

Noiseless and Sign Criterion

Algebraic criterion *on sign vector*

$$\Omega = D_J^+(\Phi^*\Phi A - \mathrm{Id})D_I$$

$$F(s) = \min_{w \in \mathrm{Ker}\, D_J} \|\Omega s - w\|_\infty$$

Noiseless and Sign Criterion

Algebraic criterion *on sign vector*

$$\Omega = D_J^+(\Phi^* \Phi A - \mathrm{Id}) D_I$$

$$F(s) = \min_{w \in \mathrm{Ker}\, D_J} \|\Omega s - w\|_\infty$$

(convex → computable)

Noiseless and Sign Criterion

Algebraic criterion *on sign vector*

$$\Omega = D_J^+(\Phi^*\Phi A - \mathrm{Id})D_I$$

$$F(s) = \min_{w \in \mathrm{Ker}\, D_J} \|\Omega s - w\|_\infty$$

(convex → computable)

Theorem 2

Let $x_0 \in \mathbb{R}^N$ be a fixed vector, and $J = I^c$ where $I = I(D^* x_0)$.
Suppose that $\mathrm{Ker}\,\Phi \cap \mathcal{G}_J = \{0\}$.
If $F\left(\mathrm{sign}\left(D_I^* x_0\right)\right) < 1$ then x_0 is identifiable.

Noiseless and Sign Criterion

Algebraic criterion *on sign vector*

$$\Omega = D_J^+(\Phi^*\Phi A - \mathrm{Id})D_I$$

$$F(s) = \min_{w \in \mathrm{Ker}\, D_J} \|\Omega s - w\|_\infty$$

(convex → computable)

Theorem 2

Let $x_0 \in \mathbb{R}^N$ be a fixed vector, and $J = I^c$ where $I = I(D^* x_0)$.
Suppose that $\mathrm{Ker}\,\Phi \cap \mathcal{G}_J = \{0\}$.
If $F\left(\mathrm{sign}\left(D_I^* x_0\right)\right) < 1$ then x_0 is identifiable.

Specializes to Fuchs results for synthesis ($D = \mathrm{Id}$)

Nam et al. Results

“Cosparse” model [Nam 2011]

Only other work on analysis recovery

Nam et al. Results

"Cosparse" model [Nam 2011]

Only other work on analysis recovery

$\Gamma = (MD_J)^+ MD_I$ M^* orthonormal basis of $\operatorname{Ker}\Phi$

$$G(s) = \|\Gamma s\|_\infty$$

Nam et al. Results

“Cosparse” model [Nam 2011]

Only other work on analysis recovery

$\Gamma = (MD_J)^+ MD_I$ M^* orthonormal basis of $\operatorname{Ker}\Phi$

$$G(s) = \|\Gamma s\|_\infty$$

Theorem

Let $x_0 \in \mathbb{R}^N$ be a fixed vector, and $J = I^c$ where $I = I(D^* x_0)$. Suppose that $\operatorname{Ker}\Phi \cap \mathcal{G}_J = \{0\}$.

If $G\left(\operatorname{sign}\left(D_I^* x_0\right)\right) < 1$ then x_0 is identifiable.

Nam et al. Results

“Cosparse” model [Nam 2011]

Only other work on analysis recovery

$\Gamma = (MD_J)^+ MD_I$ M^* orthonormal basis of $\operatorname{Ker}\Phi$

$$G(s) = \|\Gamma s\|_\infty$$

Theorem

Let $x_0 \in \mathbb{R}^N$ be a fixed vector, and $J = I^c$ where $I = I(D^* x_0)$. Suppose that $\operatorname{Ker}\Phi \cap \mathcal{G}_J = \{0\}$.

If $G\left(\operatorname{sign}\left(D_I^* x_0\right)\right) < 1$ then x_0 is identifiable.

More intrinsic criterion

$\longrightarrow$ *But* no noise robustness, even for small ones

Sketch of the Proof

Idea: Study $\mathcal{P}(y, \lambda)$ for $\lambda \approx 0$

$$x_\lambda(\Phi x_0) = A\Phi^*\Phi x_0 - \lambda A D_I s$$

Sketch of the Proof

Idea: Study $\mathcal{P}(y,\lambda)$ for $\lambda \approx 0$

$$x_\lambda(\Phi x_0) = A\Phi^*\Phi x_0 - \lambda A D_I s$$

λ small enough to have $\text{sign}(D^* x_\lambda(\Phi x_0)) = \text{sign}(D^* x_0)$

Sketch of the Proof

Idea: Study $\mathcal{P}(y,\lambda)$ for $\lambda \approx 0$

$$x_\lambda(\Phi x_0) = A\Phi^*\Phi x_0 - \lambda A D_I s$$

λ small enough to have $\text{sign}(D^* x_\lambda(\Phi x_0)) = \text{sign}(D^* x_0)$

$\lim_{\lambda\to 0} x_\lambda(\Phi x_0) = A\Phi^*\Phi x_0 = x_0$

Sketch of the Proof

Idea: Study $\mathcal{P}(y, \lambda)$ for $\lambda \approx 0$

$$x_\lambda(\Phi x_0) = A\Phi^* \Phi x_0 - \lambda A D_I s$$

λ small enough to have $\text{sign}(D^* x_\lambda(\Phi x_0)) = \text{sign}(D^* x_0)$

$$\lim_{\lambda \to 0} x_\lambda(\Phi x_0) = A\Phi^* \Phi x_0 = x_0$$

$x_\lambda(\Phi x_0)$ solution of $\mathcal{P}(\lambda)$ and $x_\lambda(\Phi x_0) \xrightarrow[\lambda \to 0]{} x_0(\Phi x_0)$

$$\Downarrow$$

$$x_0(\Phi x_0) \text{ solution of } \mathcal{P}(0)$$

Sketch of the Proof

Idea: Study $\mathcal{P}(y, \lambda)$ for $\lambda \approx 0$

$$x_\lambda(\Phi x_0) = A\Phi^*\Phi x_0 - \lambda A D_I s$$

λ small enough to have $\text{sign}(D^* x_\lambda(\Phi x_0)) = \text{sign}(D^* x_0)$

$$\lim_{\lambda \to 0} x_\lambda(\Phi x_0) = A\Phi^*\Phi x_0 = x_0$$

$$x_\lambda(\Phi x_0) \text{ solution of } \mathcal{P}(\lambda) \text{ and } x_\lambda(\Phi x_0) \xrightarrow[\lambda \to 0]{} x_0(\Phi x_0)$$

$$\Downarrow$$

$$x_0(\Phi x_0) \text{ solution of } \mathcal{P}(0)$$

$$F(\text{sign}(D^* x_\lambda(\Phi x_0)) < 1 \Rightarrow x_\lambda(\Phi x_0) \text{ unique solution}$$

Small Noise Recovery

Suppose we observe $y = \Phi x_0 + w$

$$\text{Does } \underset{\Phi x = y}{\operatorname{argmin}} \|D^* x\|_1 \text{ recovers } x_0 + A\Phi^* w \text{ ?}$$

Small Noise Recovery

Suppose we observe $y = \Phi x_0 + w$

$$\text{Does } \underset{\Phi x = y}{\text{argmin}}\, \|D^* x\|_1 \text{ recovers } x_0 + A\Phi^* w \text{ ?}$$

Generalization of Theorem 2 : Yes, if $\|w\|$ small enough

$$\text{Condition : } \operatorname{sign}(D^* x_\lambda(y)) = \operatorname{sign}(D^* x_0)$$

Small Noise Recovery

Suppose we observe $y = \Phi x_0 + w$

$$\text{Does } \underset{\Phi x = y}{\operatorname{argmin}} \|D^* x\|_1 \text{ recovers } x_0 + A\Phi^* w \text{ ?}$$

Generalization of Theorem 2 : Yes, if $\|w\|$ small enough

$$\text{Condition : } \operatorname{sign}(D^* x_\lambda(y)) = \operatorname{sign}(D^* x_0)$$

$F(\operatorname{sign}(D^* x_0)) < 1$ gives

- *identifiability*
- *small noise robustness*

Small Noise Recovery

Suppose we observe $y = \Phi x_0 + w$

$$\text{Does } \underset{\Phi x = y}{\text{argmin}} \, \|D^* x\|_1 \text{ recovers } x_0 + A\Phi^* w \text{ ?}$$

Generalization of Theorem 2 : Yes, if $\|w\|$ small enough

$$\text{Condition : } \operatorname{sign}(D^* x_\lambda(y)) = \operatorname{sign}(D^* x_0)$$

$F(\operatorname{sign}(D^* x_0)) < 1$ gives

- *identifiability*
- *small noise robustness*

Question: And for an arbitrary noise ?

Noisy and Support Criterion

Settings: $y = \Phi x_0 + w$, with w bounded noise.

Noisy and Support Criterion

Settings: $y = \Phi x_0 + w$, with w bounded noise.

identifiability of vector → identifiability of support

Noisy and Support Criterion

Settings: $y = \Phi x_0 + w$, with w bounded noise.

identifiability of vector → identifiability of support

$$\text{ARC}(I) = \max_{x \in \mathcal{G}_J} F(\text{sign}(D_I^* x))$$

Noisy and Support Criterion

Settings: $y = \Phi x_0 + w$, with w bounded noise.

identifiability of vector → identifiability of support

$$\mathrm{ARC}(I) = \max_{x \in \mathcal{G}_J} F(\mathrm{sign}(D_I^* x))$$

Theorem 3

Suppose $\mathrm{ARC}(I) < 1$ and $\lambda > K \dfrac{\|w\|}{1 - \mathrm{ARC}(I)}$ then $x_\lambda(y)$ is the unique solution of $\mathcal{P}(y, \lambda)$ and

$$\|x_\lambda(\bar{y}) - x_0\| = O(\lambda)$$

Remember !

$F(s) = \min_{w \in \text{Ker}\, D_J} \|\Omega s - w\|_\infty$	$\text{ARC}(I) = \max_{x \in \mathcal{G}_J} F(\text{sign}(D_I^* x))$
sign	support
Noiseless	Noisy
Vector identifiability	Support identifiability

From Theory to Numerics

We give a sufficient condition for identifiability.

From Theory to Numerics

We give a sufficient condition for identifiability.

How far are we from a necessary condition ?

Overview

- Analysis vs. Synthesis Regularization
- Local Parameterization of Analysis Regularization
- Identifiability and Stability
- Numerical Evaluation
- Perspectives

Proximal Operator

f l.s.c convex function from C convex of an Hilbert $\mathbb{H}$ in $\mathbb{R}$.

Proximal Operator

f l.s.c convex function from C convex of an Hilbert $\mathbb{H}$ in $\mathbb{R}$.

Proximal operator

$$\operatorname{prox}_f(x) = \underset{u \in \mathbb{R}^N}{\operatorname{argmin}} \left\{ f(u) + \frac{1}{2}\|u - x\|_2^2 \right\}$$

Proximal Operator

f l.s.c convex function from C convex of an Hilbert $\mathbb{H}$ in $\mathbb{R}$.

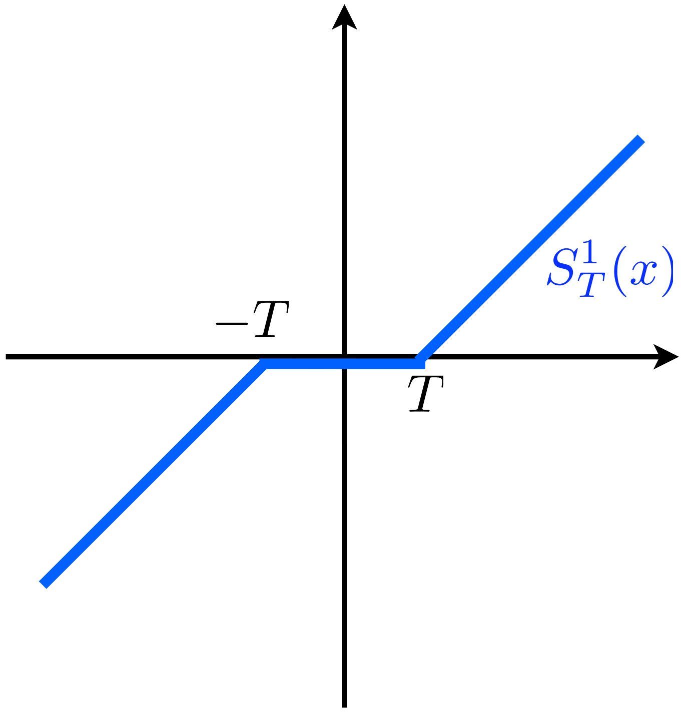
Proximal operator

$$\operatorname{prox}_f(x) = \underset{u \in \mathbb{R}^N}{\operatorname{argmin}} \left\{ f(u) + \frac{1}{2}\|u - x\|_2^2 \right\}$$

Fundamental examples:

$$\operatorname{prox}_{\|\cdot\|_1} = S_T^1.$$

$$\operatorname{prox}_{i_C} = P_C$$



How to Solve These Regularizations ?

Primal-dual schemes

$$\min_{x \in \mathbb{R}^N} \mathcal{L}(K(x)) \quad \text{where} \quad \left\{ \begin{array}{l} \mathcal{L}(g,u) = \frac{1}{2}\|y-g\|^2 + \lambda \|u\|_1 \\ K(x) = (\Phi x, D^* x) \end{array} \right.$$

How to Solve These Regularizations ?

Primal-dual schemes

$$\min_{x \in \mathbb{R}^N} \mathcal{L}(K(x)) \quad \text{where} \quad \left\{ \begin{array}{l} \mathcal{L}(g,u) = \frac{1}{2}\|y-g\|^2 + \lambda\|u\|_1 \\ K(x) = (\Phi x, D^* x) \end{array} \right.$$

Alternating Direction Method of Multipliers [Chambolle, Pock]

$$\begin{aligned} u_n &= \operatorname{prox}_{\sigma \mathcal{L}^*}(u_{n-1} + \sigma K(z_{n-1})) \\ x_n &= \operatorname{prox}_{\tau G}(x_{n-1} - \tau K^*(u_n)) \\ z_n &= x_n + \theta(x_n - x_{n-1}) \end{aligned}$$

How to Solve These Regularizations ?

Primal-dual schemes

$$\min_{x \in \mathbb{R}^N} \mathcal{L}(K(x)) \quad \text{where} \quad \begin{cases} \mathcal{L}(g,u) = \frac{1}{2}\|y-g\|^2 + \lambda\|u\|_1 \\ K(x) = (\Phi x, D^* x) \end{cases}$$

Alternating Direction Method of Multipliers [Chambolle, Pock]

$$u_n = \text{prox}_{\sigma\mathcal{L}^*}(u_{n-1} + \sigma K(z_{n-1}))$$

$$x_n = \text{prox}_{\tau G}(x_{n-1} - \tau K^*(u_n))$$

$$z_n = x_n + \theta(x_n - x_{n-1})$$

For $\mathcal{P}(y,0)$, $\|y-g\|^2 \rightarrow i_{\{y\}}$

Computing Criterions

Unconstrained formulation

$$\begin{tikzcd} F(s) = \min\limits_{w\in\mathbb{R}^N} & \|\Omega s - w\|_\infty \arrow[d] & + & i_{\mathcal{D}}(w) \arrow[d] \\ & \mathrm{P}_{\|\cdot\|_1=1} & \text{Prox} & \mathrm{P}_{\mathcal{D}} \end{tikzcd}$$

Computing Criterions

Unconstrained formulation

$$\begin{tikzcd} F(s) = \min_{w\in\mathbb{R}^N} \|\Omega s - w\|_\infty \arrow[dd] & + & i_{\mathcal{D}}(w) \arrow[dd] \\ & \text{Prox} & \\ \mathrm{P}_{\|\cdot\|_1=1} & & \mathrm{P}_{\mathcal{D}} \end{tikzcd}$$

ARC difficult to compute (non-convex)

$$\begin{array}{rrll} & \mathrm{ARC}(I) & = \max_{x\in\mathcal{G}_J} F(\mathrm{sign}(D_I^* x)) & \text{non-convex} \\ \leqslant & \mathrm{wARC}(I) & = \max_{s\in\{-1,1\}^{|J|}} F(s) & \text{non-convex} \\ \leqslant & \mathrm{oARC}(I) & = \|\Omega\|_{\infty\to\infty} & \text{easy} \end{array}$$

More on Signal Models

Signal model : *"Union of subspace"*

$$\Theta = \bigcup_{k \in \{1 \ldots P\}} \Theta_k \quad \text{where} \quad \Theta_k = \{\mathcal{G}_J \setminus \dim \mathcal{G}_J = k\}$$

More on Signal Models

Signal model : *"Union of subspace"*

$$\Theta = \bigcup_{k \in \{1 \ldots P\}} \Theta_k \quad \text{where} \quad \Theta_k = \{\mathcal{G}_J \setminus \dim \mathcal{G}_J = k\}$$

Sparsity $\|D^* x_0\|_0$ is not a good parameter

More on Signal Models

Signal model : *"Union of subspace"*

$$\Theta = \bigcup_{k\in\{1\ldots P\}} \Theta_k \quad \text{where} \quad \Theta_k = \{\mathcal{G}_J \setminus \dim \mathcal{G}_J = k\}$$

Sparsity $\|D^* x_0\|_0$ is not a good parameter

D redundant Gaussian i.i.d matrix $N \times P$

$$\|D^* x_0\|_0 < P - N \Rightarrow x_0 = 0 \;!$$

More on Signal Models

Signal model : *"Union of subspace"*

$$\Theta = \bigcup_{k \in \{1 \ldots P\}} \Theta_k \quad \text{where} \quad \Theta_k = \{\mathcal{G}_J \setminus \dim \mathcal{G}_J = k\}$$

Sparsity $\|D^* x_0\|_0$ is not a good parameter

D redundant Gaussian i.i.d matrix $N \times P$

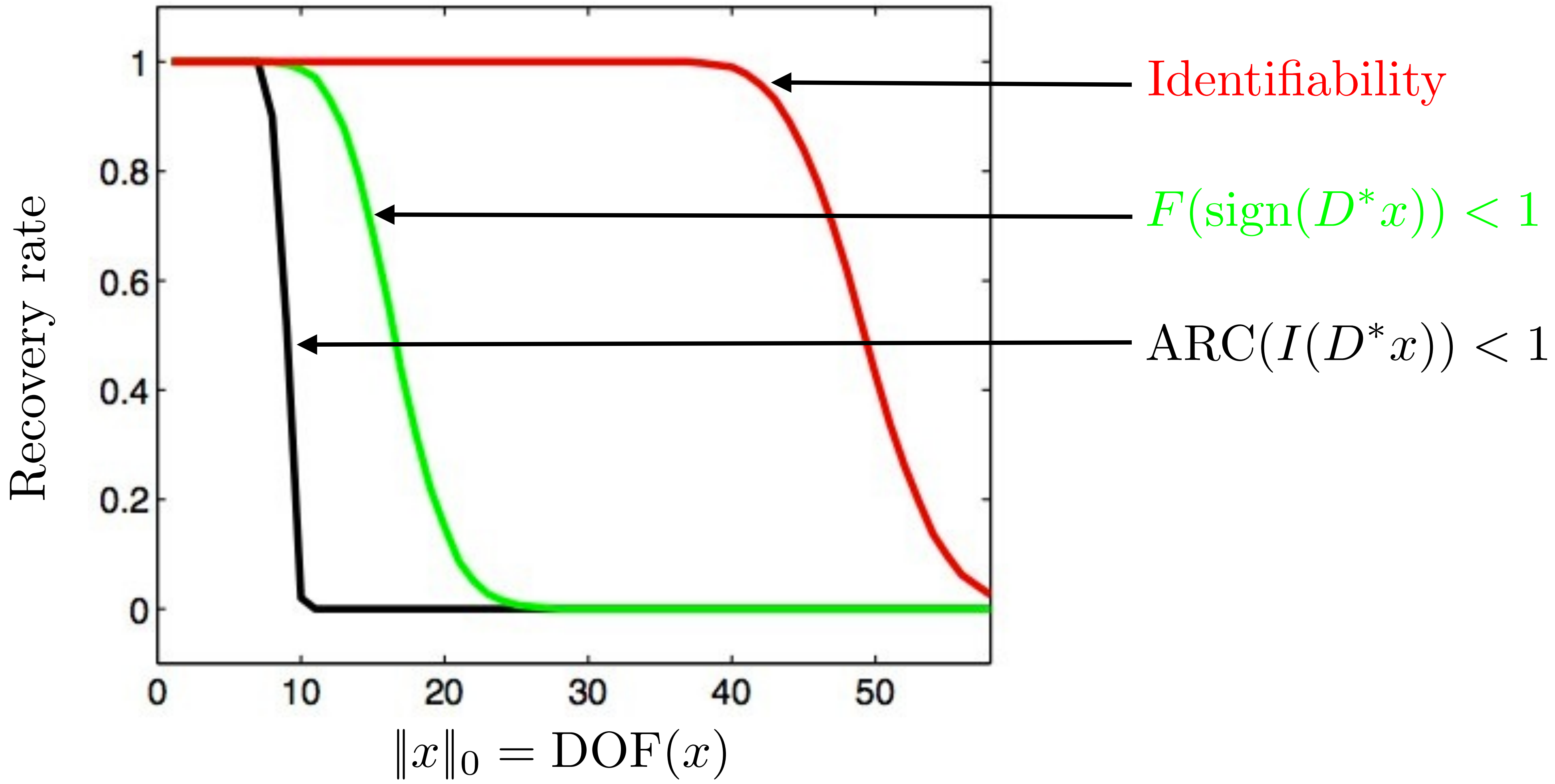
$$\|D^* x_0\|_0 < P - N \Rightarrow x_0 = 0 \;!$$

Good one : $\mathrm{DOF}(x) = \dim \mathcal{G}_J$

Random Settings

1) Synthesis results

Compressed sensing : $Q \ll N$



Credit to C. Dossal

Random Settings

2) Analysis results $\qquad D, \Phi$ Gaussian i.i.d random matrices

Random Settings

2) Analysis results $\qquad$ D, Φ Gaussian i.i.d random matrices

Many dependancies between columns

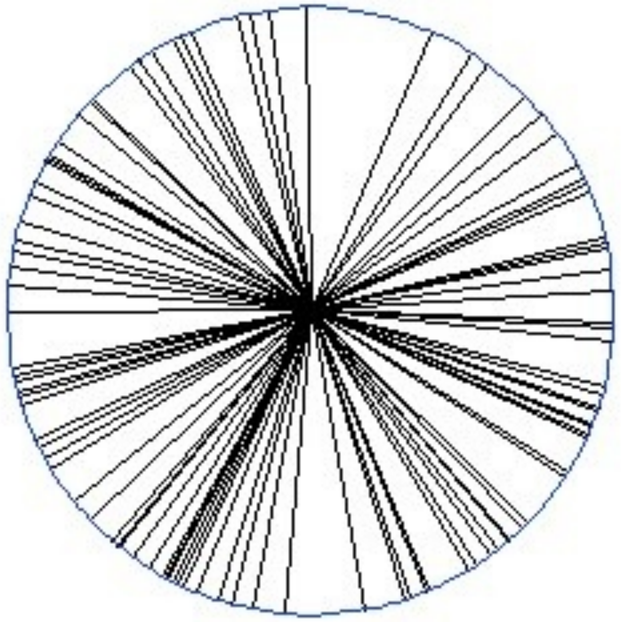
$\longrightarrow$ Strong unstability

Random Settings

2) Analysis results D, Φ Gaussian i.i.d random matrices

Many dependancies between columns

$\longrightarrow$ Strong unstability



Close to ℓ_2 ball !

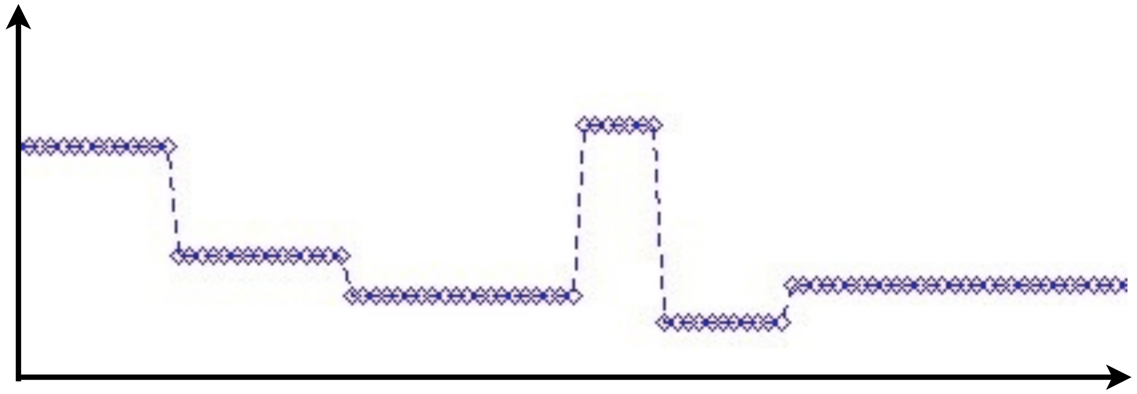
Limits : TV Instability

$D^* = \nabla, \Phi = \mathrm{Id}$

Limits : TV Instability

$D^* = \nabla, \Phi = \mathrm{Id}$

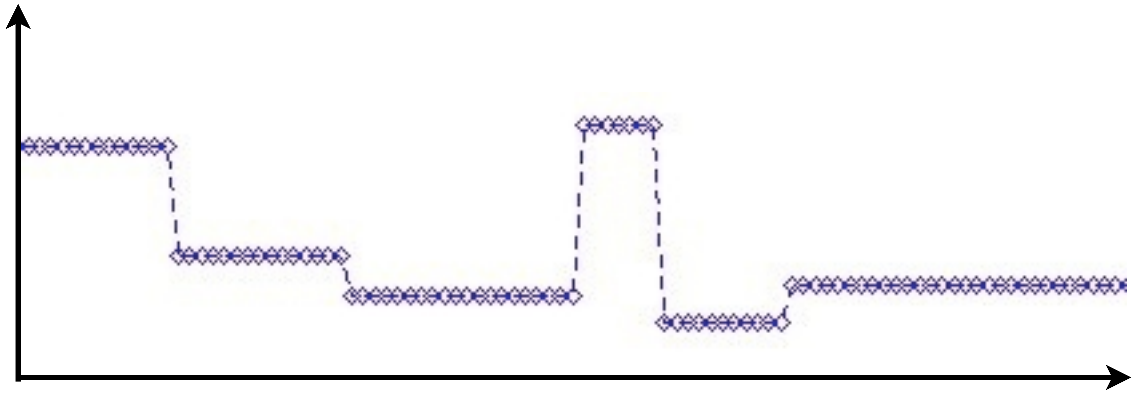
Θ_k : piecewise constant signals with $k-1$ step.



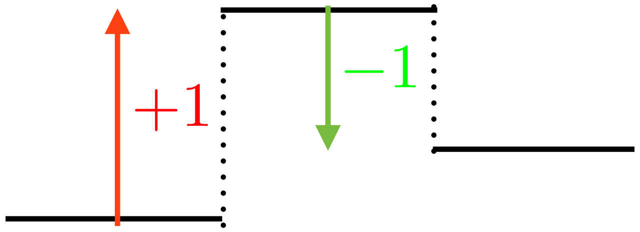
Limits : TV Instability

$$D^* = \nabla, \Phi = \mathrm{Id}$$

Θ_k : piecewise constant signals with $k-1$ step.



“*Box*”

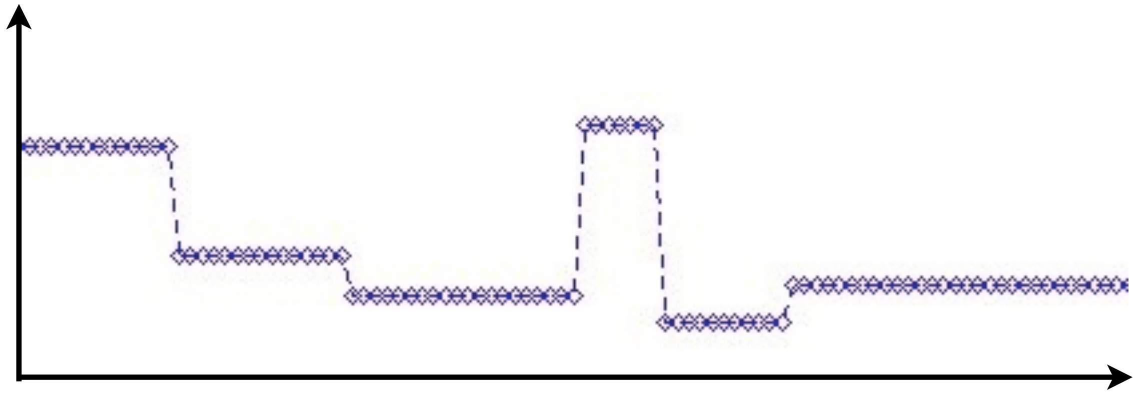


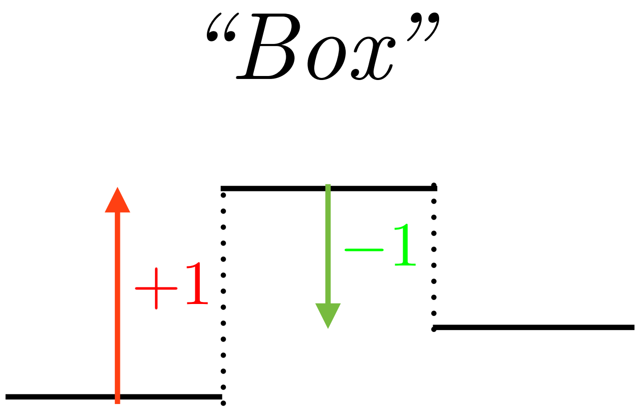
$$F(s) = 1 - \varepsilon$$

Limits : TV Instability

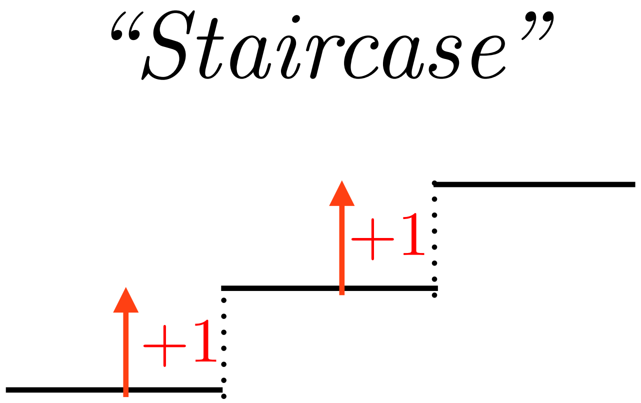
$D^* = \nabla, \Phi = \mathrm{Id}$

Θ_k : piecewise constant signals with $k-1$ step.





$$F(s) = 1 - \varepsilon$$



$$F(s) = 1$$

No noise stability
even for small one

Fused Lasso

$$\underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 \quad \text{subject to} \quad \left\{ \begin{array}{l} \|\nabla x\|_1 \leqslant s_1 \\ \|x\|_1 \leqslant s_2 \end{array} \right.$$

Fused Lasso

$$\underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 \quad \text{subject to} \quad \begin{cases} \|\nabla x\|_1 \leqslant s_1 \\ \|x\|_1 \leqslant s_2 \end{cases}$$

Signal Model: Characteristic functions sum

$$\Theta_2 : x_0 = \mathbf{1}_{[a,b]} + \mathbf{1}_{[c,d]}$$

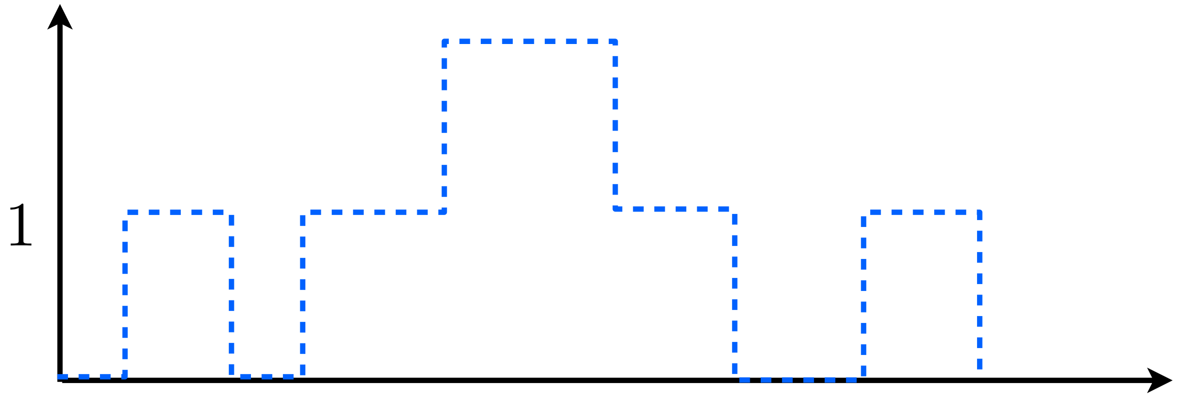
Fused Lasso

$$\underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 \quad \text{subject to} \quad \begin{cases} \|\nabla x\|_1 \leqslant s_1 \\ \|x\|_1 \leqslant s_2 \end{cases}$$

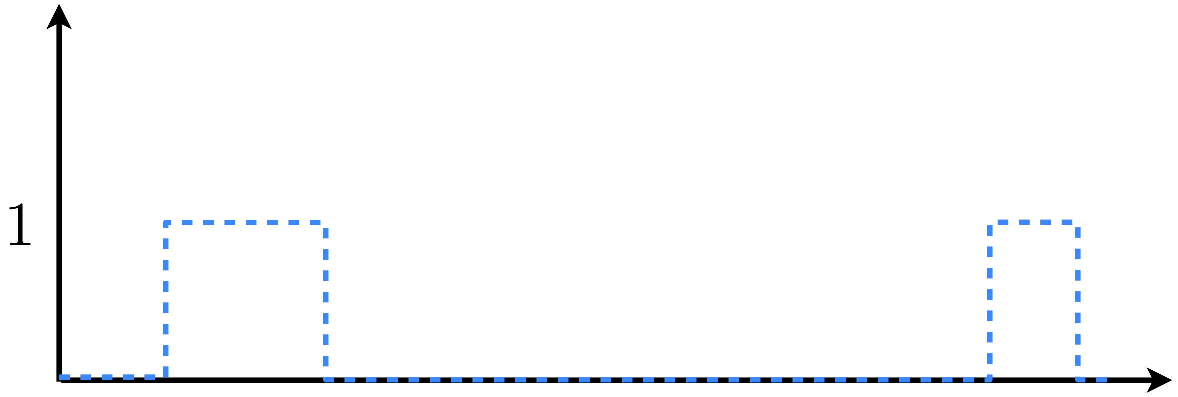
Signal Model: Characteristic functions sum

$$\Theta_2 : x_0 = \mathbf{1}_{[a,b]} + \mathbf{1}_{[c,d]}$$

Overlap



No overlap



Fused Lasso

$$\underset{x \in \mathbb{R}^N}{\text{argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 \quad \text{subject to} \quad \left\{ \begin{array}{l} \|\nabla x\|_1 \leqslant s_1 \\ \|x\|_1 \leqslant s_2 \end{array} \right.$$

$$[a, b] \cap [c, d] \neq \emptyset \Rightarrow F(x_0) \geqslant 1$$

no noise robustness

Fused Lasso

Ω_{DIF}

$\varepsilon \mathrm{Id}$

$$\underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 \quad \text{subject to} \quad \left\{ \begin{array}{l} \|\nabla x\|_1 \leqslant s_1 \\ \|x\|_1 \leqslant s_2 \end{array} \right.$$

$[a, b] \cap [c, d] = \emptyset \Rightarrow$ 2 situations

Fused Lasso

$$\begin{array}{|c|c|}\hline \Omega_{\text{DIF}} & \varepsilon \mathrm{Id} \\ \hline \end{array} \qquad \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 \quad \text{subject to} \quad \left\{ \begin{array}{l} \|\nabla x\|_1 \leqslant s_1 \\ \|x\|_1 \leqslant s_2 \end{array} \right.$$

$[a, b] \cap [c, d] = \emptyset \Rightarrow$ 2 situations

$$|c - b| \leqslant \xi(\varepsilon)$$

$$F(\operatorname{sign}(D^* x_0)) \geqslant 1$$

no noise robustness

Fused Lasso

$$\underset{x\in\mathbb{R}^N}{\text{argmin}}\ \frac{1}{2}\|y-\Phi x\|_2^2 \quad \text{subject to} \quad \left\{ \begin{array}{l} \|\nabla x\|_1 \leqslant s_1 \\ \|x\|_1 \leqslant s_2 \end{array} \right.$$

$$[a,b]\cap[c,d]=\emptyset \Rightarrow \ 2 \text{ situations}$$

$\|c-b\| \leqslant \xi(\varepsilon)$	$\|c-b\| > \xi(\varepsilon)$
$F(\text{sign}(D^* x_0)) \geqslant 1$	$F(\text{sign}(D^* x_0)) = \text{ARC}(I) < 1$
no noise robustness	strong noise robustness

Fused Lasso

Ω_{DIF} $\varepsilon \mathrm{Id}$

$$\underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 \quad \text{subject to} \quad \begin{cases} \|\nabla x\|_1 \leqslant s_1 \\ \|x\|_1 \leqslant s_2 \end{cases}$$

$$[a, b] \cap [c, d] = \emptyset \Rightarrow \ 2 \text{ situations}$$

$|c - b| \leqslant \xi(\varepsilon)$

$F(\mathrm{sign}(D^* x_0)) \geqslant 1$

no noise robustness

$|c - b| > \xi(\varepsilon)$

$F(\mathrm{sign}(D^* x_0)) = \mathrm{ARC}(I) < 1$

strong noise robustness

Haar : similar results

Take-Away Messages

Take-Away Messages

— Analysis regularization *is* robust

Take-Away Messages

— Analysis regularization *is* robust

— Geometry (union of subspaces) : key concept for recovery

Take-Away Messages

— Analysis regularization *is* robust

— Geometry (union of subspaces) : key concept for recovery

— Sparsity is not univoquely defined

Overview

- Analysis vs. Synthesis Regularization
- Local Parameterization of Analysis Regularization
- Identifiability and Stability
- Numerical Evaluation
- Perspectives

What’s Next ?

— Support identifiability with Gaussian, Poisson noise

Deterministic theorem → treat the noise as a random variable

What's Next ?

— Support identifiability with Gaussian, Poisson noise

Deterministic theorem → treat the noise as a random variable

— Total Variation identifiability

Existence of a better criterion to ensure noisy recovery ?

What's Next ?

— Support identifiability with Gaussian, Poisson noise

Deterministic theorem → treat the noise as a random variable

— Total Variation identifiability

Existence of a better criterion to ensure noisy recovery ?

— Continuous model

Work initiated by Chambolle in TV

What's Next ?

— Support identifiability with Gaussian, Poisson noise

Deterministic theorem → treat the noise as a random variable

— Total Variation identifiability

Existence of a better criterion to ensure noisy recovery ?

— Continuous model

Work initiated by Chambolle in TV

— Larger class of priors J

Block sparsity $\| \cdot \|_{p,q}$

What's Next ?

— Support identifiability with Gaussian, Poisson noise

Deterministic theorem → treat the noise as a random variable

— Total Variation identifiability

Existence of a better criterion to ensure noisy recovery ?

— Continuous model

Work initiated by Chambolle in TV

— Larger class of priors J

Block sparsity $\| \cdot \|_{p,q}$

— Real-world recovery results

Almost equal support recovery

Thanks

Joint work with

- Gabriel Peyré (CEREMADE, Dauphine)
- Charles Dossal (IMB, Bordeaux I)
- Jalal Fadili (GREYC, ENSICAEN)

Any questions ?

An Affine Implicit Mapping

$$x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

An Affine Implicit Mapping

$$x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

$$s = \mathrm{sign}(D_I^* x(y))$$

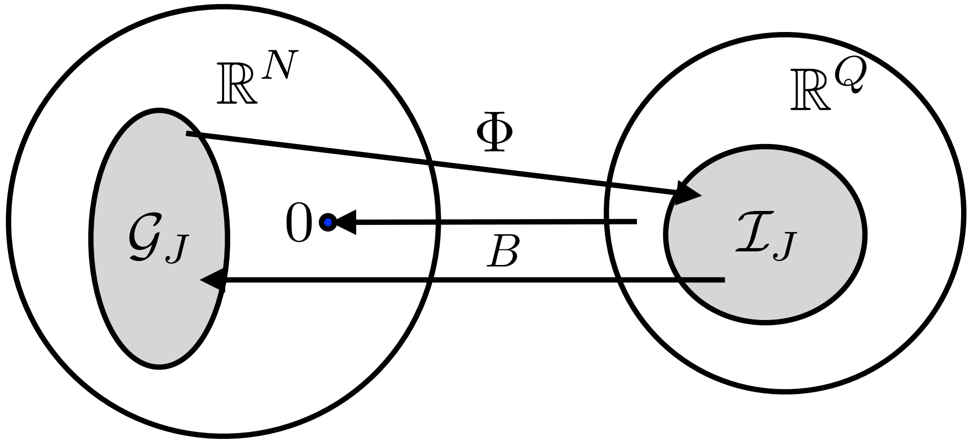
An Affine Implicit Mapping

$$x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

$$s = \mathrm{sign}(D_I^* x(y))$$

$$B = A\Phi^* \text{ inverse of } \Phi \text{ on } \mathcal{G}_J$$

$$\mathcal{G}_J \overset{B}{\cong} \mathcal{I}_J = \Phi(\mathcal{G}_J)$$



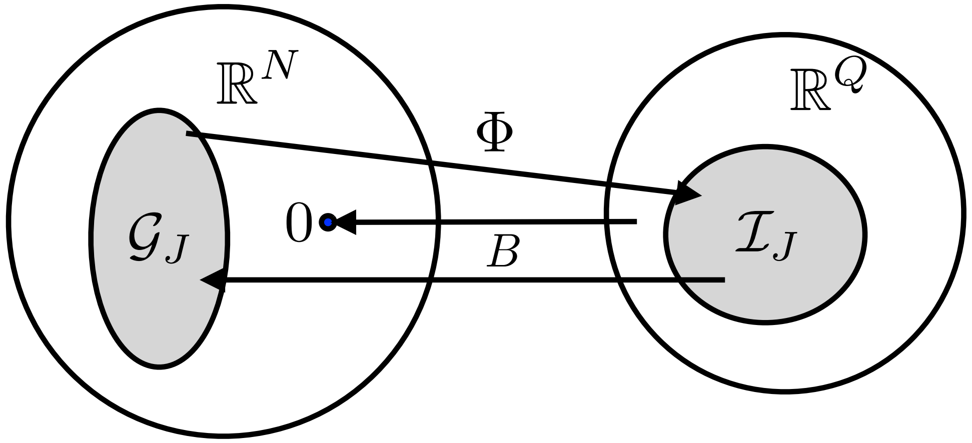
An Affine Implicit Mapping

$$x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

$$s = \text{sign}(D_I^* x(y))$$

$$B = A\Phi^* \text{ inverse of } \Phi \text{ on } \mathcal{G}_J$$

$$\mathcal{G}_J \overset{B}{\cong} \mathcal{I}_J = \Phi(\mathcal{G}_J)$$



$$B : \begin{cases} B_{|\mathcal{I}_J} = \left(\Phi_{|\mathcal{G}_J}\right)^{-1} \\ B_{|\mathcal{I}_J^\perp} = 0 \end{cases}$$

U BON of $\mathcal{G}_J$

$$B = U(U^*\Phi^*\Phi U)^{-1}U^*\Phi^*$$

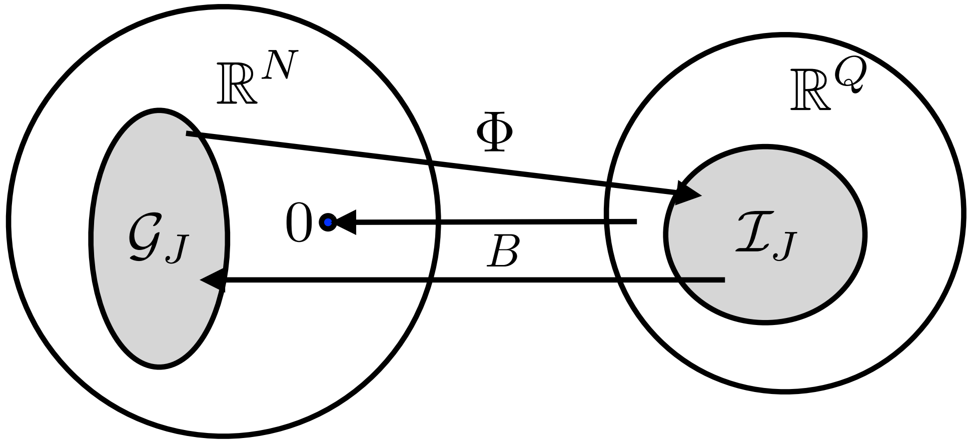
An Affine Implicit Mapping

$$x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

$$s = \text{sign}(D_I^* x(y))$$

$$B = A\Phi^* \text{ inverse of } \Phi \text{ on } \mathcal{G}_J$$

$$\mathcal{G}_J \overset{B}{\cong} \mathcal{I}_J = \Phi(\mathcal{G}_J)$$



$$B : \begin{cases} B_{|\mathcal{I}_J} = \left(\Phi_{|\mathcal{G}_J}\right)^{-1} \\ B_{|\mathcal{I}_J^\perp} = 0 \end{cases}$$

U BON of $\mathcal{G}_J$

$$B = U(U^*\Phi^*\Phi U)^{-1} U^* \Phi^*$$

Efficient computation $\quad y = Bx = \underset{D^* z = 0}{\text{argmin}} \|z - \Phi x\|_2^2$

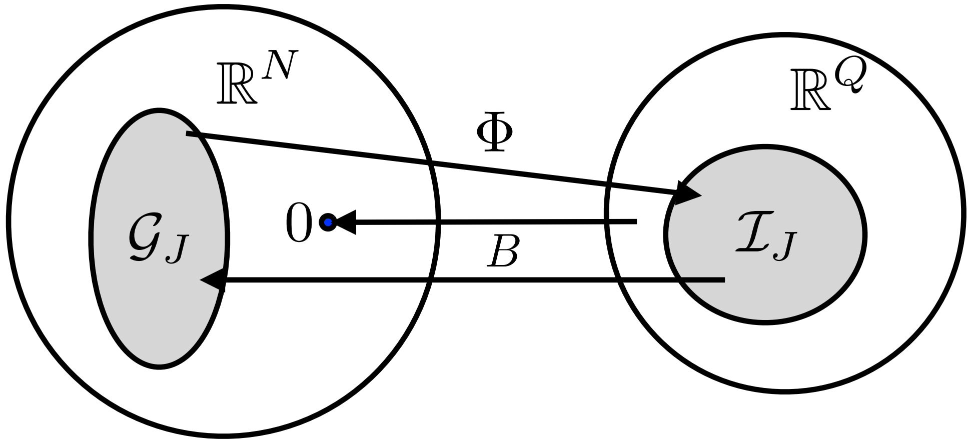
An Affine Implicit Mapping

$$x(\bar{y}) = A\Phi^* \bar{y} - \lambda A D_I s$$

$$s = \text{sign}(D_I^* x(y))$$

$$B = A\Phi^* \text{ inverse of } \Phi \text{ on } \mathcal{G}_J$$

$$\mathcal{G}_J \overset{B}{\cong} \mathcal{I}_J = \Phi(\mathcal{G}_J)$$



$$B : \begin{cases} B_{|\mathcal{I}_J} = \left(\Phi_{|\mathcal{G}_J}\right)^{-1} \\ B_{|\mathcal{I}_J^\perp} = 0 \end{cases}$$

U BON of $\mathcal{G}_J$

$$B = U(U^*\Phi^*\Phi U)^{-1} U^* \Phi^*$$

Efficient computation $\quad y = Bx = \underset{D^* z = 0}{\text{argmin}} \|z - \Phi x\|_2^2$

$$C \begin{pmatrix} y \\ \mu \end{pmatrix} = \begin{pmatrix} \Phi x \\ 0 \end{pmatrix} \quad \text{where} \quad C = \begin{pmatrix} \Phi^*\Phi & D \\ D^* & 0 \end{pmatrix}$$