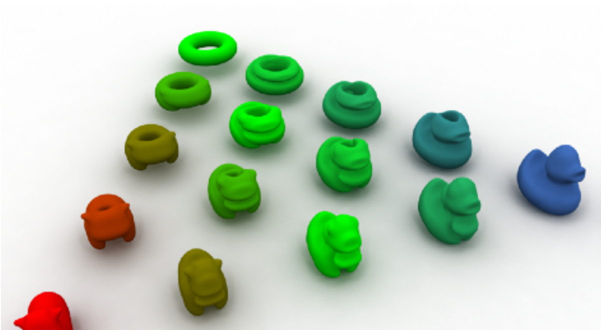
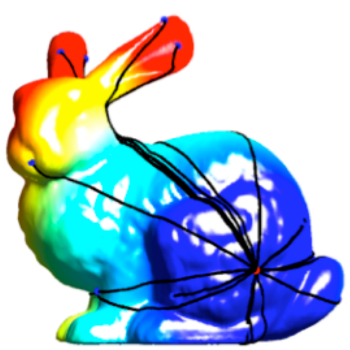
Convex Optimization

Gabriel Peyré

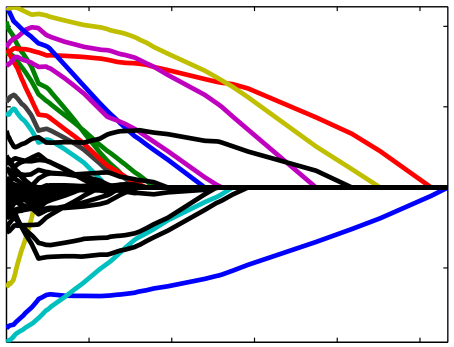
www.numerical-tours.com

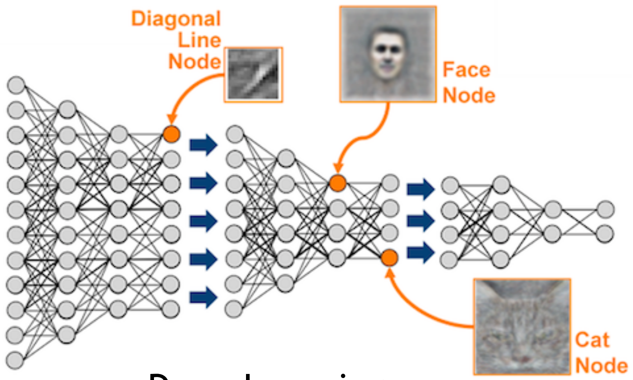
Organized by: Mérouane Debbah & Gabriel Peyré


Optimal Transport

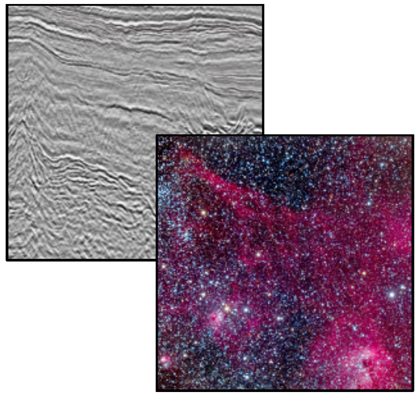

Geodesics

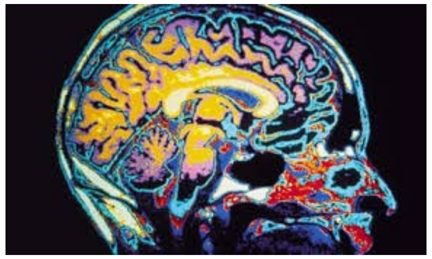

Meshes

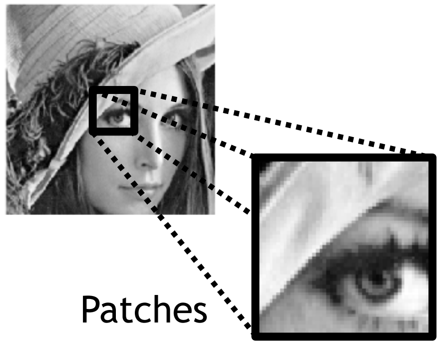

Optimization

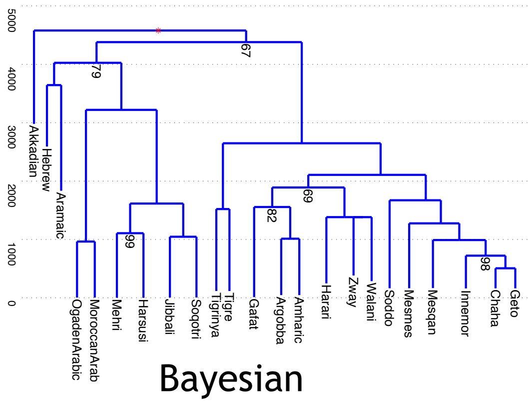


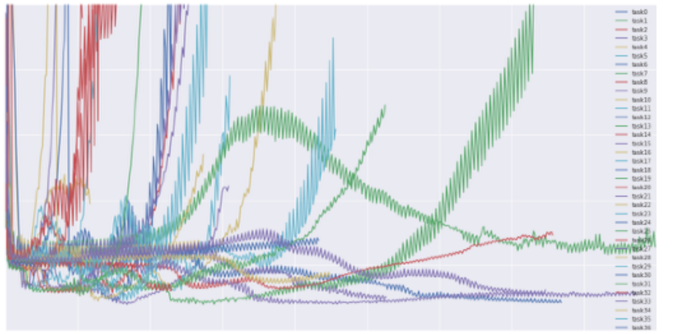
Deep Learning


Sparsity


Neuro-imaging


Patches


Bayesian


Parallel/Stochastic

Alexandre Allauzen, Paris-Sud.
Pierre Alliez, INRIA.
Guillaume Charpiat, INRIA.
Emilie Chouzenoux, Paris-Est.

Nicolas Courty, IRISA.
Laurent Cohen, CNRS Dauphine.
Marco Cuturi, ENSAE.
Julie Delon, Paris 5.

Fabian Pedregosa, INRIA.
Julien Tierny, CNRS and P6.
Robin Ryder, Paris-Dauphine.
Gael Varoquaux, INRIA.

Jalal Fadili, ENSICaen.
Alexandre Gramfort, INRIA.
Matthieu Kowalski, Supelec.
Jean-Marie Mirebeau, CNRS,P-Sud.

Examples

Optimization at the heart of:

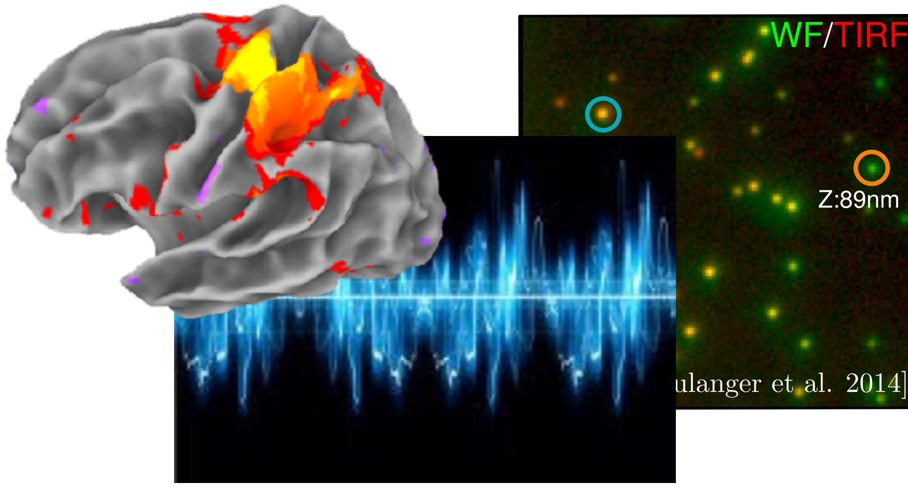
– imaging sciences (denoising, inversion, . . .)

– telecom (network design, routing, . . .)

– machine learning (classification, clustering, . . .)

$$\min_x \; \{f(x) \; ; \; x \in C\}$$

objective constraints

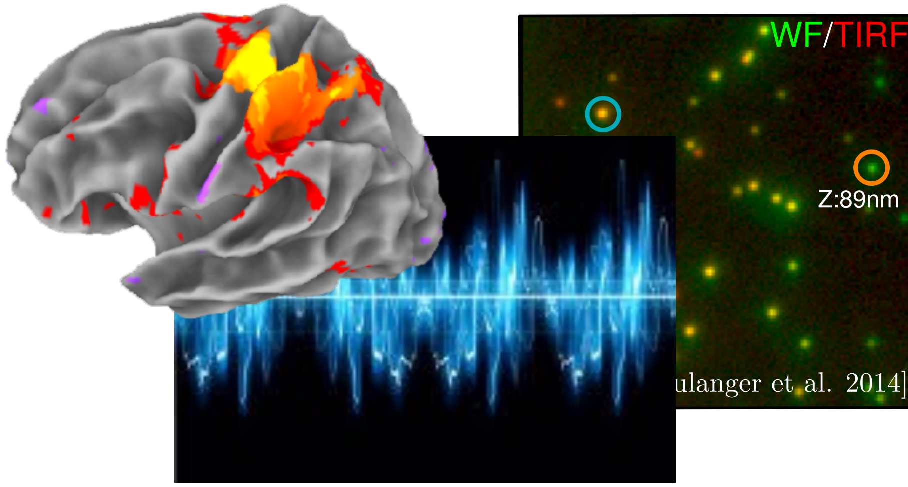


Examples

Optimization at the heart of:
- imaging sciences (denoising, inversion, ...)
- telecom (network design, routing, ...)
- machine learning (classification, clustering, ...)

$$\min_x \; \{ f(x) \; ; \; x \in C \}$$

objective constraints



(Penalized) regression:

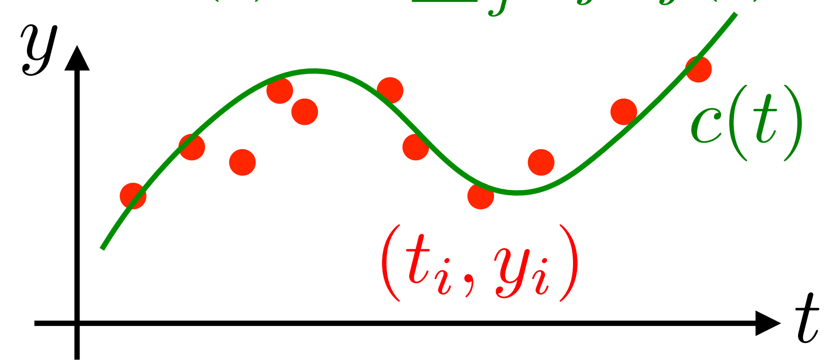
$$f(x) \overset{\text{def.}}{=} \|Ax - y\|^2 + \lambda R(x) \qquad A_{i,j} \overset{\text{def.}}{=} \varphi_j(t_i)$$

$$C \overset{\text{def.}}{=} \mathbb{R}^N$$

regularization $R(x) = \begin{cases} \|x\|_2^2 = \sum_j |x_j|^2, \\ \|x\|_1 = \sum_j |x_j| \\ \dots \end{cases}$

$\rightarrow$ Unconstraint optimization problem.

$$c(t) \overset{\text{def.}}{=} \sum_j x_j \varphi_j(t)$$



Examples

Optimization at the heart of:

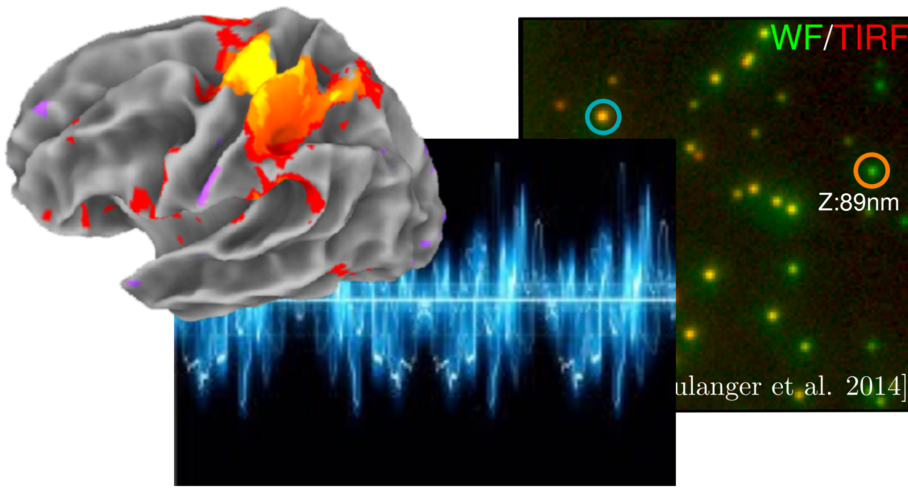
– imaging sciences (denoising, inversion, ...)

– telecom (network design, routing, ...)

– machine learning (classification, clustering, ...)

$$\min_x \; \{f(x) \; ; \; x \in C\}$$

objective constraints



(Penalized) regression:

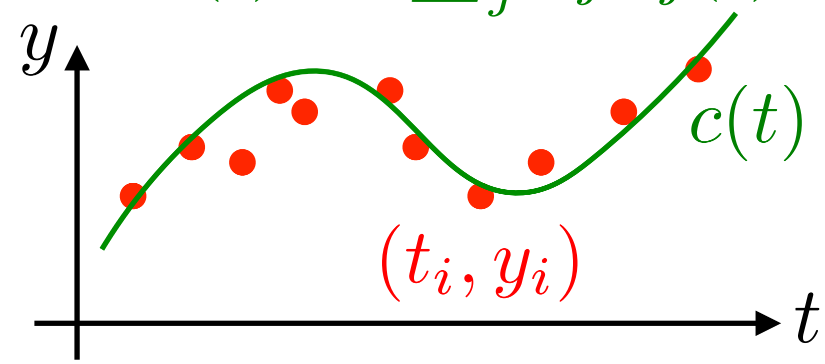
$$f(x) \overset{\text{def.}}{=} \|Ax - y\|^2 + \lambda R(x) \qquad A_{i,j} \overset{\text{def.}}{=} \varphi_j(t_i)$$

$$C \overset{\text{def.}}{=} \mathbb{R}^N$$

regularization $R(x) = \begin{cases} \|x\|_2^2 = \sum_j |x_j|^2, \\ \|x\|_1 = \sum_j |x_j| \\ \ldots \end{cases}$

→ Unconstraint optimization problem.

$$c(t) \overset{\text{def.}}{=} \sum_j x_j \varphi_j(t)$$

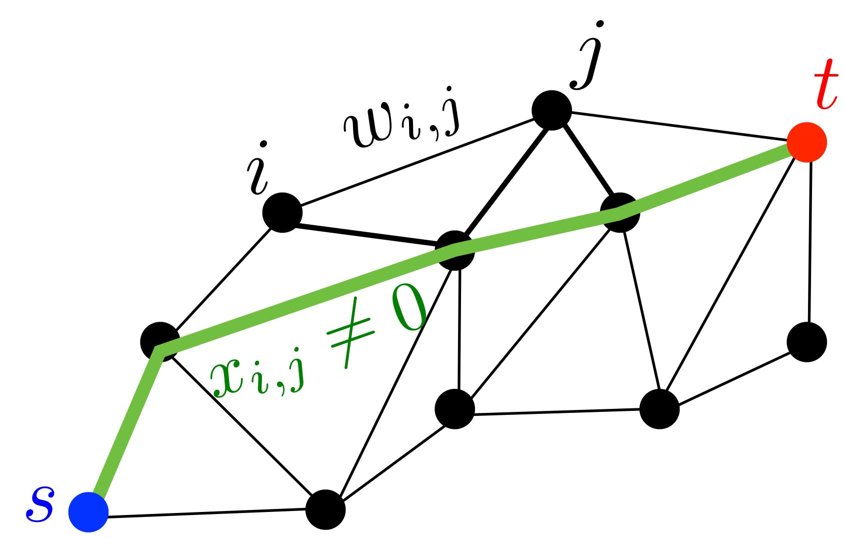


Min-cost flow problem

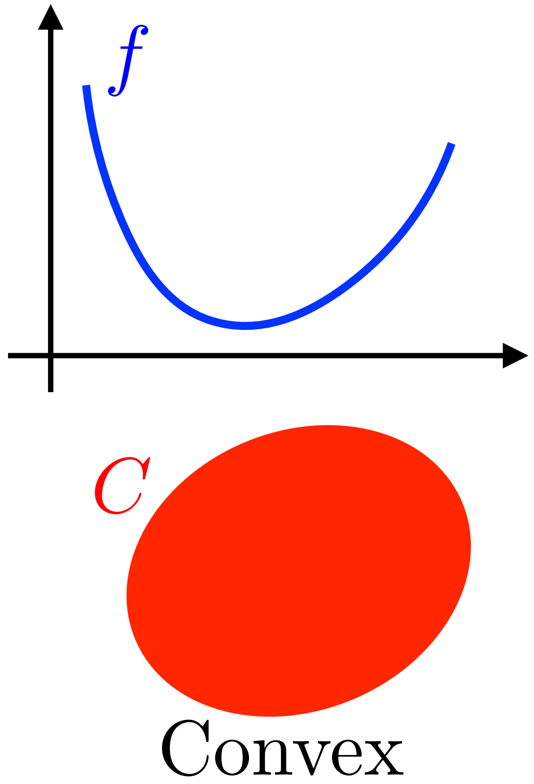
$$f(x) \overset{\text{def.}}{=} \sum_{i,j} w_{i,j} |x_{i,j}|$$

$$C \overset{\text{def.}}{=} \{x \; ; \; \text{div}(x) = \delta_s - \delta_t\}$$

→ Linear programming.

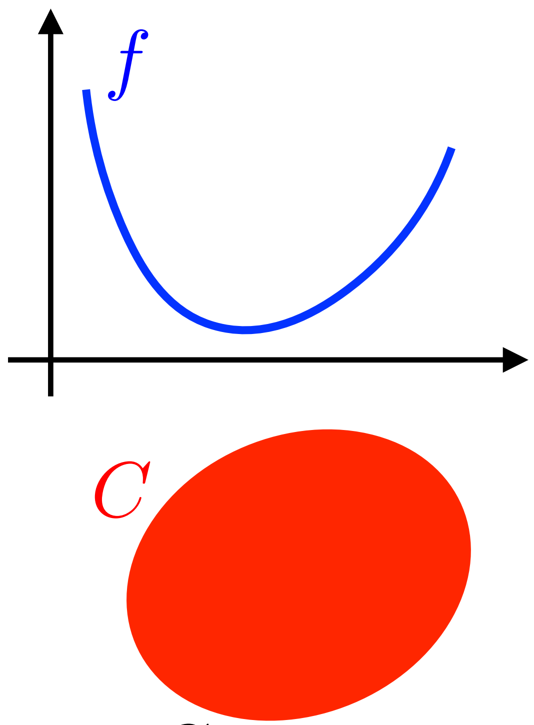


Non-smooth vs. Non-convex

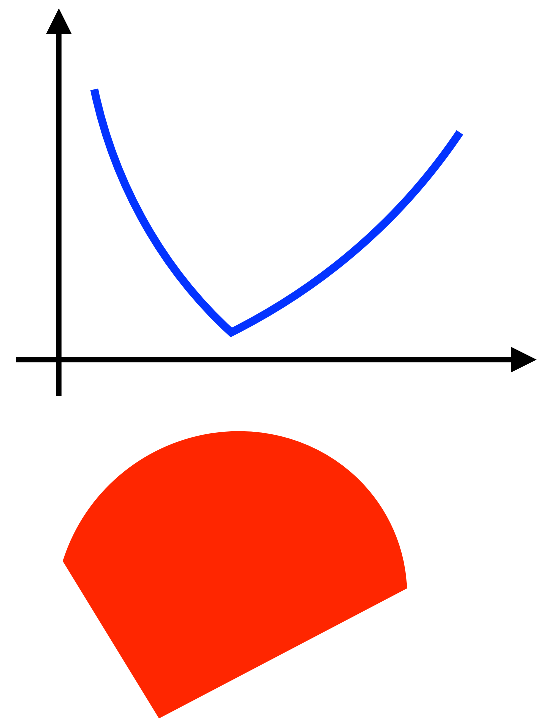


Convexity is nice: no local minimum, duality, guaranteed algorithms.

Non-smooth vs. Non-convex



Convex

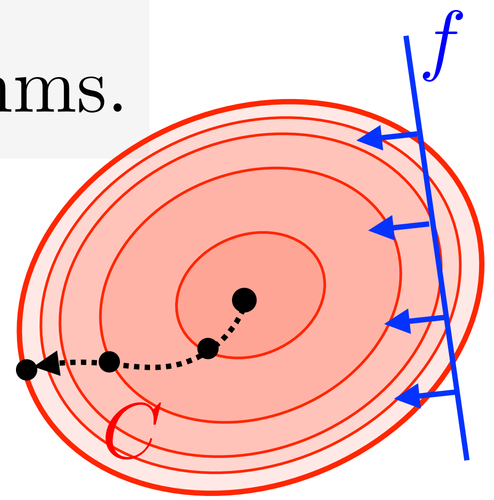


Non-smooth

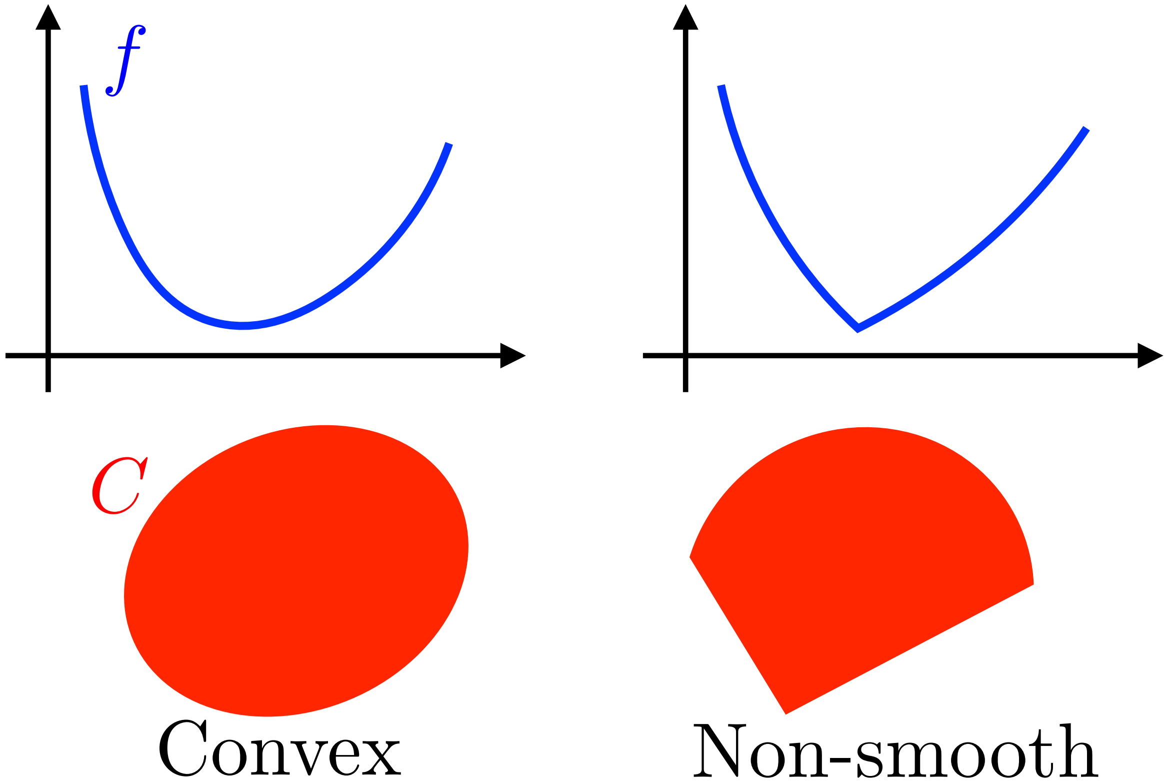
Convexity is nice: no local minimum, duality, guaranteed algorithms.

Many non-smooth problem are now routinely solved:

$\rightarrow$ interior points methods (small size + high-precision).



Non-smooth vs. Non-convex

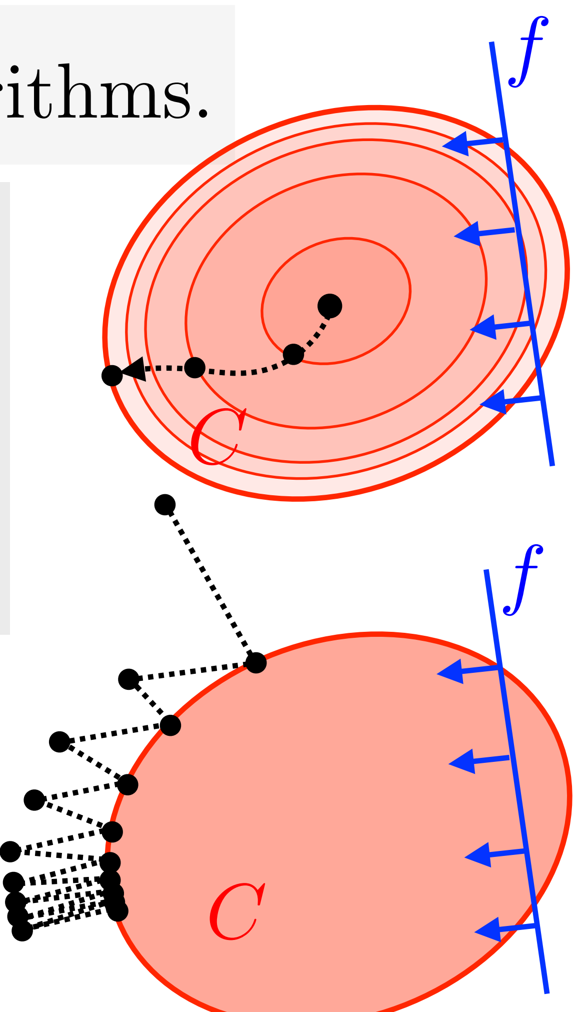


Convexity is nice: no local minimum, duality, guaranteed algorithms.

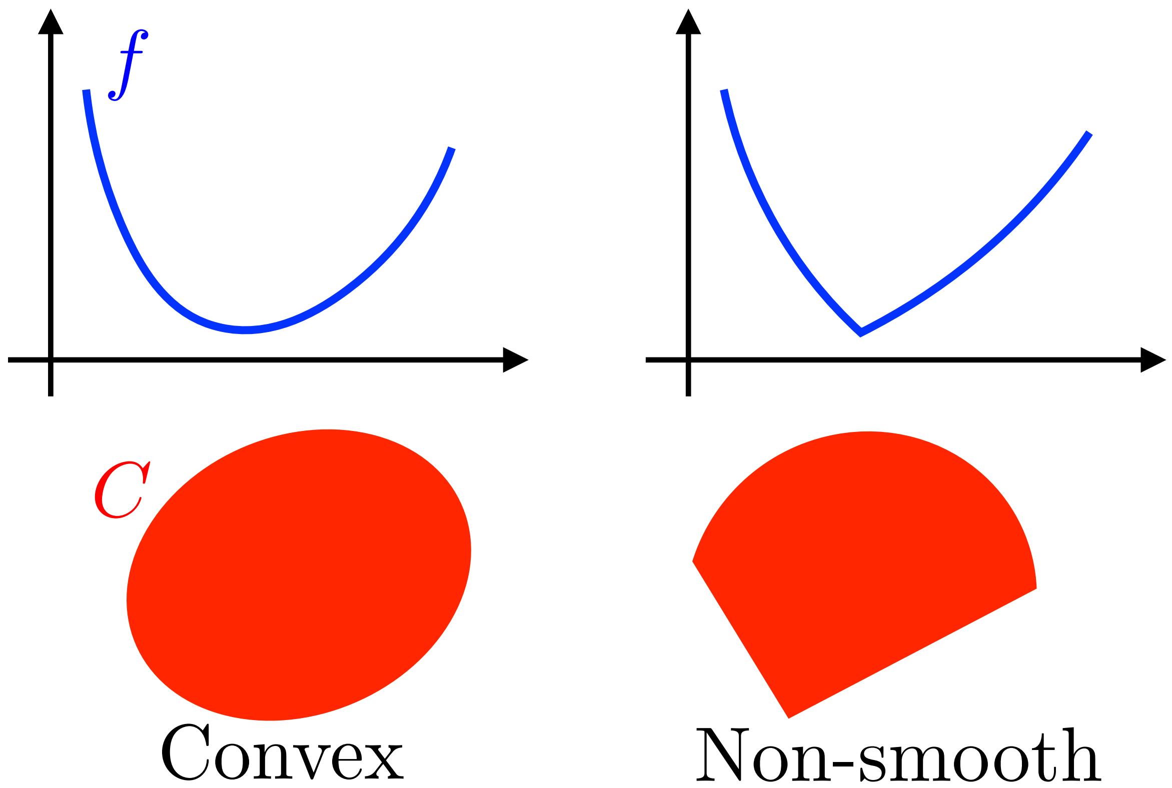
Many non-smooth problem are now routinely solved:

→ interior points methods (small size + high-precision).

→ first order schemes (gradient, proximal, Frank-Wolfe).



Non-smooth vs. Non-convex



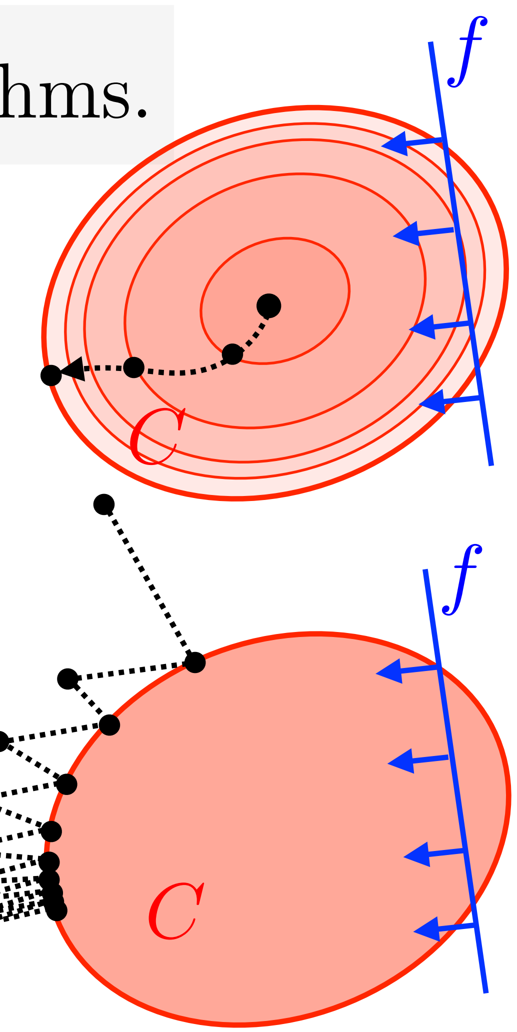
Convexity is nice: no local minimum, duality, guaranteed algorithms.

Many non-smooth problem are now routinely solved:

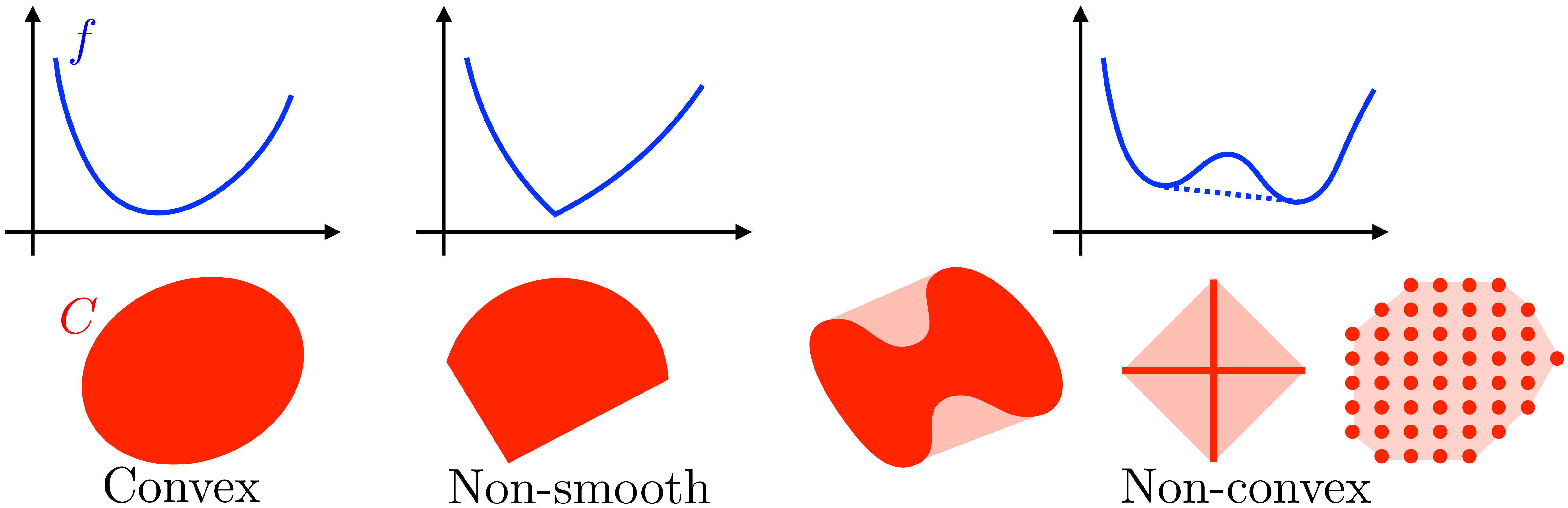
$\rightarrow$ interior points methods (small size + high-precision).

$\rightarrow$ first order schemes (gradient, proximal, Frank-Wolfe).

$\rightarrow$ structured problems (sparse, low-rank, etc.)



Non-smooth vs. Non-convex



Convexity is nice: no local minimum, duality, guaranteed algorithms.

Many non-smooth problem are now routinely solved:

$\rightarrow$ interior points methods (small size + high-precision).

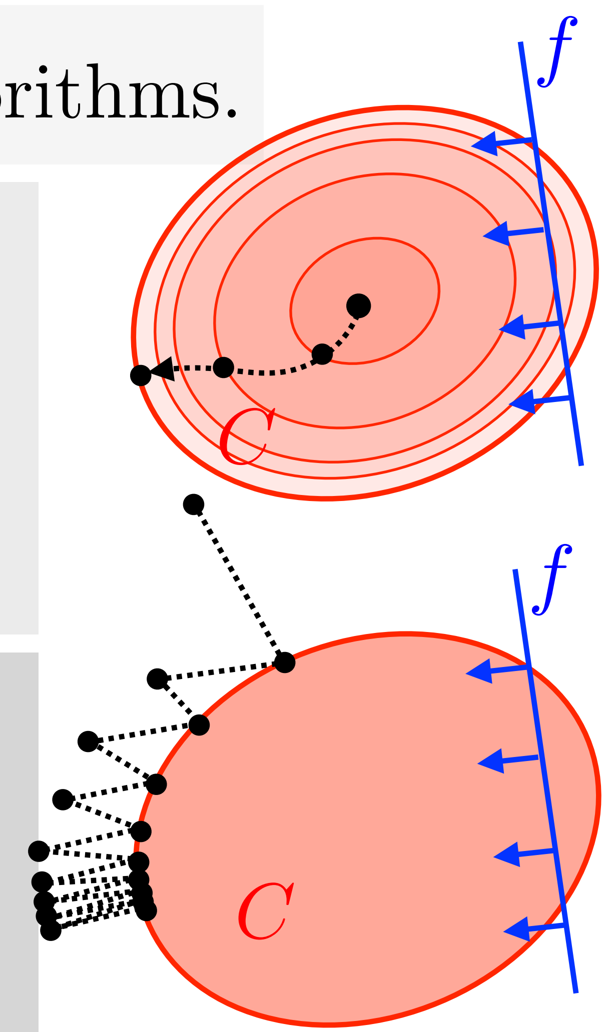
$\rightarrow$ first order schemes (gradient, proximal, Frank-Wolfe).

$\rightarrow$ structured problems (sparse, low-rank, etc.)

Non-convexity is hard: convexification is sometime possible.

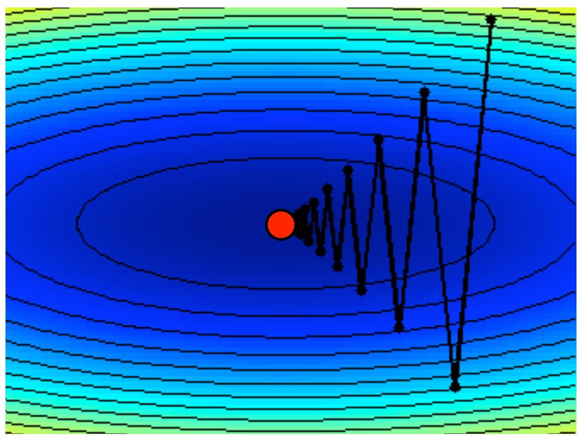
$\rightarrow$ sparsity with ℓ^1, low-rank with nuclear norm.

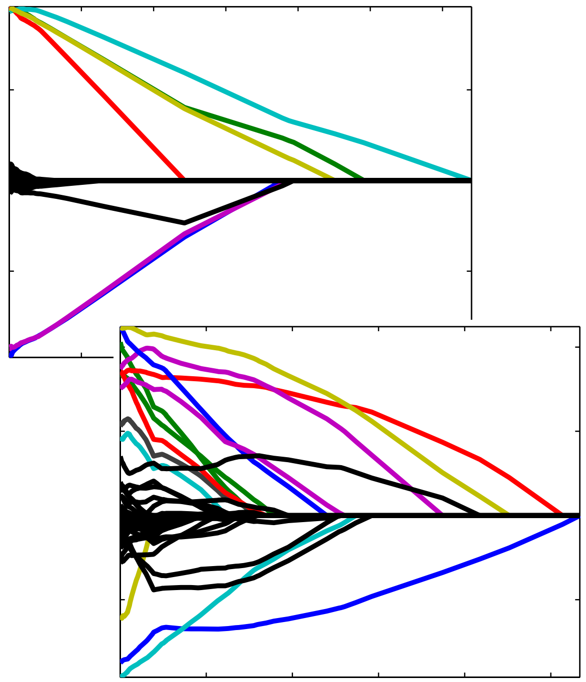
$\rightarrow$ Lasserre SDP hierarchy (small size).

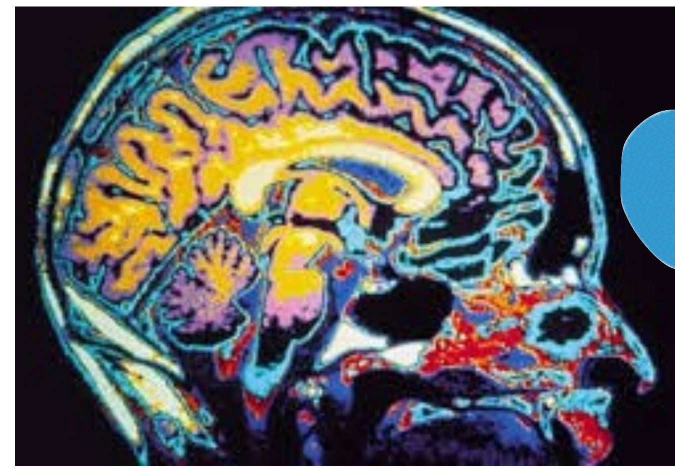


What's next

Alexandre Gramfort: sparsity, applications in ML/imaging.







Matthieu Kowalski: proximal methods, applications in audio.