Séminaire de Calcul Scientifique du CERMICS

École des Ponts
ParisTech

# ‘Off-the-grid’ methods

Yohann De Castro (Orsay-INRIA)

26 juin 2018

CERMICS
June 26th, 2018

Yohann De Castro (Orsay - INRIA)

# « Off-The-Grid » methods

UNIVERSITÉ PARIS SUD

Inria

Yohann DE CASTRO

**Orthant**

$$x \geq 0$$

Vectors

- **High-Dimensional Statistics**

UNIVERSITÉ PARIS SUD
Inria

Yohann DE CASTRO

**Orthant**

$x \geq 0$

Vectors

**SDP**

$M \succeq 0$

Matrices

- **High-Dimensional Statistics**
- **Matrix completion, recovery, factorisation**

Yohann DE CASTRO

| **Orthant** | **SDP** | **Lasserre's hierarchies** |
| --- | --- | --- |
| $x \geq 0$ | $M \succeq 0$ | $\mu \geq 0$ |
| Vectors | Matrices | Measures |

- **High-Dimensional Statistics**
- **Matrix completion, recovery, factorisation**
- **Global Optimization**
- **Sparse Deconvolution**
- **Computational Limits**

Yohann DE CASTRO

**Orthant**

$$x \geq 0$$

Vectors

**SDP**

$$M \succeq 0$$

Matrices

**Lasserre's hierarchies**

$$\mu \geq 0$$

Measures

- **High-Dimensional Statistics**

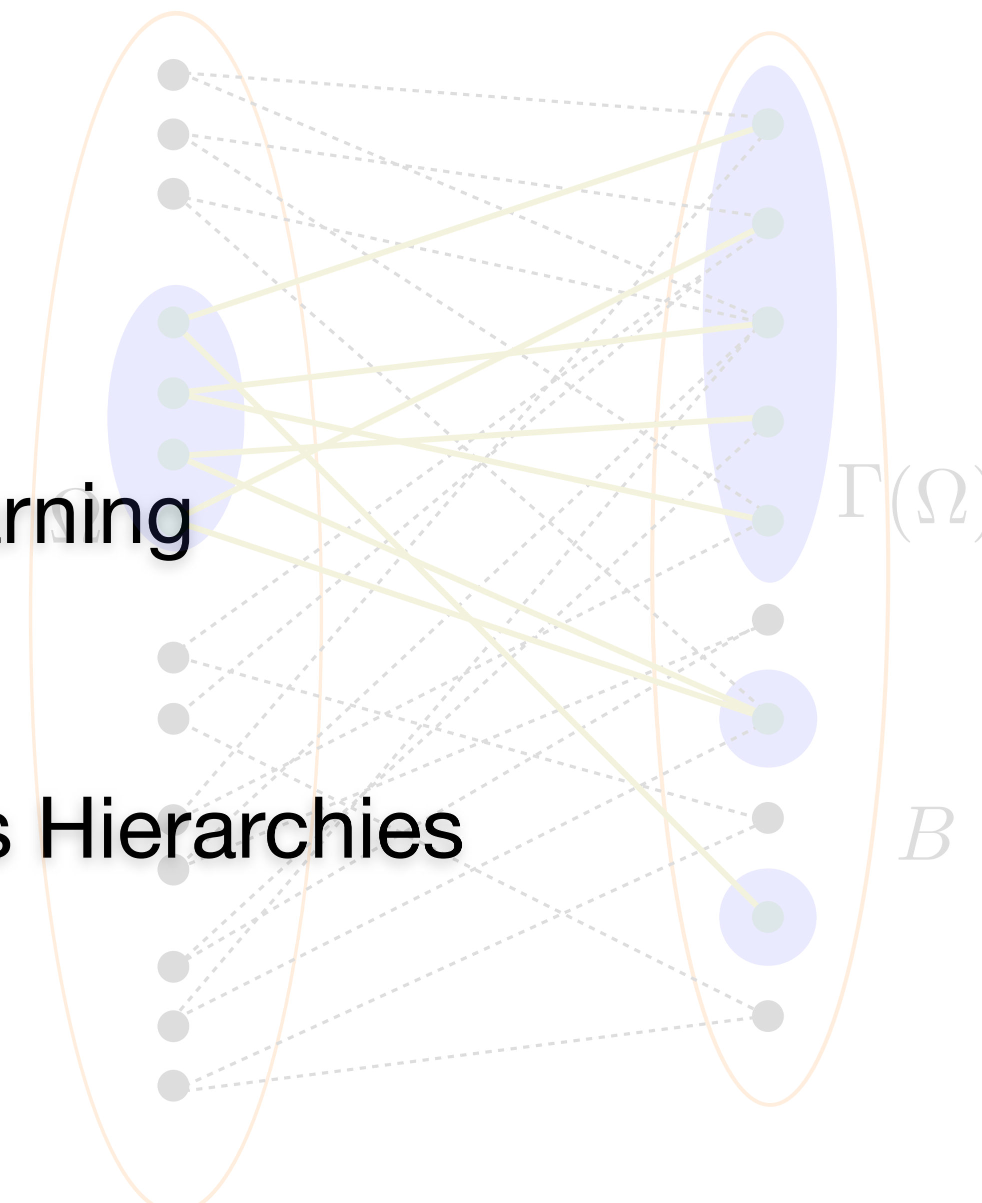
- **Matrix completion, recovery, factorisation**

- **Global Optimization**
- **Sparse Deconvolution**
- **Computational Limits**

Grids

« Off-The-Grid » Methods

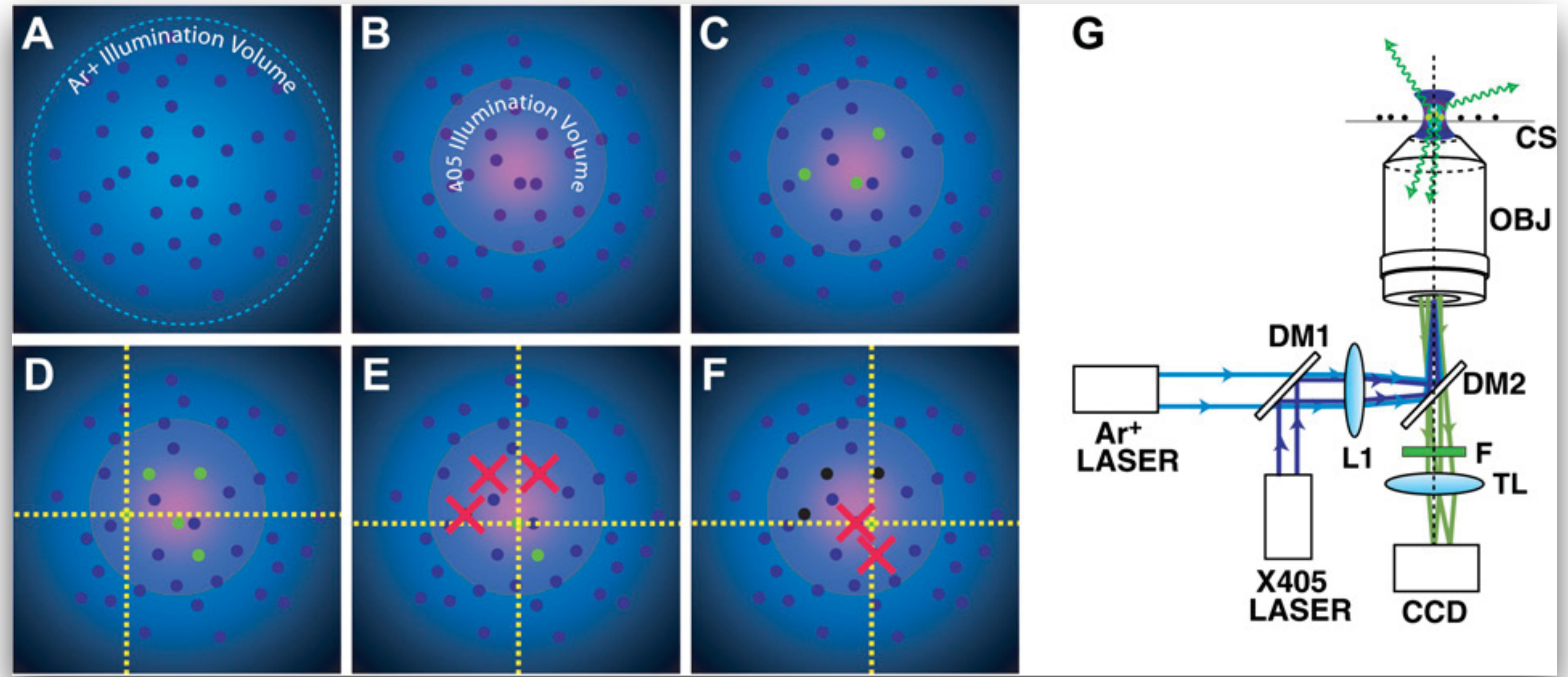
- « Off-The-Grid » Statistical Learning
- The Meta Algorithm: Lasserre's Hierarchies





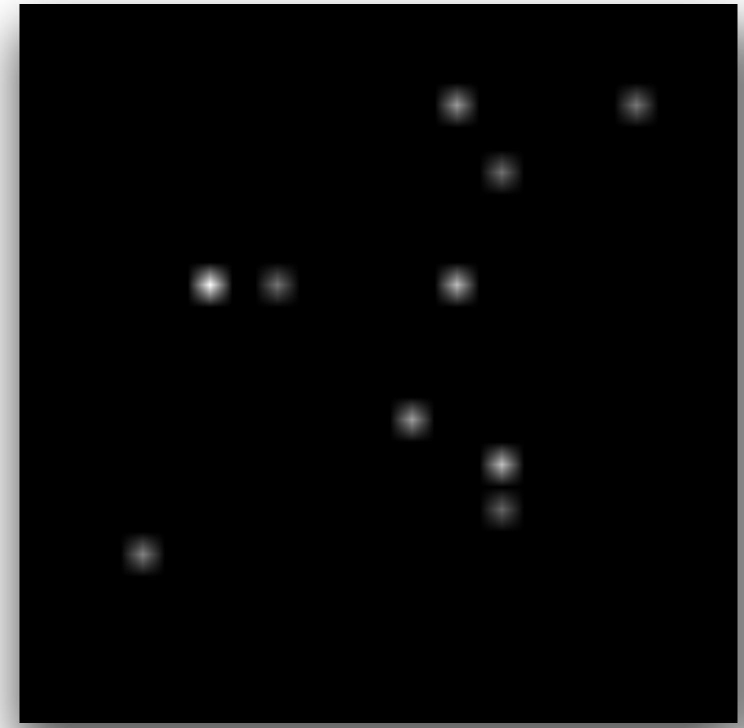

UNIVERSITÉ PARIS SUD
Inria

# Sparse Deconvolution

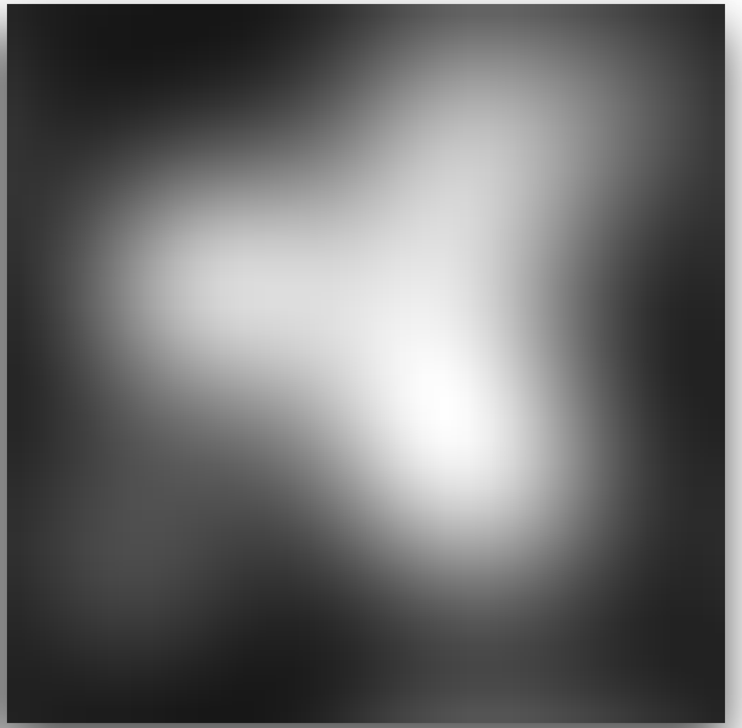



- S. Hess, T. Girirajan, M. Mason, Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy, *Biophysical Journal* (2004).

# Sparse Deconvolution



(a) Heat source $u_0$

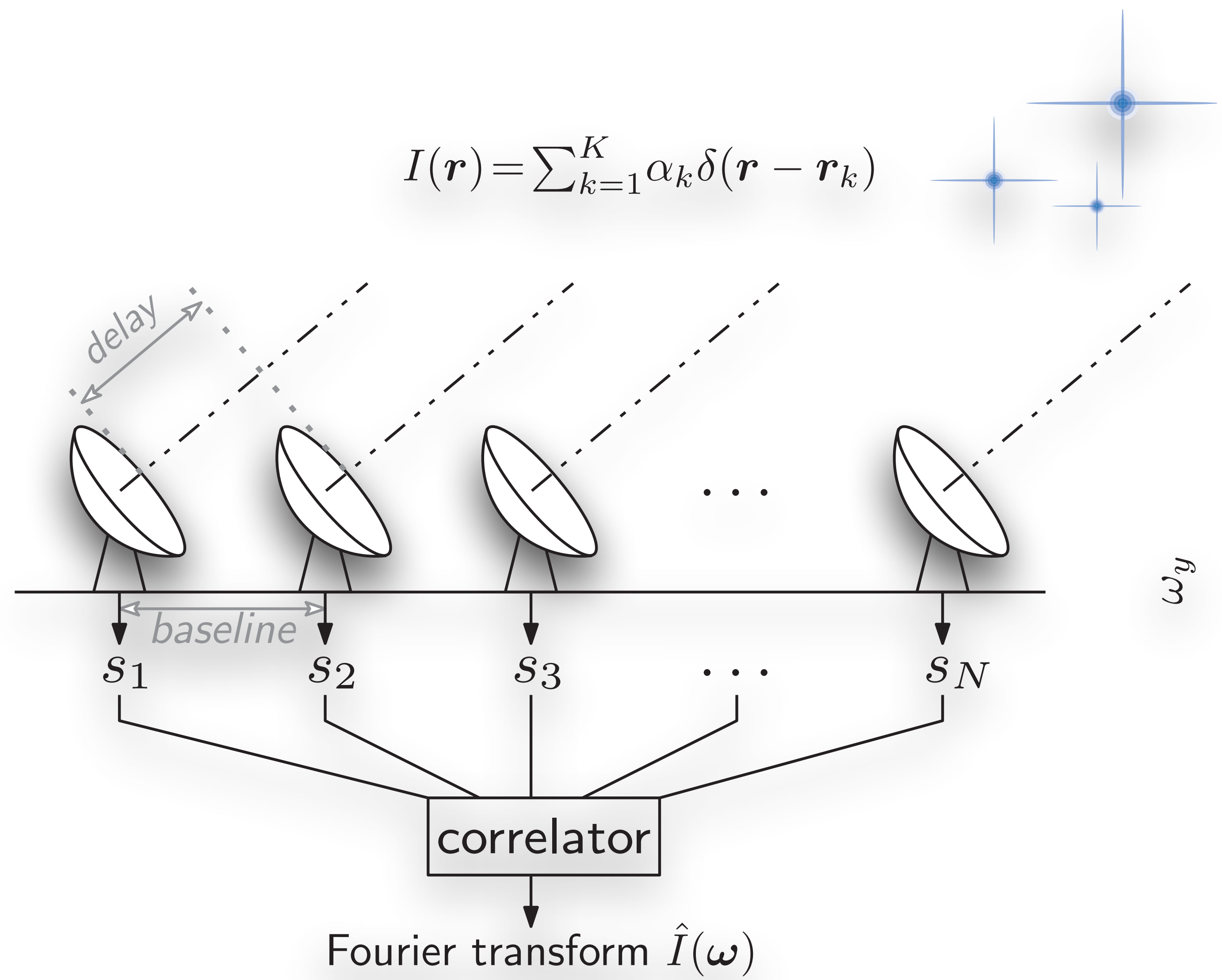


(b) $Au_0 = f$

- Y. Li, S. Osher, R. Tsai, Heat Source Identification based on L1 Constrained Minimization, *Inverse Problems and Imaging* (2014).

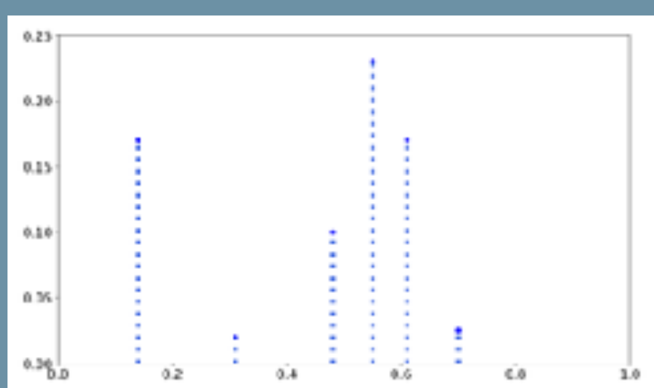
Yohann DE CASTRO

# Sparse Deconvolution




- H. Pan, T. Blu, M. Vetterli, Towards Generalized FRI Sampling With an Application to Source Resolution in Radioastronomy, *IEEE trans. on Signal Processing* (2017).

# Sparse Deconvolution



$x$

Yohann DE CASTRO

# Sparse Deconvolution

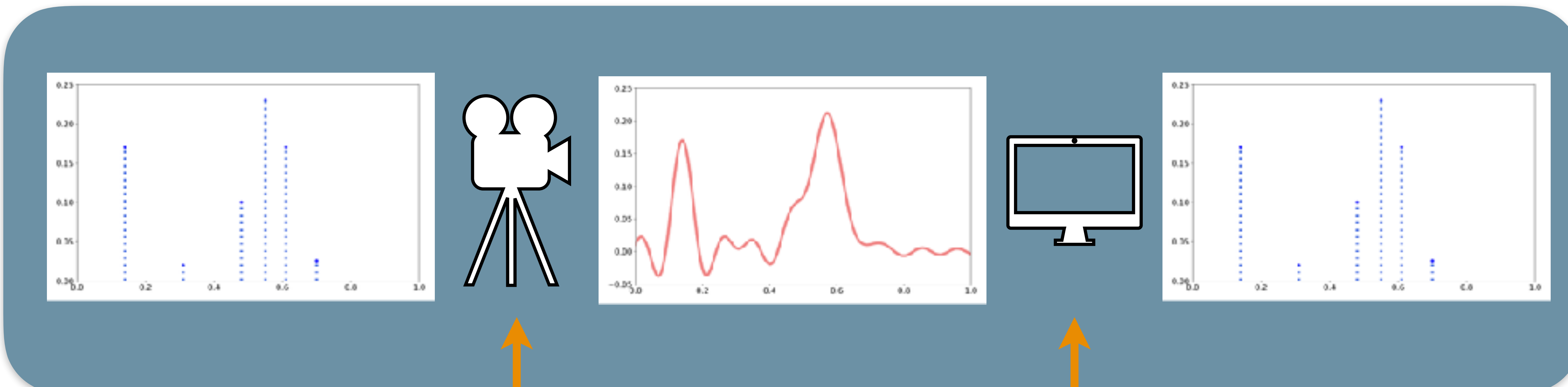
$x$ $y$

$$y = \Phi(x)$$

Yohann DE CASTRO

# Sparse Deconvolution

$$\|x\|_{\mathrm{TV}} = \min_{x=x_+-x_-\ ,x_\pm\geq 0} \int \mathrm{d}x_+ + \int \mathrm{d}x_-$$



$x$

$y$

**Perfect Recovery**

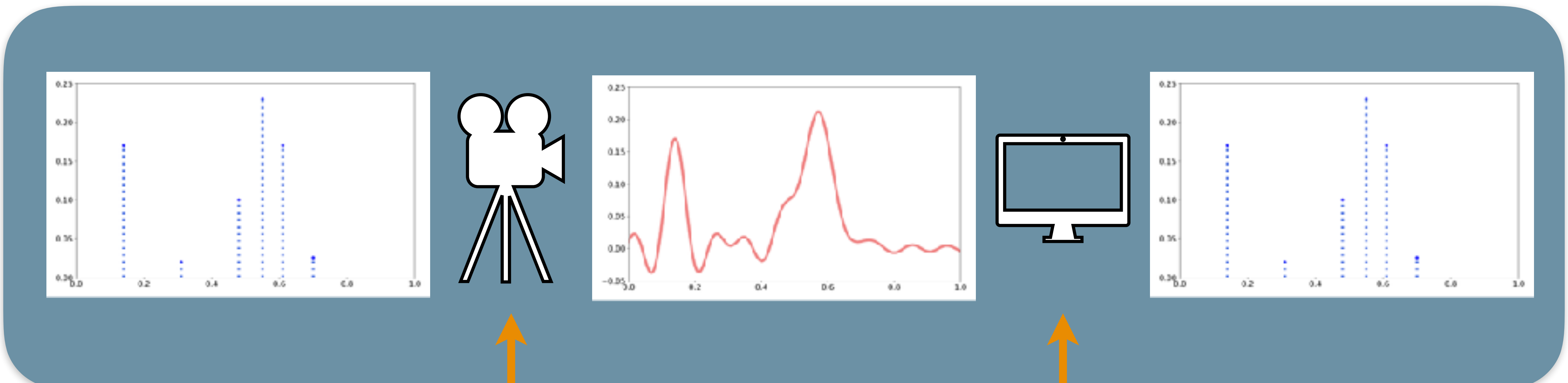
$$y = \Phi(x)$$

$$\min_{y=\Phi(x)} \|x\|_{\mathrm{TV}}$$

Yohann DE CASTRO

# Sparse Deconvolution

$$||x||_{\mathrm{TV}} = \min_{x=x_+-x_-\,,x_\pm\geq 0} \int \mathrm{d}x_+ + \int \mathrm{d}x_-$$



$x$

$y$

Perfect Recovery

$$y = \Phi(x)$$

$$\min_{y=\Phi(x)} ||x||_{\mathrm{TV}}$$

Discrete Measure

# Sparse Deconvolution

$$\min_{y=\Phi(x)} ||x||_{\mathrm{TV}} = \min_{y=\Phi(x_+)-\Phi(x_-)\,,\ x_\pm\geq 0} \langle 1, x_+\rangle + \langle 1, x_-\rangle$$

# Sparse Deconvolution

$$\min_{y=\Phi(x)} ||x||_{\mathrm{TV}} = \min_{y=\Phi(x_+)-\Phi(x_-)\,,\ x_\pm\geq 0} \langle 1, x_+\rangle + \langle 1, x_-\rangle$$

- **[Cone]** Moments of Nonnegative Measures
- **[Algorithms]** Lasserre's Hierarchies, Frank-Wolfe

# Sparse Deconvolution

$$\min_{y=\Phi(x)} ||x||_{\mathrm{TV}} = \min_{y=\Phi(x_+)-\Phi(x_-)\,,\ x_\pm \geq 0} \langle 1, x_+ \rangle + \langle 1, x_- \rangle$$

**Linear Conic with « Moments » Constraints**

- **[Cone]** Moments of Nonnegative Measures
- **[Algorithms]** Lasserre's Hierarchies, Frank-Wolfe

# Robust Moment Problem

Robustness

- $y = \Phi(x) + \eta$ :

$$\bar{x} \in \arg\min_v \left\{ \frac{1}{2} ||y - \Phi(v)||_2^2 + \lambda ||v||_{\mathrm{T}V} \right\}$$

# Robust Moment Problem

Robustness

- $y = \Phi(x) + \eta$ :

$$\bar{x} \in \arg\min_{v} \left\{ \frac{1}{2} ||y - \Phi(v)||_2^2 + \lambda ||v||_{\mathrm{TV}} \right\}$$

**[DC et al; Candès and Fernandez-Granda; Recht; …] Spike Detection**

- Analysis via « **Dual Certificate** » **[DC, Gamboa; Candès, Tao; …]**

# Robust Moment Problem

- $y = \Phi(x) + \eta$ :

Robustness

$$\bar{x} \in \arg\min_{v} \left\{ \frac{1}{2} ||y - \Phi(v)||_2^2 + \lambda ||v||_{TV} \right\}$$

**[DC et al; Candès and Fernandez-Granda; Recht; …] Spike Detection**

- Analysis via « **Dual Certificate** » **[DC, Gamboa; Candès, Tao; …]**
- « Minimal Separation Condition » on the support of the spikes
- Optimal error bounds

**+ Support Recovery**

# Robust Moment Problem

- $y = \Phi(x) + \eta$ :

Robustness

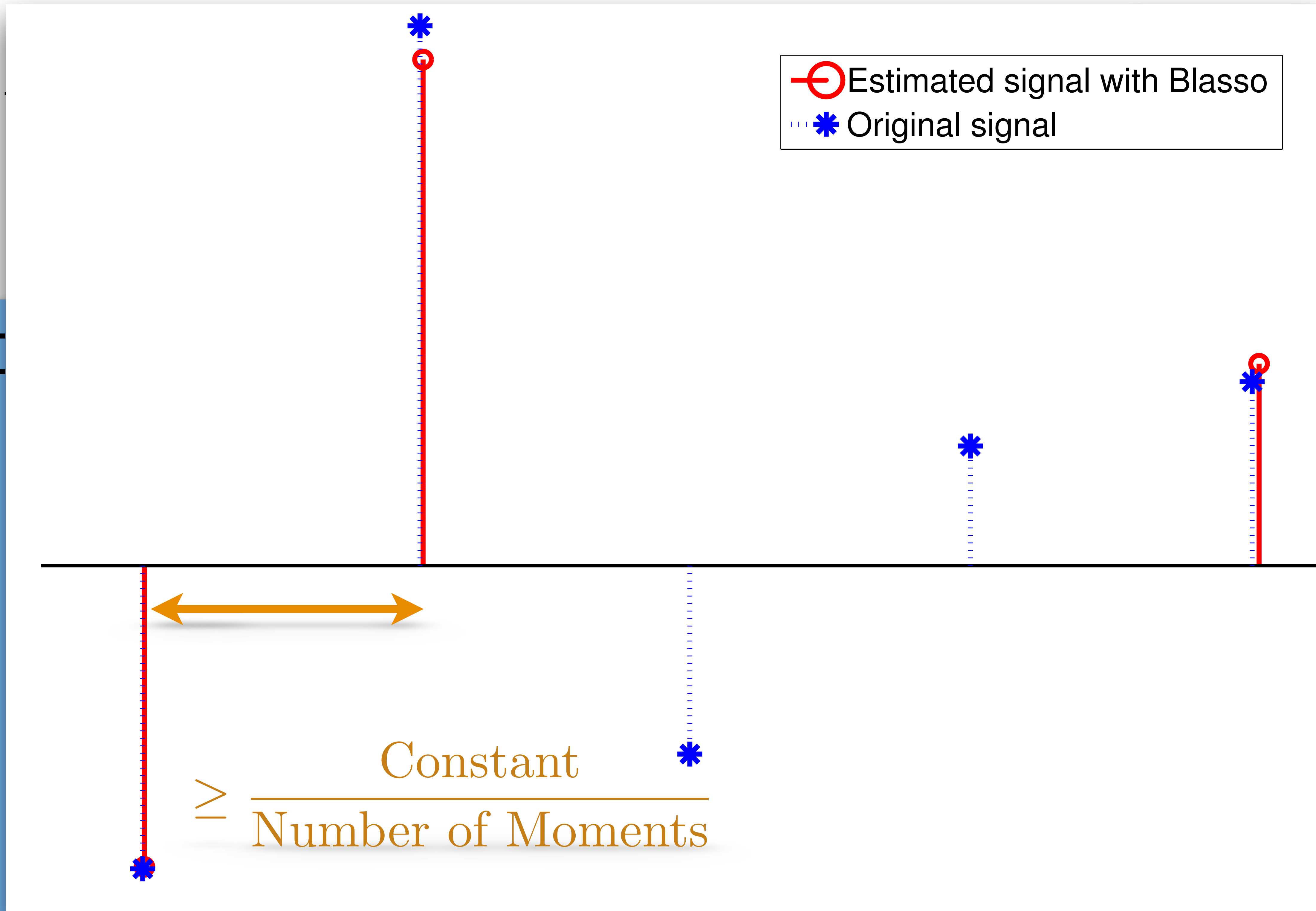
$$\bar{x} \in \arg\min_{v} \left\{ \frac{1}{2} ||y - \Phi(v)||_2^2 + \lambda ||v||_{TV} \right\}$$

**[DC et al; Candès and Fernandez-Granda; Recht; …] Spike Detection**

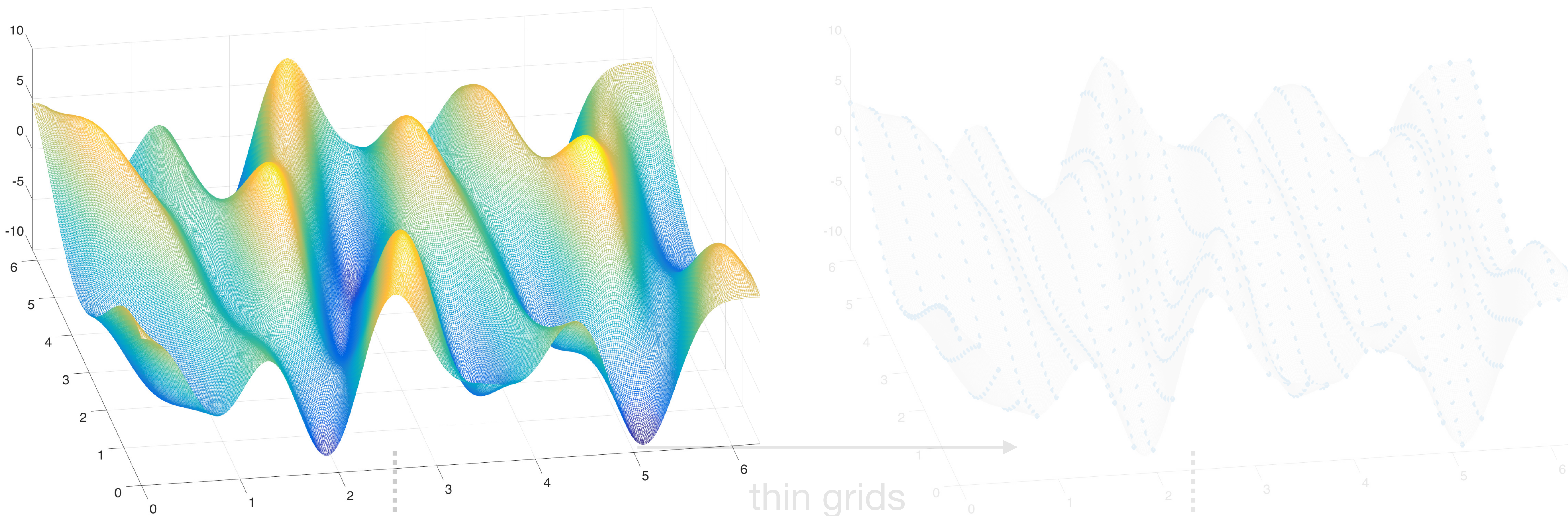
- Analysis via « **Dual Certificate** » **[DC, Gamboa; Candès, Tao; …]**
- « Minimal Separation Condition » on the support of the spikes
- Optimal error bounds
- **[DC, Gamboa, Henrion, Lasserre]**
  Tractable Algorithm via « Lasserre's hierarchies »

**+ Support Recovery**

# Robust Moment Problem



# Grid vs Off-the-grid [ADCM18]



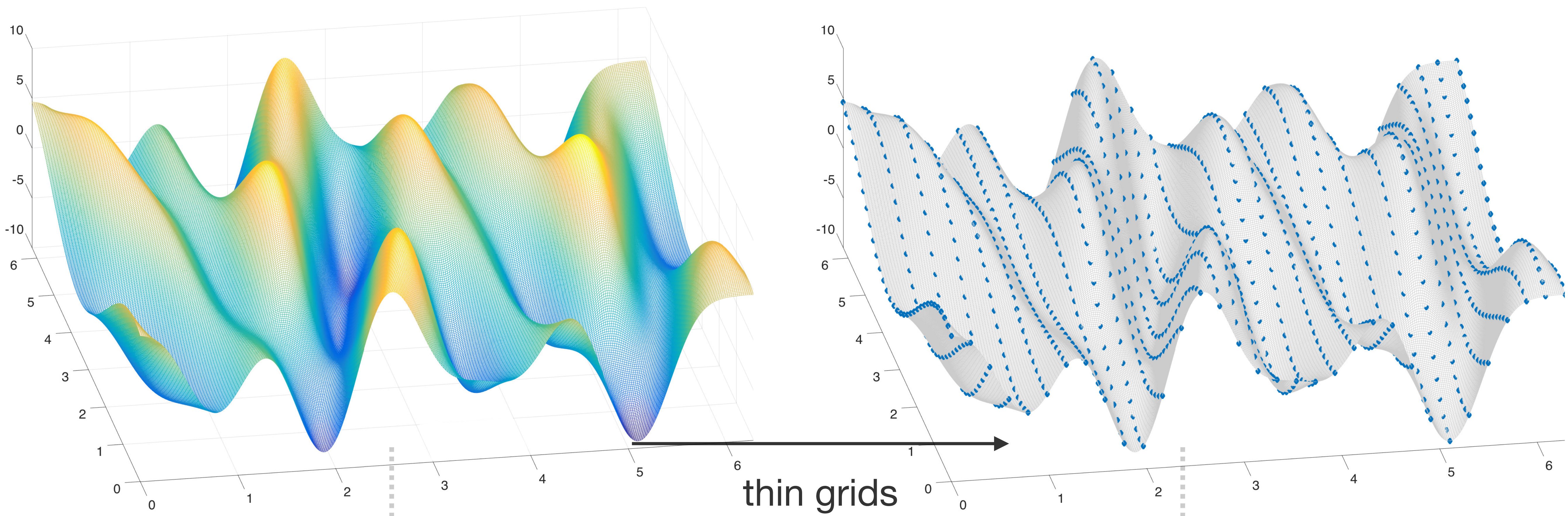

**Inference from the continuous process**

Inference from $p$ grid points

Grid-less approach

Grid approach

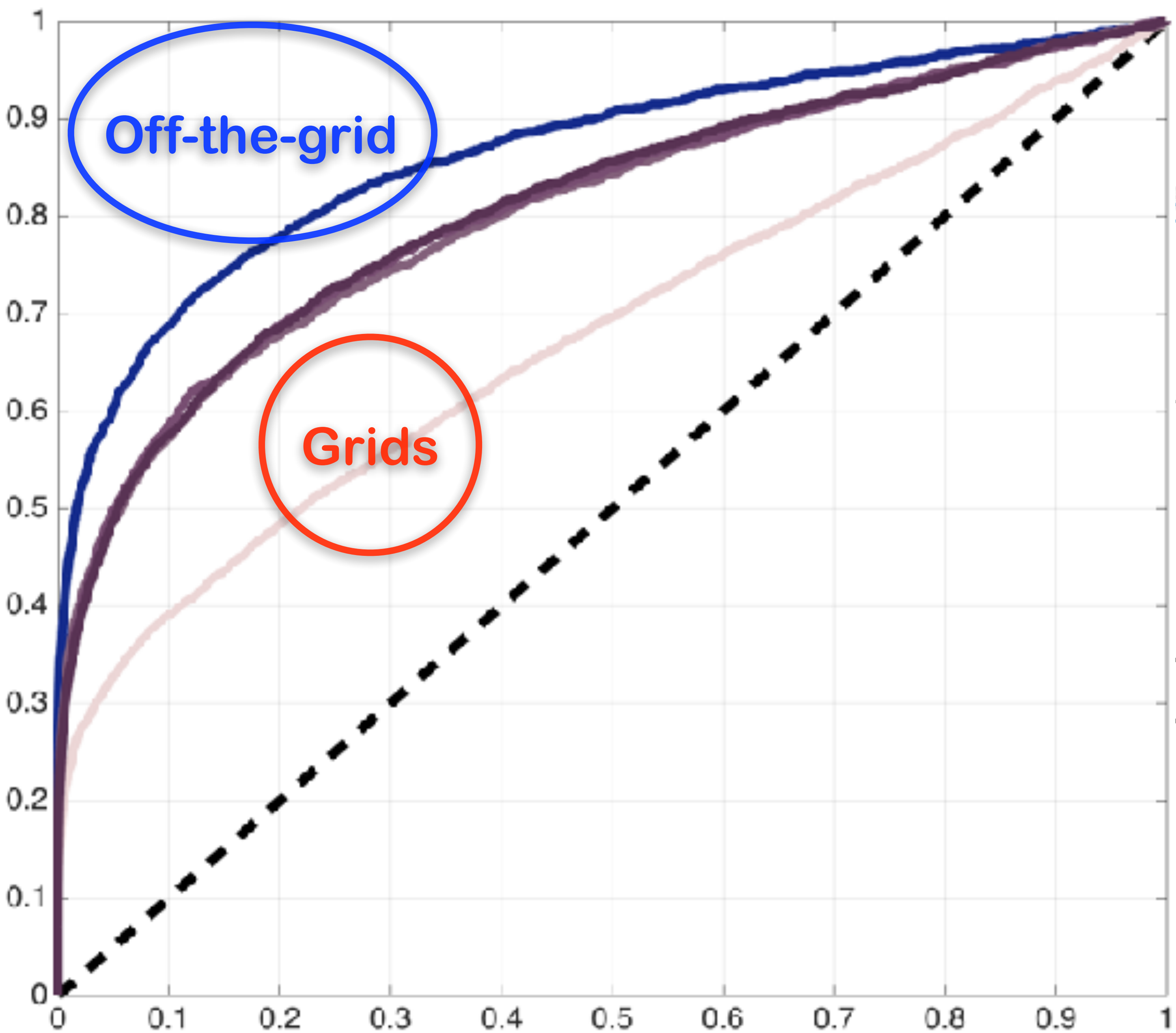
# Grid vs Off-the-grid [ADCM18]

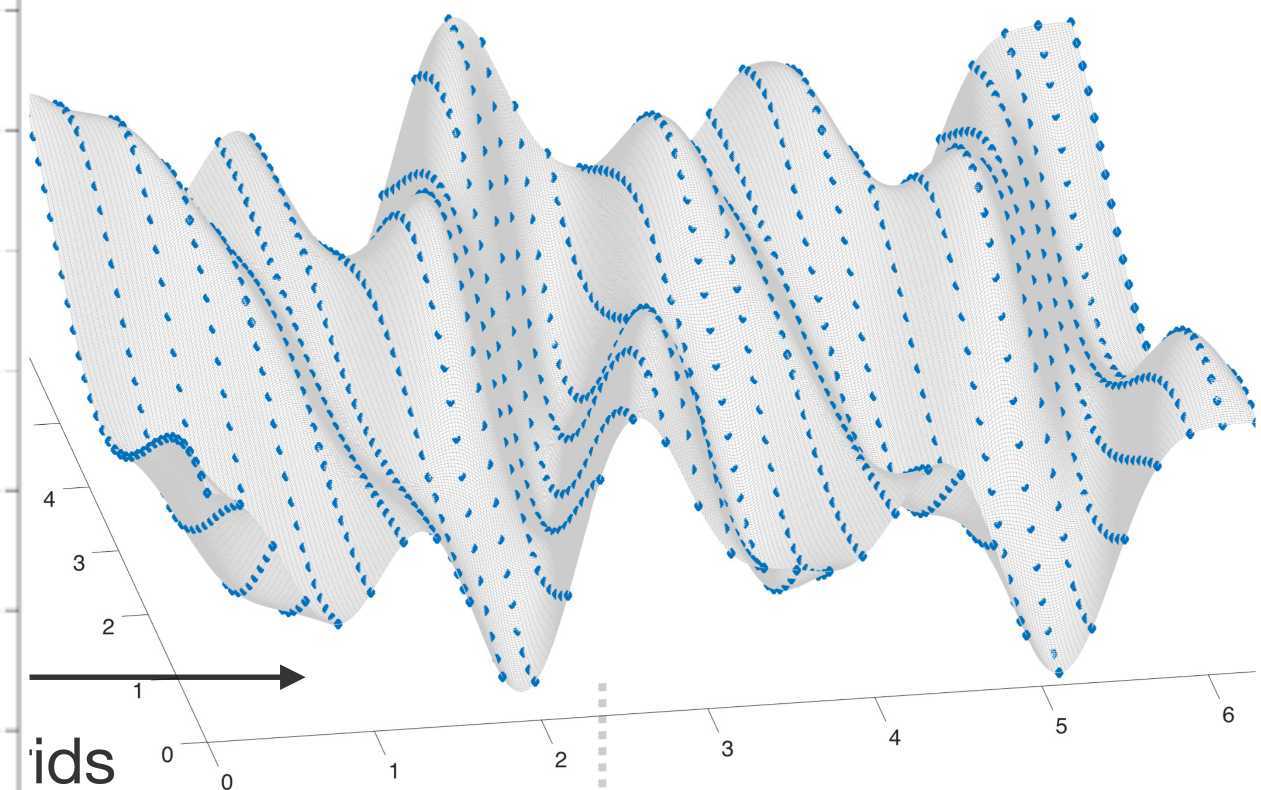



**Inference from the continuous process** | **Inference from $p$ grid points**

Grid-less approach | Grid approach

# Grid vs Off-the-grid [ADCM18]



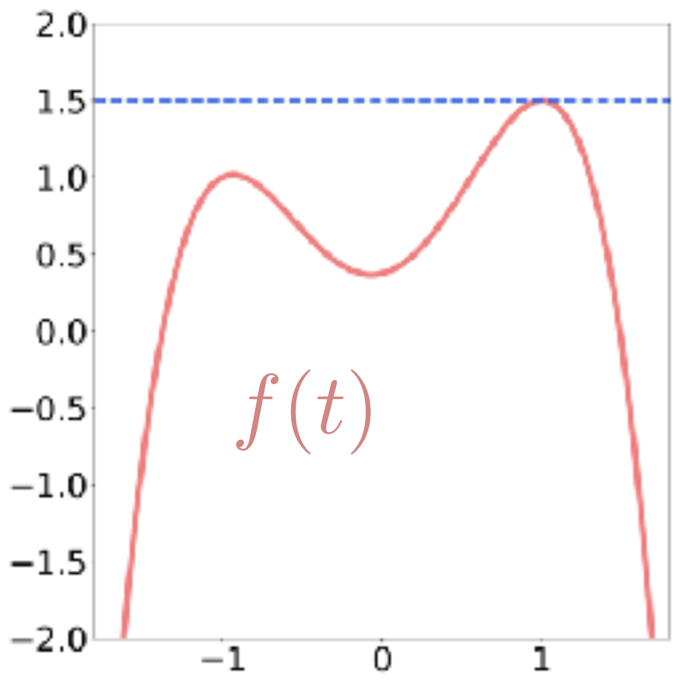

**Inference from $p$ grid points**

**Grid approach**

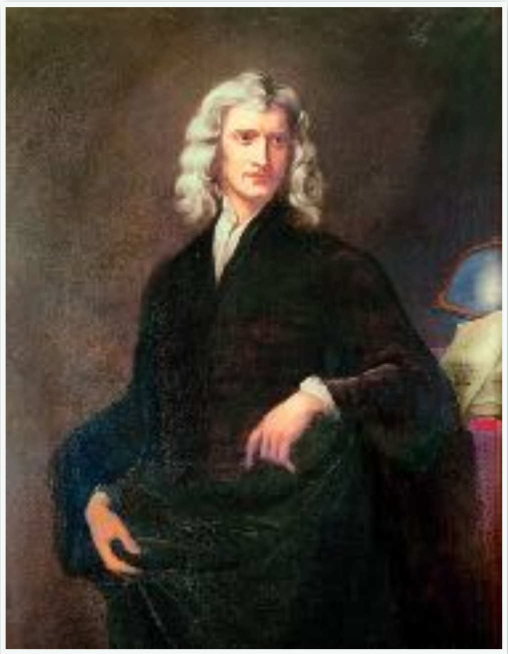
Under one Sparse alternative

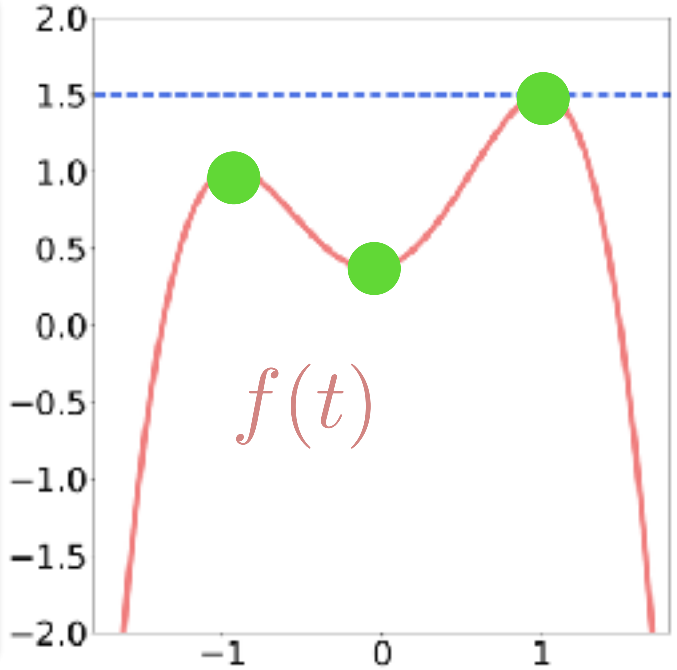
Yohann DE CASTRO

- « Off-The-Grid » Statistical Learning
- The Meta Algorithm: Lasserre's Hierarchies


2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−1
0
1
$f(t)$

Yohann DE CASTRO

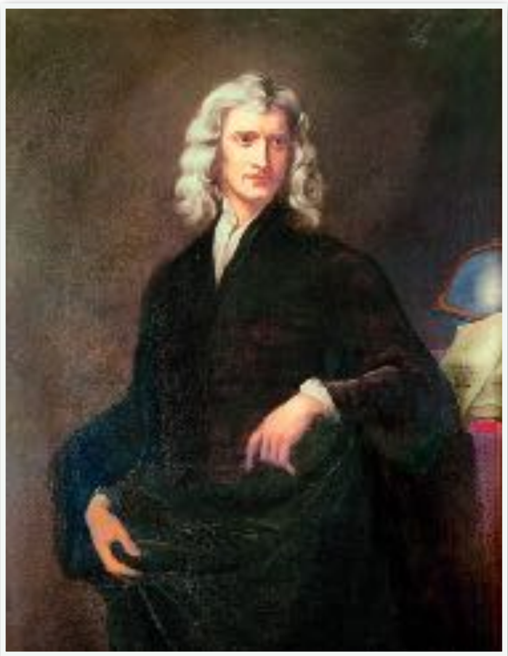


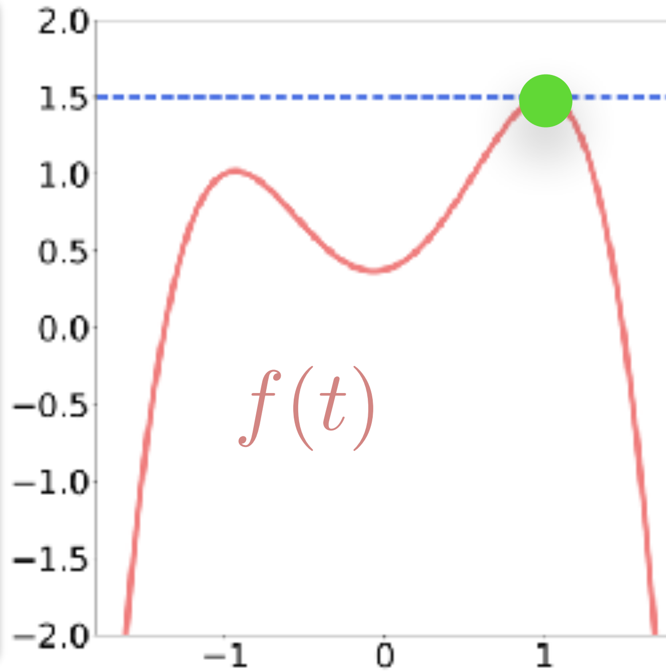



**Newton's approach**

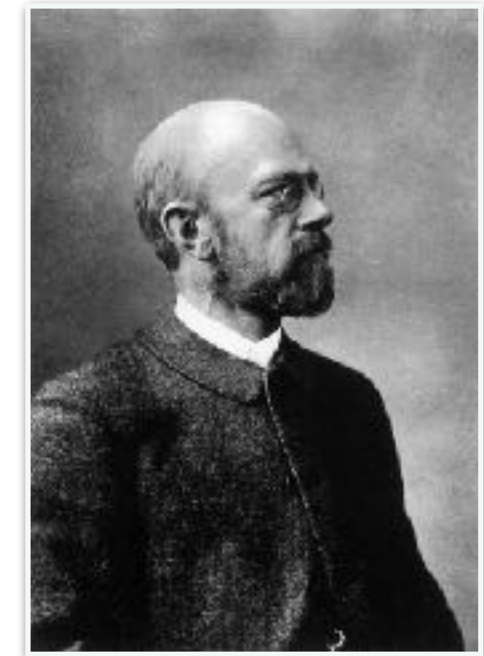
- Search through **local** optima
- Well suited for **convex** functions
- BUT simple non-convex function may have an exponential number of bad optima, e.g.

$$f(t_1) + \ldots + f(t_n)$$






$$f(t) = \frac{3}{2} - \frac{(t-1)^2}{8} - (t^2-1)^2$$



**Newton's approach**

- Search through **local** optima
- Well suited for **convex** functions
- BUT simple non-convex function may have an exponential number of bad optima, e.g.

$$f(t_1) + \ldots + f(t_n)$$

**Hilbert's approach**

- **Globally** decompose the function into simple pieces
- Lasserre's hierarchy algorithm: **find decomposition efficiently**
- Strongest provable guarantees for **LOTS** of ML problems

 Yohann DE CASTRO

# Global Optimization

$$\min_{t\in\Omega} f(t) = \min_{\mu\geq 0\,,\ \mu(\Omega)=1} \int_\Omega f(t)\mu(\mathrm{d}t)$$

# Global Optimization

$$\min_{t\in\Omega} f(t) = \min_{\mu\geq 0\,,\ \mu(\Omega)=1} \int_\Omega f(t)\mu(\mathrm{d}t)$$

**Framework**

$$\Omega = \Big\{ t \in \mathbb{R}^D \ :\ \sum g_{\alpha,m} t^\alpha \geq 0\,,\ 1 \leq m \leq M \Big\}$$

Semi-Algebraic

$$f(t) = \sum f_\alpha t^\alpha$$

Polynomial

# Global Optimization

$$\min_{t\in\Omega} f(t) = \min_{\mu\geq 0\,,\ \mu(\Omega)=1} \int_\Omega f(t)\mu(\mathrm{d}t)$$

**Framework**

$$\Omega = \left\{ t \in \mathbb{R}^D \ :\ \sum g_{\alpha,m} t^\alpha \geq 0\,,\ 1 \leq m \leq M \right\}$$

Semi-Algebraic

$$f(t) = \sum f_\alpha t^\alpha$$

Polynomial

$$\min_{t\in\Omega} \sum f_\alpha t^\alpha = \min_{\mu\geq 0\,,\ \mu(\Omega)=1} \sum f_\alpha \int_\Omega t^\alpha \mu(\mathrm{d}t) = \min_{(m_\alpha)\in\mathcal{M}} \sum f_\alpha m_\alpha$$

Yohann DE CASTRO

# Global Optimization

$$\min_{t\in\Omega} f(t) = \min_{\mu\geq 0\,,\ \mu(\Omega)=1} \int_\Omega f(t)\mu(\mathrm{d}t)$$

**Framework**

$$\Omega = \left\{t \in \mathbb{R}^D \ : \ \sum g_{\alpha,m} t^\alpha \geq 0\,,\ 1 \leq m \leq M\right\}$$

Semi-Algebraic

$$f(t) = \sum f_\alpha t^\alpha$$

Polynomial

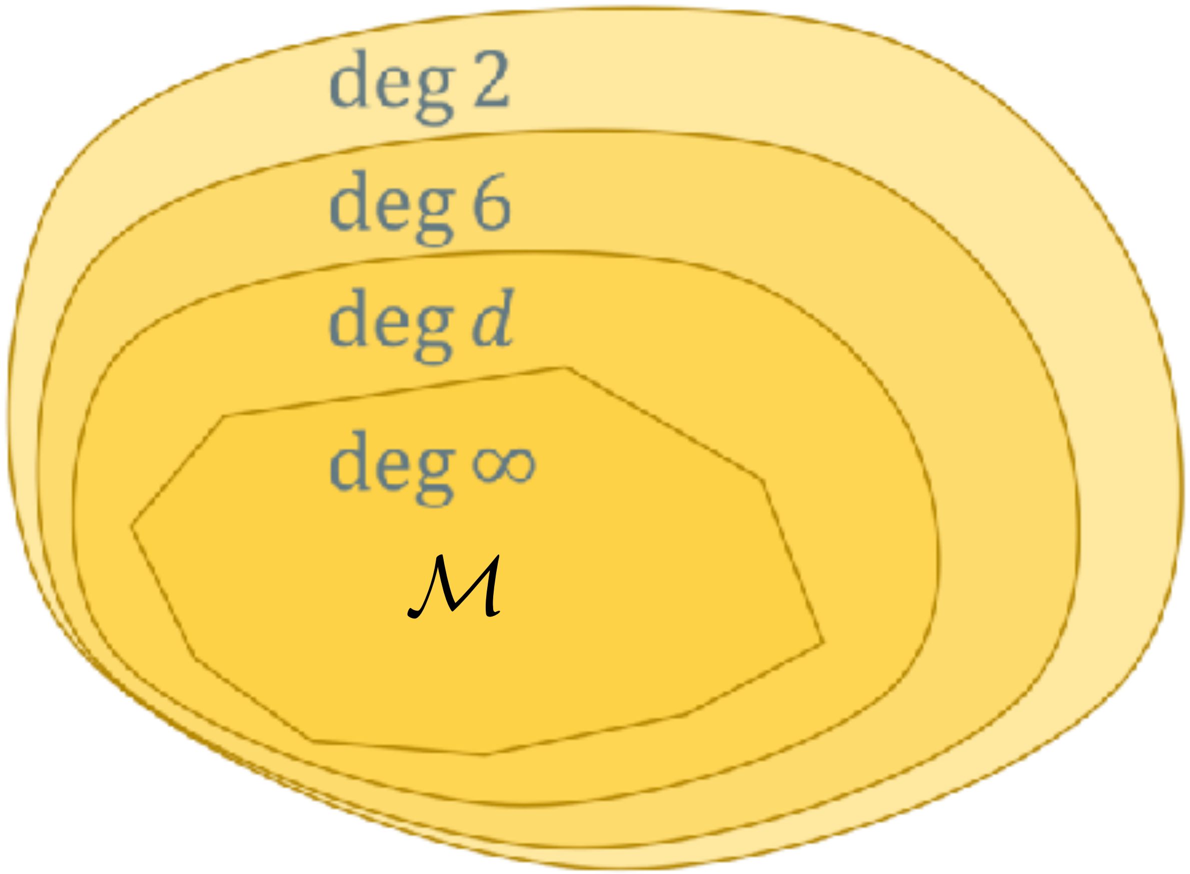
$$\min_{t\in\Omega} \sum f_\alpha t^\alpha = \min_{\mu\geq 0\,,\ \mu(\Omega)=1} \sum f_\alpha \int_\Omega t^\alpha \mu(\mathrm{d}t) = \min_{(m_\alpha)\in\mathcal{M}} \sum f_\alpha m_\alpha$$

Yohann DE CASTRO

$$\min_{t\in\Omega}\sum f_\alpha t^\alpha = \min_{\mu\geq 0\,,\ \mu(\Omega)=1}\sum f_\alpha \int_\Omega t^\alpha \mu(\mathrm{d}t) = \min_{(m_\alpha)\in\mathcal{M}}\sum f_\alpha m_\alpha$$

Yohann DE CASTRO

$$\mathcal{M} = \liminf_d \mathcal{M}^{\mathrm{SOS}(d)} \subseteq \ldots \subseteq \mathcal{M}^{\mathrm{SOS}(d)} \subseteq \mathcal{M}^{\mathrm{SOS}(d-2)} \subseteq \ldots \subseteq \mathcal{M}^{\mathrm{SOS}(2)}$$

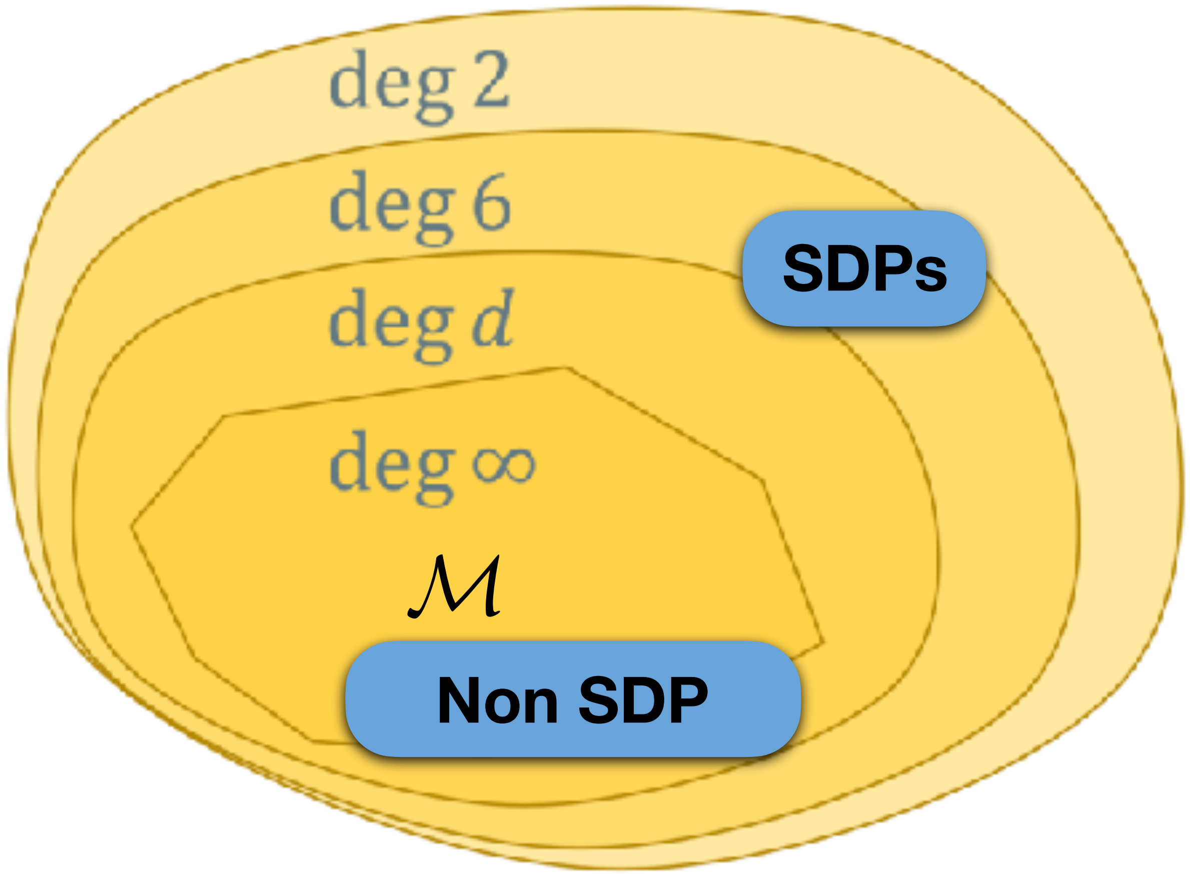



$$\min_{t\in\Omega} \sum f_\alpha t^\alpha = \min_{\mu\geq 0\,,\ \mu(\Omega)=1} \sum f_\alpha \int_\Omega t^\alpha \mu(\mathrm{d}t) = \min_{(m_\alpha)\in\mathcal{M}} \sum f_\alpha m_\alpha$$

Yohann DE CASTRO

$$\mathcal{M} = \liminf_d \mathcal{M}^{\mathrm{SOS}(d)} \subseteq \ldots \subseteq \mathcal{M}^{\mathrm{SOS}(d)} \subseteq \mathcal{M}^{\mathrm{SOS}(d-2)} \subseteq \ldots \subseteq \mathcal{M}^{\mathrm{SOS}(2)}$$

Lasserre's hierarchies give a sequence of nested OUTER **SDP approximations** of the cone of moments of nonnegative measures




$$\min_{t\in\Omega} \sum f_\alpha t^\alpha = \min_{\mu\geq 0\,,\ \mu(\Omega)=1} \sum f_\alpha \int_\Omega t^\alpha \mu(\mathrm{d}t) = \min_{(m_\alpha)\in\mathcal{M}} \sum f_\alpha m_\alpha$$

# Optimality of Lasserre's hierarchies?

**Under Khot's Unique Game Conjecture:**

NP-hard to achieve better approximation guarantees than Lasserre's hierarchies for class of worst-case discrete optimization problems [Khot'02, Raghavendra'08]

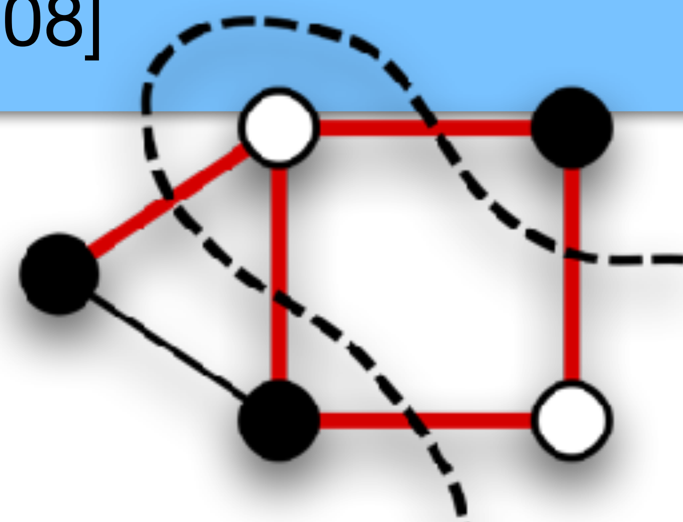
- Max Cut problem
- Stable Set problem
- Planted Clique problem
- Deg. 4 on the Sphere
- Tensor PCA
- **LOTS** of problems

Yohann DE CASTRO

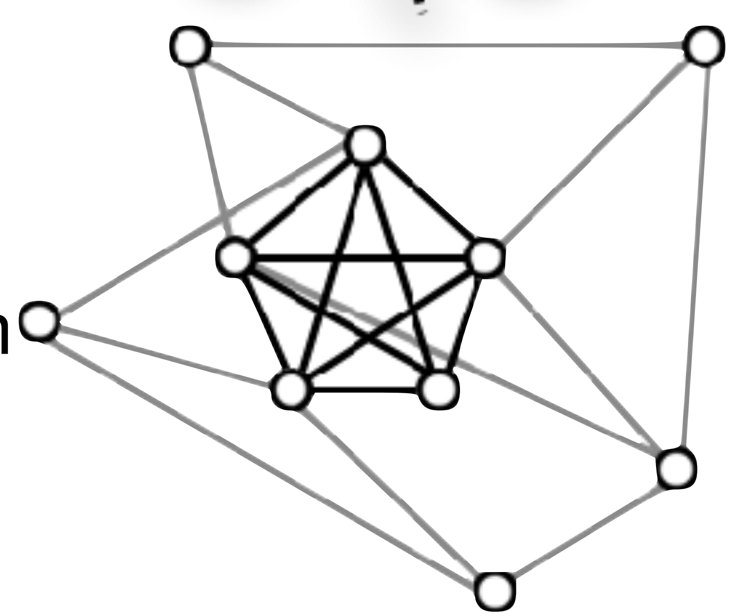
# Optimality of Lasserre's hierarchies?

**Under Khot's Unique Game Conjecture:**

NP-hard to achieve better approximation guarantees than Lasserre's hierarchies for class of worst-case discrete optimization problems [Khot'02, Raghavendra'08]



- Max Cut problem
- Stable Set problem
- Planted Clique problem



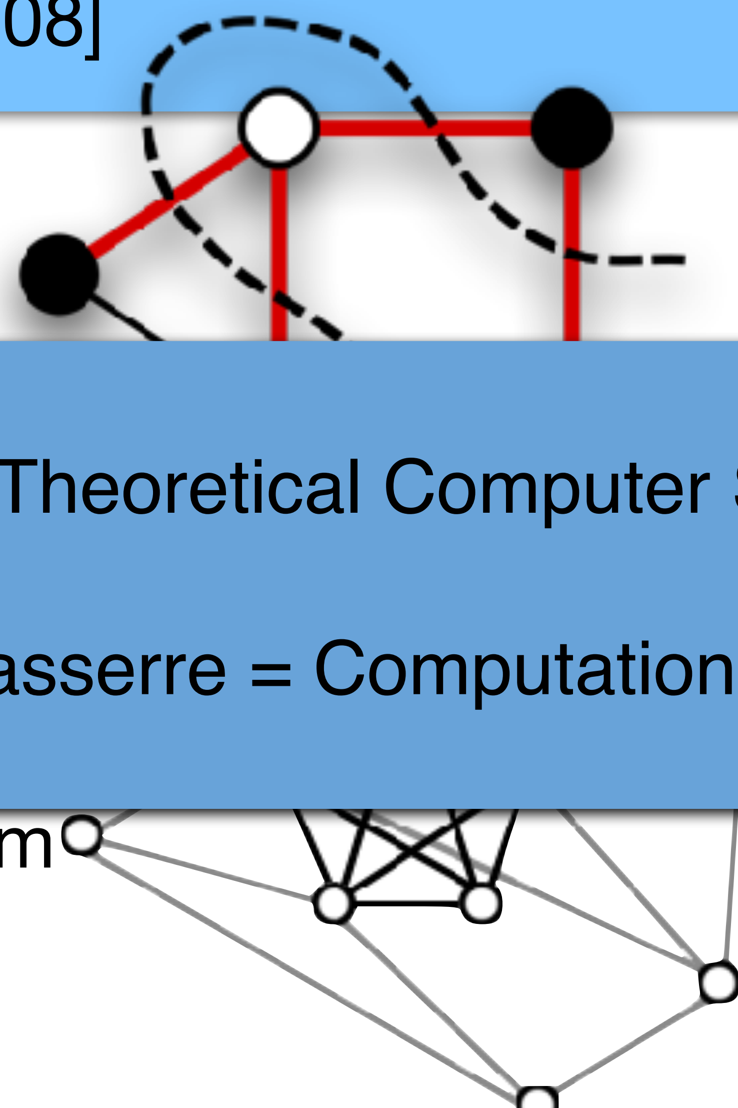
- Deg. 4 on the Sphere
- Tensor PCA
- **LOTS** of problems

Yohann DE CASTRO

# Optimality of Lasserre's hierarchies?

**Under Khot's Unique Game Conjecture:**

NP-hard to achieve better approximation guarantees than Lasserre's hierarchies for class of worst-case discrete optimization problems [Khot'02, Raghavendra'08]



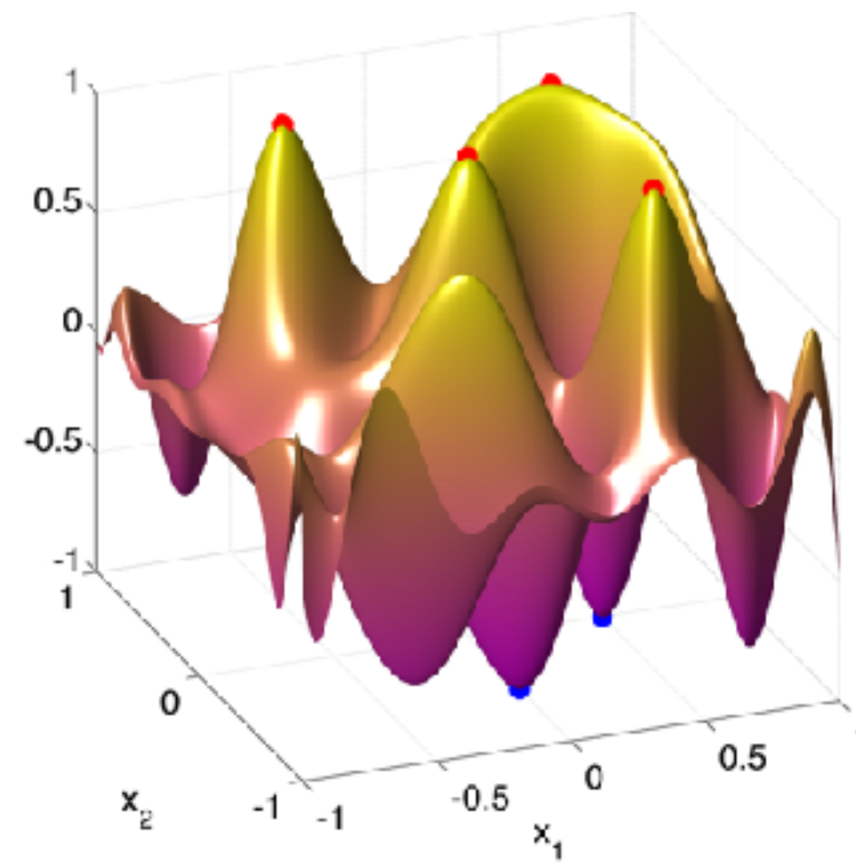
- Max Cut problem
- Stable Set probler
- Planted Clique problem

- Deg. 4 on the Sphere
- ensor PCA
- **LOTS** of problems

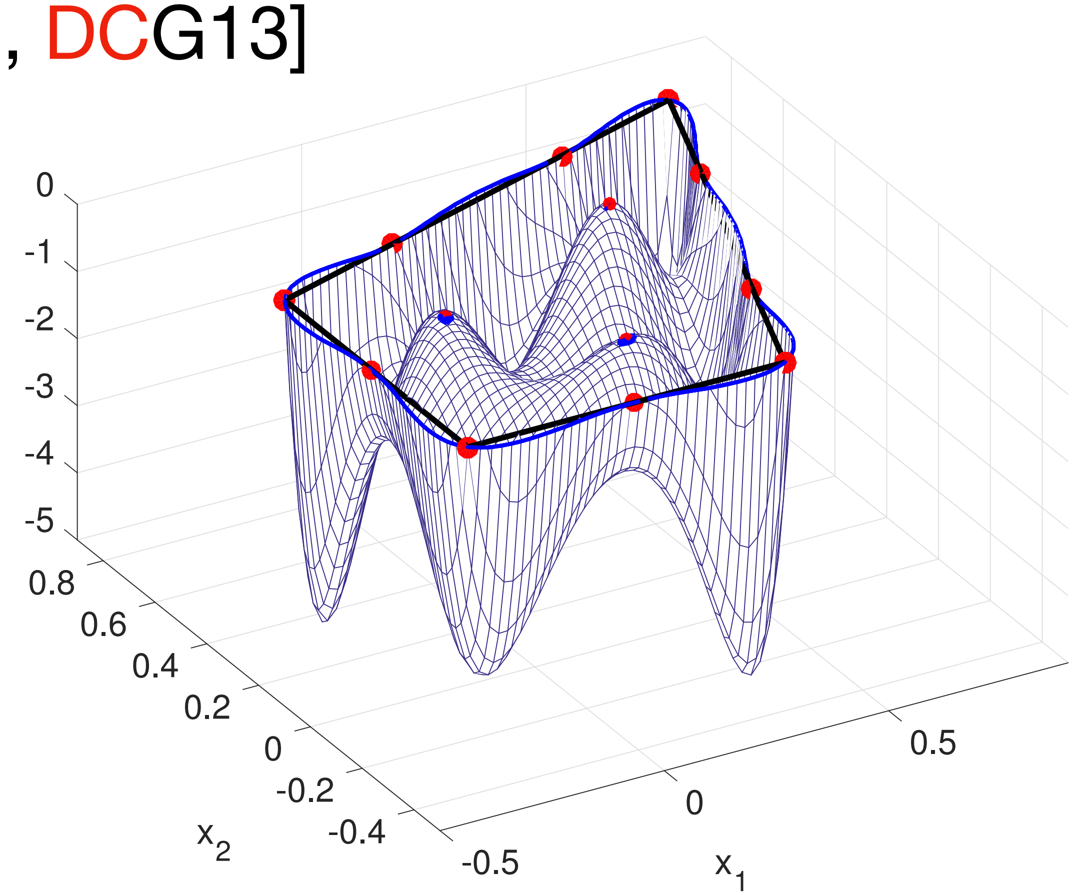
In Theoretical Computer Science

Lasserre = Computational Limit

# ML Applications




- Combinatorial Optimization on Graphs
- Tensor Decompositions
- Sparse Deconvolution [DCGHL16, ADCG14, DCG13]
- Gaussian Mixtures [DCGMM18+]
- Optimal Designs [DCGHHL18]




Step by Step
Sketch of Proof
in Convex Duality

# Convex Objective

$$\inf_{(m_\alpha)\in\mathcal{M}} \sum_{\alpha=0}^{s-1} f_\alpha m_\alpha \longrightarrow \inf_{(m_\alpha)\in\mathcal{M}} F\Big(\sum_{\alpha=0}^{s-1} m_\alpha H_\alpha\Big)$$

- Convex objective: $F$
- Symmetric matrices: $H_0, \ldots, H_{s-1}$
- Moment matrix: $M_s(m_\alpha) : (m_\alpha) \mapsto \sum_{\alpha=0}^{s-1} m_\alpha H_\alpha$

# Convex Objective

$$\inf_{(m_\alpha)\in\mathcal{M}} \sum_{\alpha=0}^{s-1} f_\alpha m_\alpha \longrightarrow \inf_{(m_\alpha)\in\mathcal{M}} F(\sum_{\alpha=0}^{s-1} m_\alpha H_\alpha)$$

- Convex objective: $F$
- Symmetric matrices: $H_0, \ldots, H_{s-1}$
- Moment matrix: $M_s(m_\alpha) : (m_\alpha) \mapsto \sum_{\alpha=0}^{s-1} m_\alpha H_\alpha$

# Convex Objective

$$\inf_{(m_\alpha)\in\mathcal{M}} \sum_{\alpha=0}^{s-1} f_\alpha m_\alpha \longrightarrow \inf_{(m_\alpha)\in\mathcal{M}} F\Big(\sum_{\alpha=0}^{s-1} m_\alpha H_\alpha\Big)$$

- Convex objective: $F$
- Symmetric matrices: $H_0, \dots, H_{s-1}$
- Moment matrix: $M_s(m_\alpha) : (m_\alpha) \mapsto \sum_{\alpha=0}^{s-1} m_\alpha H_\alpha$

**Primal**

$$\inf_{(m_\alpha)} \left\{ F(M_s(m_\alpha)) \;:\; m = (m_\alpha) \in \mathcal{M},\; Am = b \right\}.$$

Yohann DE CASTRO

# Fenchel Dual

$$F^\star(W) := \sup_{X \in \mathbb{S}_d} \left\{ \langle X, W \rangle - F(X) \right\}$$

**Key Property 1**

$$\forall W\,,\ \exists X_W \ s.t.\ \nabla F(X_W) = W\,, \quad F^\star(W) = \langle X_W, W \rangle - F(X_W)$$

$$\forall X\,,\ \exists W_X \ s.t.\ \nabla F^\star(W_X) = X\,, \quad F(X) = \langle X, W_X \rangle - F^\star(W)$$

# Fenchel Dual

$$F^\star(W) := \sup_{X\in\mathbb{S}_d} \left\{\langle X, W\rangle - F(X)\right\}$$

**Key Property 1**

$$\forall W\,,\ \exists X_W\ s.t.\ \nabla F(X_W) = W\,, \quad F^\star(W) = \langle X_W, W\rangle - F(X_W)$$

$$\forall X\,,\ \exists W_X\ s.t.\ \nabla F^\star(W_X) = X\,, \quad F(X) = \langle X, W_X\rangle - F^\star(W_X)$$

Yohann DE CASTRO

# Dual Cone

$$\mathcal{M}^\star := \Big\{ c \in \mathrm{R}^s \ : \ \forall m = (m_\alpha)_{\alpha=0}^{s-1} \in \mathcal{M}\,,\ \langle m, c\rangle \geq 0 \Big\}$$

**Key Property 2**

$$c \in \mathcal{M}^\star \Leftrightarrow \sum_\alpha c_\alpha t^\alpha \geq 0 \text{ on } \Omega$$

$$\mathcal{M}^\star = \limsup_d \mathcal{M}^{\star,\mathrm{SOS}(d)} \supseteq \ldots \supseteq \mathcal{M}^{\star,\mathrm{SOS}(d)} \supseteq \ldots \supseteq \mathcal{M}^{\star,\mathrm{SOS}(2)}$$

**INNER SDP approximations of nonnegative polynomials**

Yohann DE CASTRO

# Dual Cone

$$\langle m, c\rangle = \int_\Omega \big(\sum c_\alpha t^\alpha\big) \mathrm{d}\mu(t)$$

$$\mathcal{M}^\star := \Big\{ c \in \mathrm{R}^s \ : \ \forall m = (m_\alpha)_{\alpha=0}^{s-1} \in \mathcal{M}\,,\ \langle m, c\rangle \geq 0 \Big\}$$

**Key Property 2**

$$c \in \mathcal{M}^\star \Leftrightarrow \sum_\alpha c_\alpha t^\alpha \geq 0 \text{ on } \Omega$$

$$\mathcal{M}^\star = \limsup_d \mathcal{M}^{\star,\mathrm{SOS}(d)} \supseteq \ldots \supseteq \mathcal{M}^{\star,\mathrm{SOS}(d)} \supseteq \ldots \supseteq \mathcal{M}^{\star,\mathrm{SOS}(2)}$$

**INNER SDP approximations of nonnegative polynomials**

 Yohann DE CASTRO

# Dual Cone

$$\langle m, c\rangle = \int_\Omega (\sum c_\alpha t^\alpha)\mathrm{d}\mu(t)$$

$$\mathcal{M}^\star := \left\{ c \in \mathrm{R}^s \;:\; \forall m = (m_\alpha)_{\alpha=0}^{s-1} \in \mathcal{M}\,,\; \langle m, c\rangle \geq 0 \right\}$$

**Key Property 2**

$$c \in \mathcal{M}^\star \Leftrightarrow \sum_\alpha c_\alpha t^\alpha \geq 0 \text{ on } \Omega$$

$$\mathcal{M}^\star = \limsup_d \mathcal{M}^{\star,\mathrm{SOS}(d)} \supseteq \ldots \supseteq \mathcal{M}^{\star,\mathrm{SOS}(d)} \supseteq \ldots \supseteq \mathcal{M}^{\star,\mathrm{SOS}(2)}$$

**INNER SDP approximations of nonnegative polynomials**

UNIVERSITÉ PARIS SUD Inria

# Lagrangian Analysis

**Primal**

$$\inf_{(m_\alpha)} \left\{ F(M_s(m_\alpha)) \ : \ m = (m_\alpha) \in \mathcal{M},\ \ Am = b \right\}.$$

$$\mathcal{L} = F(X) - \langle W, X - M_s(m)\rangle + \langle z, Am - b\rangle - \langle m, c\rangle \tag{1}$$

$$= F(X) - \langle W, X\rangle - \langle z, b\rangle + \sum_{\alpha=0}^{s-1} m_\alpha \big(\langle W, H_\alpha\rangle + (A^\top z)_\alpha - c_\alpha\big) \tag{2}$$

(1) →

$$\inf_{X,m} \sup_{W,c,z} \mathcal{L} =$$ **Primal**

# Lagrangian Analysis

**Primal**

$$\inf_{(m_\alpha)} \left\{ F(M_s(m_\alpha)) \ : \ m = (m_\alpha) \in \mathcal{M}, \ Am = b \right\}.$$

$$\mathcal{L} = F(X) - \langle W, X - M_s(m) \rangle + \langle z, Am - b \rangle - \langle m, c \rangle \quad (1)$$

$$= F(X) - \langle W, X \rangle - \langle z, b \rangle + \sum_{\alpha=0}^{s-1} m_\alpha \big( \langle W, H_\alpha \rangle + (A^\top z)_\alpha - c_\alpha \big) \quad (2)$$

(1) → $\inf_{X,m} \sup_{W,c,z} \mathcal{L} =$ **Primal**

# Lagrangian Analysis

**Dual**

$$\sup_{W\in\mathbb{S}_d,z\in\mathbb{R}^m} \left\{\langle z,b\rangle - F^\star(W) \;:\; (\langle W,H_\alpha\rangle)_\alpha + A^\top z \in \mathcal{M}^\star\right\}.$$

$$\mathcal{L} = F(X) - \langle W, X - M_s(m)\rangle + \langle z, Am - b\rangle - \langle m, c\rangle \qquad (1)$$

$$= F(X) - \langle W, X\rangle - \langle z, b\rangle + \sum_{\alpha=0}^{s-1} m_\alpha\big(\langle W, H_\alpha\rangle + (A^\top z)_\alpha - c_\alpha\big) \qquad (2)$$

(2) → $\sup_{W,c,z}\inf_{X,m}\mathcal{L} =$ **Dual**

**Primal**

$$\inf_{(m_\alpha)} \left\{ F(M_s(m_\alpha)) \ : \ m = (m_\alpha) \in \mathcal{M}, \ \ Am = b \right\}.$$

**Dual**

$$\sup_{W \in \mathbb{S}_d, z \in \mathbb{R}^m} \left\{ \langle z, b \rangle - F^\star(W) \ : \ (\langle W, H_\alpha \rangle)_\alpha + A^\top z \in \mathcal{M}^\star \right\}.$$

**When no duality gap:** $\langle z^\star, b \rangle - F^\star(W^\star) = F(M_s(m^\star))$

**Complementarity leads to:**

$$\langle W^\star, M_s(m^\star) \rangle - \langle z^\star, b \rangle = \sum_{\alpha=0}^{s-1} m^\star_\alpha (\langle W^\star, H_\alpha \rangle + (A^\top z^\star)_\alpha) = 0$$

$$\langle W^\star, M_s(m^\star) \rangle = \langle z^\star, b \rangle = F(M_s(m^\star)) + F^\star(W^\star) \quad \text{and} \quad W^\star = \nabla F(M_s(m^\star)).$$

**Primal**

$$\inf_{(m_\alpha)} \left\{ F(M_s(m_\alpha)) \ : \ m=(m_\alpha)\in\mathcal{M},\ \ Am=b \right\}.$$

**Dual**

$$\sup_{W\in\mathbb{S}_d, z\in\mathbb{R}^m} \left\{ \langle z,b\rangle - F^\star(W) \ : \ (\langle W,H_\alpha\rangle)_\alpha + A^\top z \in \mathcal{M}^\star \right\}.$$

**When no duality gap:** $\langle z^\star,b\rangle - F^\star(W^\star) = F(M_s(m^\star))$

**Complementarity leads to:**

$$\langle W^\star, M_s(m^\star)\rangle - \langle z^\star, b\rangle = \sum_{\alpha=0}^{s-1} m_\alpha^\star(\langle W^\star, H_\alpha\rangle + (A^\top z^\star)_\alpha) = 0$$

$$\langle W^\star, M_s(m^\star)\rangle = \langle z^\star, b\rangle = F(M_s(m^\star)) + F^\star(W^\star) \quad \text{and} \quad W^\star = \nabla F(M_s(m^\star)).$$

Yohann DE CASTRO

**Primal**

$$\inf_{(m_\alpha)} \left\{ F(M_s(m_\alpha)) \ : \ m = (m_\alpha) \in \mathcal{M}, \ Am = b \right\}.$$

**Dual**

$$\sup_{W \in \mathbb{S}_d, z \in \mathbb{R}^m} \left\{ \langle z, b \rangle - F^\star(W) \ : \ (\langle W, H_\alpha \rangle)_\alpha + A^\top z \in \mathcal{M}^\star \right\}.$$

**When no duality gap:** $\langle z^\star, b \rangle - F^\star(W^\star) = F(M_s(m^\star))$

**Complementarity leads to:**

$$\langle W^\star, M_s(m^\star) \rangle - \langle z^\star, b \rangle = \sum_{\alpha=0}^{s-1} m^\star_\alpha (\langle W^\star, H_\alpha \rangle + (A^\top z^\star)_\alpha) = 0$$

$$\langle W^\star, M_s(m^\star) \rangle = \langle z^\star, b \rangle = F(M_s(m^\star)) + F^\star(W^\star) \quad \text{and} \quad W^\star = \nabla F(M_s(m^\star)).$$

**Key Property 1**

Primal

$$\inf_{(m_\alpha)} \left\{ F(M_s(m_\alpha)) \ : \ m = (m_\alpha) \in \mathcal{M}, \ Am = b \right\}.$$

**Key Property 2**

Dual

$$\sup_{W \in \mathbb{S}_d, z \in \mathbb{R}^m} \left\{ \langle z, b \rangle - F^\star(W) \ : \ (\langle W, H_\alpha \rangle)_\alpha + A^\top z \in \mathcal{M}^\star \right\}.$$

When no duality gap: $\langle z^\star, b \rangle - F^\star(W^\star) = F(M_s(m^\star))$

Complementarity leads to:

$$\langle W^\star, M_s(m^\star) \rangle - \langle z^\star, b \rangle = \sum_{\alpha=0}^{s-1} m_\alpha^\star \left( \langle W^\star, H_\alpha \rangle + (A^\top z^\star)_\alpha \right) = 0$$

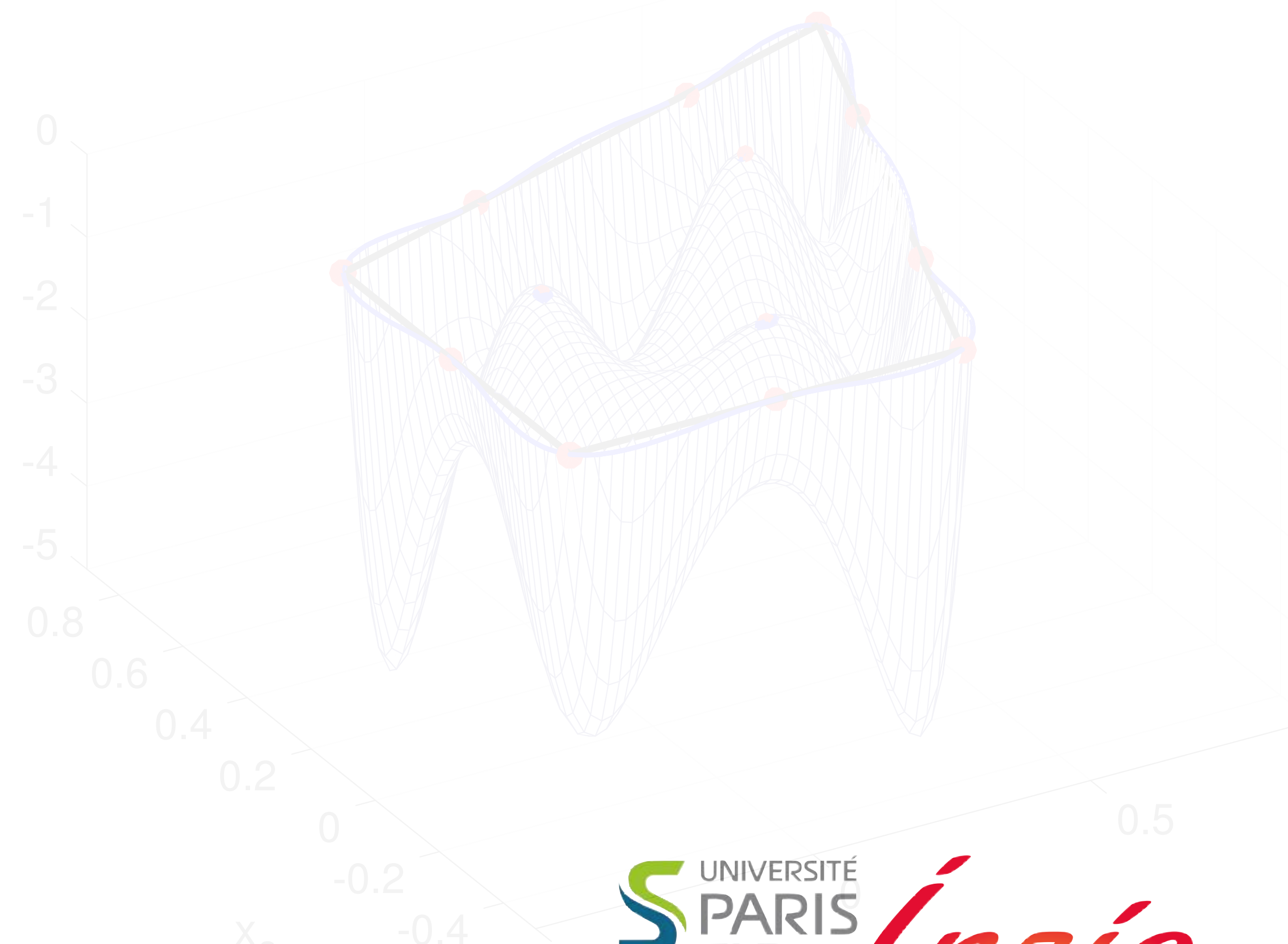
**Key Property 3**

$$\int_\Omega \underbrace{P(t)}_{\geq 0} \, \underbrace{\mathrm{d}\mu(t)}_{\geq 0} = 0$$

# An Example: Optimal Designs

$$\inf_{(m_\alpha)} \Big\{ -\log\det(M_s(m_\alpha)) \ : \ m = (m_\alpha) \in \mathcal{M},\ m_0 = 1 \Big\}\,.$$

- $F^\star(W) = -s - \log\det(-W)$ and $\nabla F(X) = -X^{-1}$
- $W^\star = -M_s(m^\star)^{-1}$
- $p^\star(t) = \sum_{\alpha=0}^{s-1} \langle M_s(m^\star)^{-1}, H_\alpha \rangle t^\alpha$
- $\int_\Omega (s - p^\star(t))\,\mathrm{d}\mu^\star(t) = 0$

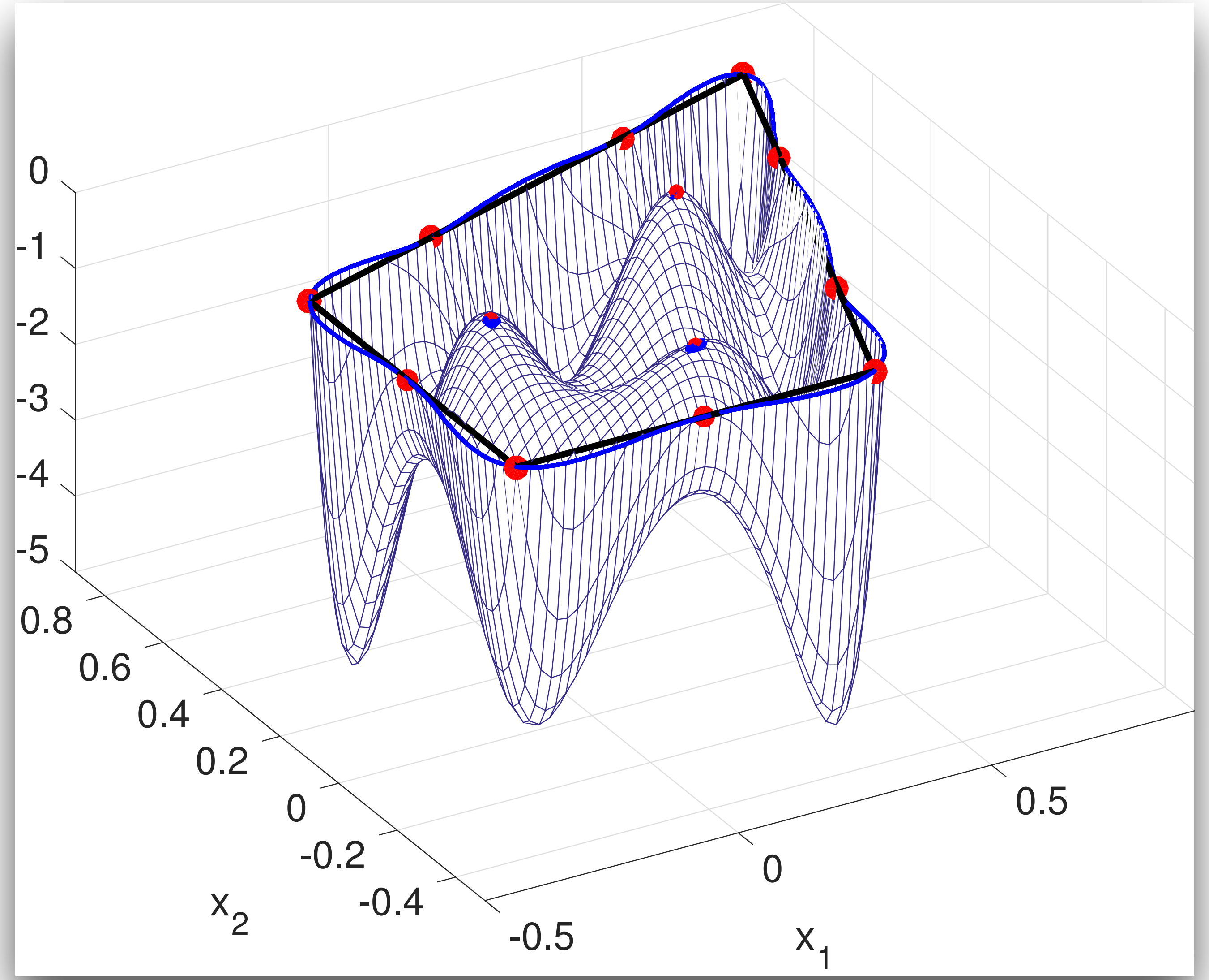


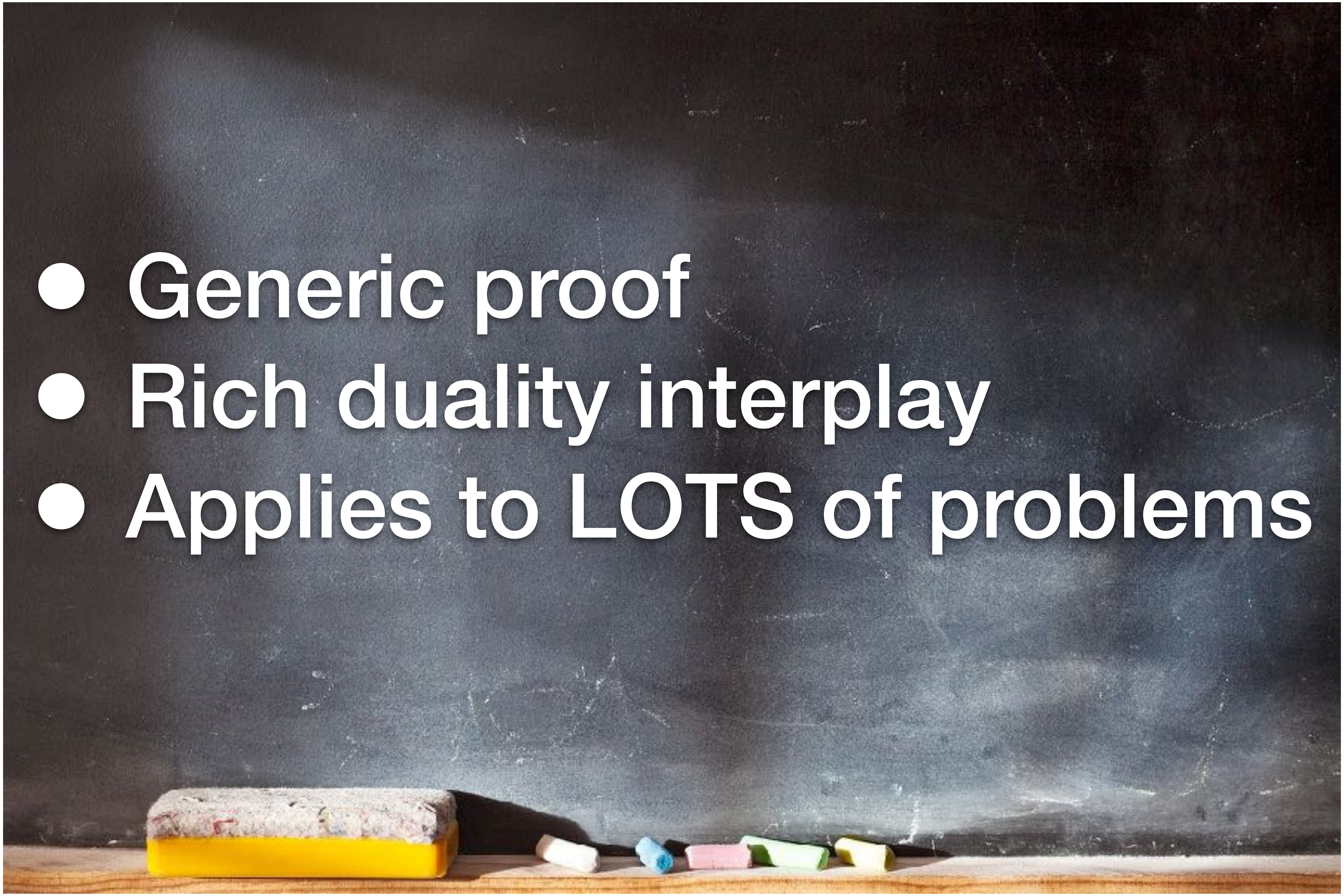
Yohann DE CASTRO

# An Example: Optimal Designs

$$\inf_{(m_\alpha)} \left\{ -\log\det(M_s(m_\alpha)) \; : \; m = (m_\alpha) \in \mathcal{M}, \; m_0 = 1 \right\}.$$

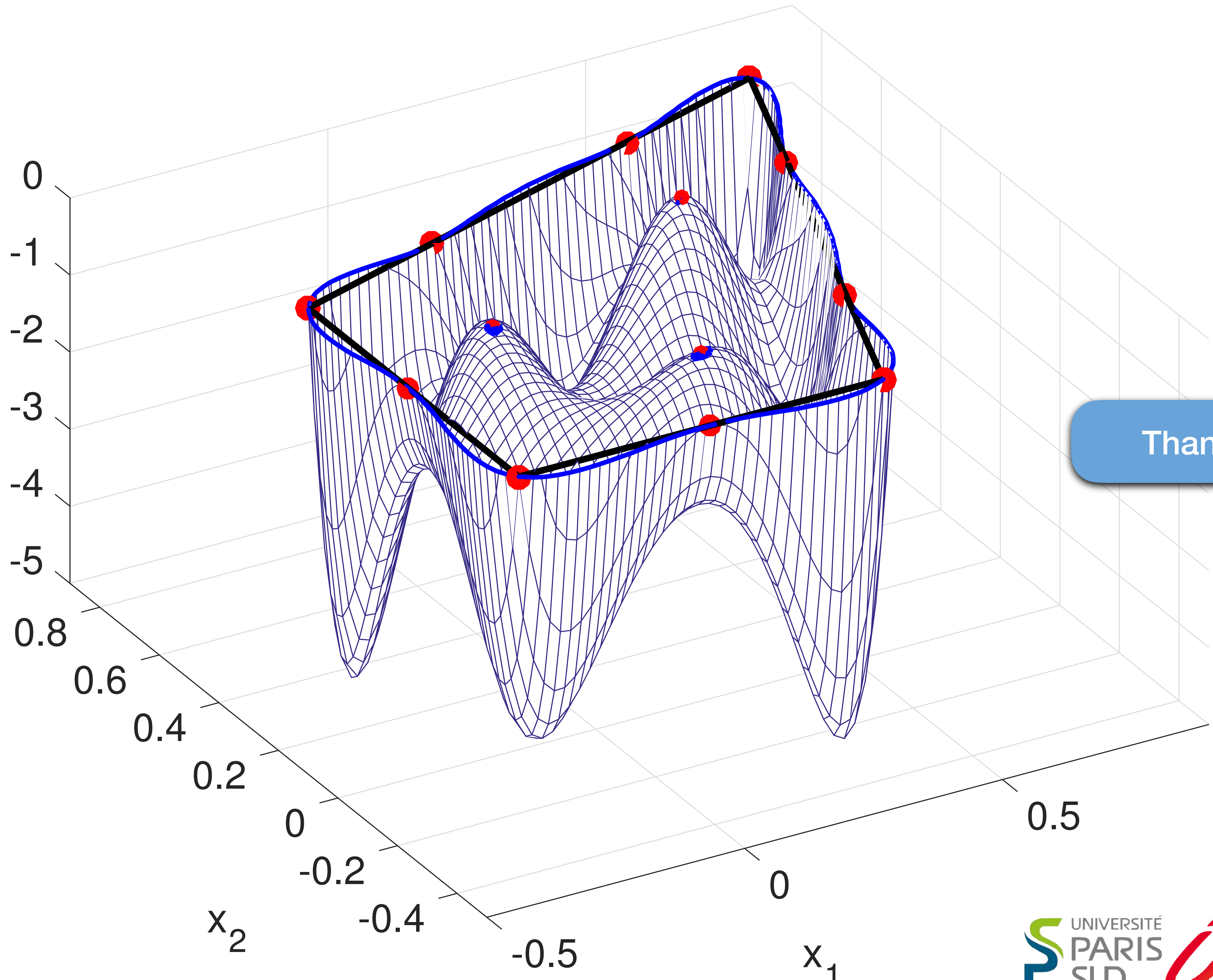
- $F^\star(W) = -s - \log\det(-W)$ and $\nabla F(X) = -X^{-1}$
- $W^\star = -M_s(m^\star)^{-1}$
- $p^\star(t) = \sum_\alpha \langle M_s(m^\star)^{-1}, H_\alpha \rangle t^\alpha$
- $\int_\Omega (s - p^\star(t)) \mathrm{d}\mu^\star(t) = 0$






UNIVERSITÉ PARIS SUD Inria

Yohann DE CASTRO




Thank You