Constrained Image Restoration Problems

Gabriele Steidl
University of Kaiserslautern
Germany

joint work with

R. Ciak, T. Teuber (University of Kaiserslautern)
B. Shafei (Fraunhofer ITWM, Kaiserslautern)
R. H. Chan (Chinese University of Hongkong)
J.-C. Pesquet (University of Paris-Est)

Outline

- $\Gamma_0(\mathbb{R}^n)$ set of proper, convex, lower semicontinuous functions

For $\Phi, \Psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ in $\Gamma_0(\mathbb{R}^n)$ consider

$$(P_{1,\tau}) \qquad \operatorname*{argmin}_{x \in \mathbb{R}^n} \{\Phi(x) \quad \text{subject to} \quad \Psi(x) \le \tau\}$$

$$(P_{2,\lambda}) \qquad \operatorname*{argmin}_{x \in \mathbb{R}^n} \{\Phi(x) + \lambda \Psi(x)\}, \qquad \lambda \ge 0$$

- $(P_{1,\tau})$: Ivanov regularization
  (in general τ can be estimated, but the problem maybe hard to solve numerically)
- $(P_{2,\lambda})$: Tihkonov regularization
  (in general the problem can be solved numerically, but the choice of λ is not straightforward)

1. Classical Example Find the minimizer $\hat{x} \in \mathbb{R}^n$ of

$$\|x\|_2^2 \quad \text{s.t.} \quad \|Hx - b\|_2^2 \le \tau \qquad \text{or} \qquad \|x\|_2^2 + \lambda \|Hx - b\|_2^2$$

Morozov's discrepancy principle (Morozov 1984 ...):

- Solution of the penalized problem is given analytically

$$\hat{x}(\lambda) = (\lambda H^\top H + I)^{-1} \lambda H^\top b$$

- Substituting this solution into the constraint gives

$$f(\lambda) \ = \ \|H\hat{x}(\lambda) - b\|_2^2 \ = \ b^\top (\lambda H H^\top + I)^{-2} b = \tau$$

  - $f(\lambda)$ convex, strictly decreasing function
  - $f(\lambda) = \tau$ has a unique solution $\hat{\lambda}$ for all $\|b_0\|_2^2 < \tau \le \|b\|_2^2$
  - Efficient computation of $\hat{\lambda}$ e.g., by Newton's method.

- Computation of $\hat{x}(\hat{\lambda})$

2. Examples from Image Restoration

- Gaussian noise:

$$\frac{1}{2}\|Hx-b\|_2^2 \quad \text{s.t.} \quad \|\nabla x\|_{2,1} \le \tau \qquad \text{or} \qquad \frac{1}{2}\|Hx-b\|_2^2 + \lambda \underbrace{\|\nabla x\|_{2,1}}_{discrete\ TV; ROF}$$

$$\|\nabla x\|_{2,1} \quad \text{s.t.} \quad \|Hx-b\|_2^2 \le \tau \qquad \text{or} \qquad \|\nabla x\|_{2,1} + \lambda\|Hx-b\|_2^2$$

- Poisson or multiplicative noise:

$$D(b,Hx) \quad \text{s.t.} \quad \|\nabla x\|_{2,1} \le \tau \qquad \text{or} \qquad \underbrace{D(b,Hx)}_{I-divergence} + \lambda\|\nabla x\|_{2,1}$$

$$\|\nabla x\|_{2,1} \quad \text{s.t.} \quad D(b,Hx) \le \tau \qquad \text{or} \qquad \|\nabla x\|_{2,1} + \lambda D(b,Hx)$$

- I-divergence = Kullback-Leibler divergence

References (constrained problem): van den Berg/Friedlander 2008, Fadili/Peyre 2011, Ng/Weiss/Yuan 2010, R. Chan/Wen 2012, Carlavan/Blanc-Feraud 2012, Friedlander et al. 2012

Outline

1. Introduction
2. General Remarks on Lagrange Multipliers
3. Homogeneous Constraints
4. Pareto Frontiers
5. I-Divergence and Anscombe Transform Constraints
6. Summary and Conclusions

Problems

$$(P_{1,\tau}) \qquad \operatorname*{argmin}_{x\in\mathbb{R}^n} \{\Phi(x) \quad \text{subject to} \quad \Psi(x) \le \tau\}$$

$$(P_{2,\lambda}) \qquad \operatorname*{argmin}_{x\in\mathbb{R}^n} \{\Phi(x) + \lambda\Psi(x)\}, \qquad \lambda \ge 0$$

Theorem (Relation between λ and τ)

Let $\Phi, \Psi \in \Gamma_0(\mathbb{R}^n)$. Assume some regularity assumptions.

i) Let $\hat{x}_\lambda$ be a minimizer of $(P_{2,\lambda})$, $\lambda > 0$.
Then $\hat{x}_\lambda$ is also a minimizer of $(P_{1,\tau})$ for $\tau := \Psi(\hat{x}_\lambda)$.
Moreover, this τ is unique if and only if $\hat{x}_\lambda$ is not a minimizer of Φ.

ii) Let $\hat{x}_\tau \in \operatorname{int}(\operatorname{dom}\Psi)$ be a minimizer of $(P_{1,\tau})$.
If $\hat{x}_\tau$ is not a minimizer of Ψ, Φ, then there exists a parameter $\lambda > 0$ such that $\hat{x}_\tau$ is also a minimizer of $(P_{2,\lambda})$.

Remark. i) does not mean that to each λ there exists a unique τ. If $\hat{x}_\lambda$ and $\tilde{x}_\lambda$ are solutions of $(P_{2,\lambda})$, then $\Psi(\hat{x}_\lambda) \neq \Psi(\tilde{x}_\lambda)$ is possible.

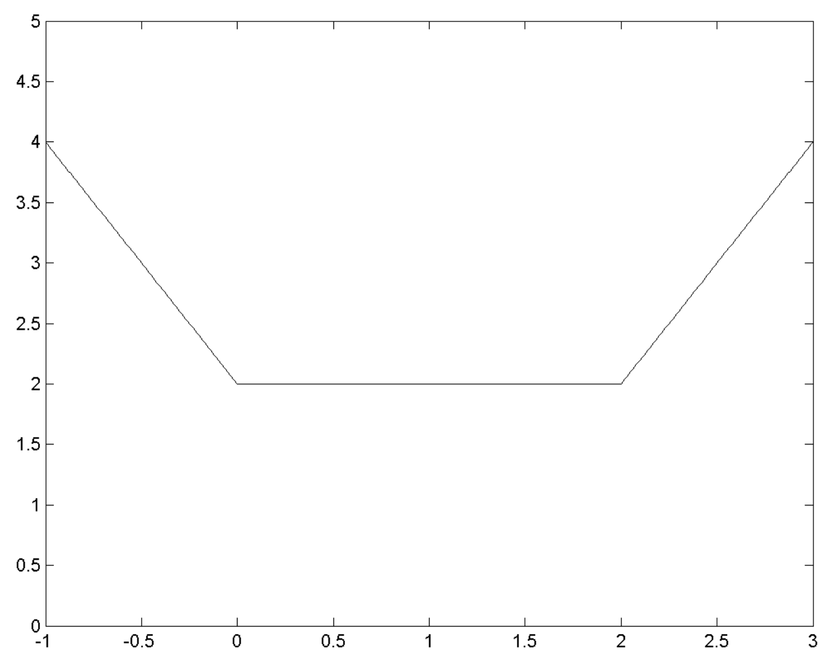
ii) does not mean that to each τ there exists a unique λ also if $\hat{x}_\tau$ is unique.

Example

$$\Phi(x) := |x-2|, \quad \Psi(x) := |x|, \quad \lambda := 1$$

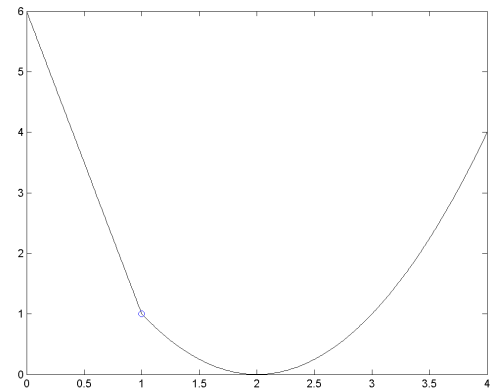
Then

$$(P_{2,1}) \qquad \Phi(x) + \Psi(x) = \begin{cases} -2(x-1) & \text{if } x < 0 \\ 2 & \text{if } x \in [0,2] \\ +2(x-1) & \text{if } x > 2 \end{cases}$$



Function $\Phi(x) + \Psi(x)$

Example $\Psi(x) := |x|, \quad \Phi(x) := \begin{cases} (x-2)^2 & \text{if } x \geq 1 \\ m(x-1)+1 & \text{if } x < 1 \end{cases}, \quad m \leq -2$



Function Φ for $m = -5$.

Thus, $\hat{x} = \tau$ is the solution of $(P_{1,\tau})$ for $\tau \in (0,2)$. On the other hand, we get

$$\operatorname*{argmin}_{x \in \mathbb{R}} \{\Phi(x) + \lambda |x|\} = \begin{cases} \{2 - \frac{\lambda}{2}\} & \text{if } \lambda \in [0,2) \\ \{1\} & \text{if } \lambda \in [2, -m) \\ [0,1] & \text{if } \lambda = -m \\ \{0\} & \text{if } \lambda \in (-m, +\infty) \end{cases}$$

Part ii) can be shown by the following lemma.

Lemma

Let $\Psi \in \Gamma_0(\mathbb{R}^n)$, $x^* \in \operatorname{dom}\Psi$, $\alpha := \Psi(x^*)$ and $S := \operatorname{lev}_\alpha \Psi = \{x \in \mathbb{R}^n : \Psi(x) \leq \alpha\}$. Then we have

$$\mathbb{R}_+ \, \partial\Psi(x^*) \subseteq \partial\iota_S(x^*) \; = \; N_S(x^*).$$

Moreover, if $x^* \in \operatorname{int}(\operatorname{dom}\Psi)$ and x^* is not a minimizer of Ψ, then equality holds true.

Proof via epigraphical projection (inf-projection) $\nu(x) := \inf_u \varphi(x,u)$

Example: Equality does in general not hold true for $x^* \in \operatorname{ri}(\operatorname{dom}\Psi)$:

$$\Psi(x_1,x_2) := \begin{cases} x_1 & \text{if } \; x_2 = 0 \\ +\infty & \text{if } \; x_2 \neq 0 \end{cases}$$

For

$$x^* := (x_1^*, 0) \in \operatorname{ri}(\operatorname{dom}\Psi), \qquad S := \operatorname{lev}_{x_1^*}\Psi = (-\infty, x_1^*) \times \{0\}$$

we get

$$\begin{aligned} \partial\iota_S(x^*) &= \{(p_1,p_2) : p_1 \in [0,+\infty), p_2 \in (-\infty,+\infty)\} \\ \partial\Psi(x^*) &= \{(1,p_2) : p_2 \in (-\infty,+\infty)\} \\ \mathbb{R}_+\partial\Psi(x^*) &= \{(0,0)\} \cup \{(p_1,p_2) : p_1 \in (0,+\infty), p_2 \in (-\infty,+\infty)\} \end{aligned}$$

Outline

Setting:

- $\Phi(x) = \Phi(x_1 + x_2) = \phi(x_1)$ (e.g. $\mathbb{R}^n = \mathcal{R}(H^*) \oplus \mathcal{N}(H)$), where $\phi : X_1 \to \mathbb{R} \cup \{+\infty\}$ is a function with

  i) $\operatorname{dom} \phi$ is an open subset of X_1 with $0 \in \overline{\operatorname{dom} \phi}$,

  ii) ϕ belongs to $\Gamma_0(X_1)$ and is strictly convex and essentially smooth,

  iii) ϕ has a minimizer

  Examples: $\Phi(x) = \|Hx - b\|_2^2$, $\Phi(x) = D(b, Hx)$

- $\Psi : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ be a positive homogeneous function, i.e.,

$$\Psi(\alpha x) = \alpha \Psi(x), \quad \alpha > 0$$

  Special homogeneous function are seminorms and indicator functions

$$\iota_S(x) := \begin{cases} 0 & \text{if } x \in S \\ +\infty & \text{otherwise} \end{cases}$$

- level sets: $\operatorname{lev}_\tau \Psi := \{x \in \mathbb{R}^n : \Psi(x) \le \tau\}$

The following holds true:

$$\Psi_1 := \iota_{lev_1 \|\cdot\|} \quad \text{and} \quad \Psi_2 = \|\cdot\|,$$
$$\Psi_1^* = \|\cdot\|_* \quad \text{and} \quad \Psi_2^* = \iota_{lev_1 \|\cdot\|_*}$$

where $\Psi^*(p) := \sup_x \{\|\langle p, x\rangle - \Phi(x)\}$ is the Fenchel dual of Ψ and $\|\cdot\|_*$ the dual norm of $\|\cdot\|$.

Primal Problems:

$$(P_{1,\tau}) \qquad \operatorname*{argmin}_{x \in \mathbb{R}^n} \{\Phi(x) \quad \text{subject to} \quad \|Lx\| \le \tau\}$$
$$(P_{2,\lambda}) \qquad \operatorname*{argmin}_{x \in \mathbb{R}^n} \{\Phi(x) \ + \ \lambda \|Lx\|\}$$

Fenchel Dual Problems:

$$(D_{1,\tau}) \qquad \operatorname*{argmin}_{p \in \mathbb{R}^m} \{\Phi^*(-L^* p) \ + \ \tau \|p\|_*\}$$
$$(D_{2,\lambda}) \qquad \operatorname*{argmin}_{p \in \mathbb{R}^m} \{\Phi^*(-L^* p) \quad \text{subject to} \quad \|p\|_* \le \lambda\}$$

KKT condition gives $\quad -L^* \hat{p} = \nabla \Phi(\hat{x})$

Theorem

Let $\mathcal{N}(H) \cap \mathcal{N}(L) = \{0\}$ and $\operatorname{argmin} \Phi \cap \mathcal{N}(L) = \emptyset$. Let

$$\tau_e := \min_{x \in \operatorname{argmin} \Phi} \|Lx\|, \qquad \lambda_e := \min_{p \in \operatorname{argmin} \Phi^*(-L^* \cdot)} \|p\|_*.$$

Then, for $\tau \in (0, \tau_e)$ and $\lambda \in (0, \lambda_e)$, the problems $(P_{1,\tau})$, $(P_{2,\lambda})$, $(D_{1,\tau})$, $(D_{2,\lambda})$ have solutions with finite minima. Further there exists a bijective mapping $g : (0, \tau_e) \to (0, \lambda_e)$ such that for $\tau \in (0, \tau_e)$ and $\lambda \in (0, \lambda_e)$ we have

$$\left\{ \begin{array}{c} \mathrm{SOL}(P_{1,\tau}) = \mathrm{SOL}(P_{2,\lambda}) \\ \mathrm{SOL}(D_{1,\tau}) = \mathrm{SOL}(D_{2,\lambda}) \end{array} \right\} \quad \text{if} \quad (\tau, \lambda) \in \operatorname{gr} g$$

For $(\tau, \lambda) \in \operatorname{gr} g$ any solutions $\hat{x}$ and $\hat{p}$ of the primal and dual problems, resp., fulfill

$$\tau = \|L\hat{x}\| \quad \text{and} \quad \lambda = \|\hat{p}\|_*.$$

- τ-λ **curve** $g : (0, \tau_e) \to (0, \lambda_e)$ is monotone decreasing and continuous

1. Joint Sparsity Problems $L := A^*$, $H := I$

Dual Problems:

$$(D_{2,\lambda}) \qquad \operatorname*{argmin}_{p\in\mathbb{R}^m} \left\{ \frac{1}{2}\|Ap-b\|_2^2 \quad \text{subject to} \quad \|p\|_{2,1} \le \lambda \right\},$$

$$(D_{1,\tau}) \qquad \operatorname*{argmin}_{p\in\mathbb{R}^m} \left\{ \frac{1}{2}\|Ap-b\|_2^2 \; + \; \tau\|p\|_{2,1} \right\}$$

where $A \in \mathbb{R}^{n,m}$, $n \ll m$ is some special matrix and $0 \neq b \in \mathcal{R}(A)$.

- $(D_{2,\lambda})$: LASSO (Tibshirani 1994)

Primal Problems: $\hat{x} = b - A\hat{p}$

$$(P_{2,\lambda}) \qquad \operatorname*{argmin}_{x\in\mathbb{R}^n} \left\{ \frac{1}{2}\|x-b\|_2^2 \; + \; \lambda\|A^*x\|_{2,\infty} \right\},$$

$$(P_{1,\tau}) \qquad \operatorname*{argmin}_{x\in\mathbb{R}^n} \left\{ \frac{1}{2}\|x-b\|_2^2 \quad \text{subject to} \quad \|A^*x\|_{2,\infty} \le \tau \right\}.$$

References: Daubechies/Fornasier/Loris 2008, Lorenz et al. 2011- , van den Berg/Friedlander 2008

Corollary. Let

$$\tau_e := \|A^* b\|_{2,\infty} \quad \text{and} \quad \lambda_e := \min_{Ap=b} \|p\|_{2,1}.$$

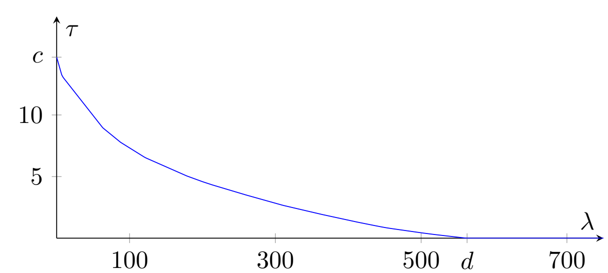
Then, for $\lambda \in (0, \lambda_e)$, a solution $\hat{p}$ of $(D_{2,\lambda})$ is also a solution of $(D_{1,\tau})$ iff

$$\tau = \| \underbrace{A^*(b - A\hat{p})}_{=:\hat{q}} \|_{2,\infty} = \|\hat{\mathbf{q}}_j\|_2 \quad \text{for all } \ j \ \text{ with } \ \hat{\mathbf{p}}_j \neq 0$$

where $\hat{\mathbf{q}}_j := (\hat{q}_{kn+j})_{k=0}^{\kappa-1}$.

Example. A obtained by taking 200 rows of the (1000,1000) DCT-II matrix

Computation by FBS (Forward-Backward Splitting), alternatively Nesterov's alg. 2005, FISTA (Beck/Teboulle 2009)



Function g^{-1} for the joint sparsity problem

2. Image Inpainting Model I H inpainting mask

$$(D_{2,\lambda}) \qquad \operatorname*{argmin}_{P \in \mathbb{R}^{N,M}} \left\{ \frac{1}{2} \|H(P) - B\|_F^2 \quad \text{subject to} \quad \|P\|_* \le \tau \right\}$$

$$(D_{1,\tau}) \qquad \operatorname*{argmin}_{P \in \mathbb{R}^{N,M}} \left\{ \frac{1}{2} \|H(P) - B\|_F^2 \; + \; \lambda \|P\|_* \right\}$$

with the nuclear norm $\|P\|_*$ of P

Reference (penalized problem): Liu et al. 2009

3. Image Inpainting Model II $L := \nabla$, H inpainting mask

$$(P_{1,\tau}) \qquad \operatorname*{argmin}_{x \in \mathbb{R}^n} \left\{ \frac{1}{2} \|Hx - b\|_2^2 \quad \text{subject to} \quad \|Lx\|_{2,1} \le \tau \right\}$$

$$(P_{2,\lambda}) \qquad \operatorname*{argmin}_{x \in \mathbb{R}^n} \left\{ \frac{1}{2} \|Hx - b\|_2^2 \; + \; \lambda \|Lx\|_{2,1} \right\}$$

Computation by ADMM (alternating direction of multipliers), PDHGMp (Arrow-Hurwitz method with extrapolation of the dual variable) (T. Chan/Zhou, Esser 2009, Chambolle/Pock 2011) and other algs. (Chen/Teboulle 1994, Combettes/Pesquet 2011, Vu 2012, Bot/Hendrich 2012)

3. Homogeneous Functions: Examples

Solutions of the TV-inpainting problem
Top: $(P_{1,\tau})$ for various values of λ; $\|P_{\mathrm{orig}}\|_*/\|\hat{P}\|_* \approx 1, 1.5$
Bottom: $(P_{1,\tau})$ for various values of τ; $TV(x_{\mathrm{orig}})/TV(\hat{x}) \approx 1, 2$

Scaled ADMM
Initialization: $q_1^{(0)} = q_2^{(0)} = 0$, $y_1^{(0)} = Hb$, $y_2^{(0)} = Lb$.
For $k = 0, \dots$ repeat until a stopping criterion is reached

$$
\begin{aligned}
x^{(k+1)} &= \operatorname*{argmin}_{x \in \mathbb{R}^n} \left\{ \|q_1^{(k)} + Hx - y_1^{(k)}\|_2^2 + \|q_2^{(k)} + Lx - y_2^{(k)}\|_2^2 \right\}, \\
y_1^{(k+1)} &= \operatorname*{argmin}_{y_1 \in \mathbb{R}^n} \left\{ \langle 1, y_1 - b \log(y_1) \rangle + \frac{\gamma}{2} \|q_1^{(k)} + Hx^{(k+1)} - y_1\|_2^2 \right\}, \\
y_2^{(k+1)} &= \operatorname*{argmin}_{y_2 \in \mathbb{R}^{2n}} \left\{ \iota_{B_{1,\tau}}(y_2) + \frac{\gamma}{2} \|q_2^{(k)} + Lx^{(k+1)} - y_2\|_2^2 \right\}, \\
q_1^{(k+1)} &= q_1^{(k)} + Hx^{(k+1)} - y_1^{(k+1)}, \\
q_2^{(k+1)} &= q_2^{(k)} + Lx^{(k+1)} - y_2^{(k+1)}.
\end{aligned}
$$

To obtain the curve g we have to compute

$$
\lambda = \|\hat{p}\|_{2,\infty} = \gamma \|\hat{q}_2\|_{2,\infty}
$$

Outline

1. Introduction
2. General Remarks on Lagrange Multipliers
3. Homogeneous Constraints
4. Pareto Frontiers
5. I-Divergence and Anscombe Transform Constraints
6. Summary and Conclusions

Multicriteria optimization (multiobjective optimization, vector optimization)

- F.Y. Edgeworth (1881), V. Pareto (1906)
- **Pareto optimum**: 'The optimum allocation of the resources of a society is not attained so long as it is possible to make at least one individual better off in his own estimation while keeping others as well off as before in their own estimation.'

Relation to our $\tau - \lambda$ problem?

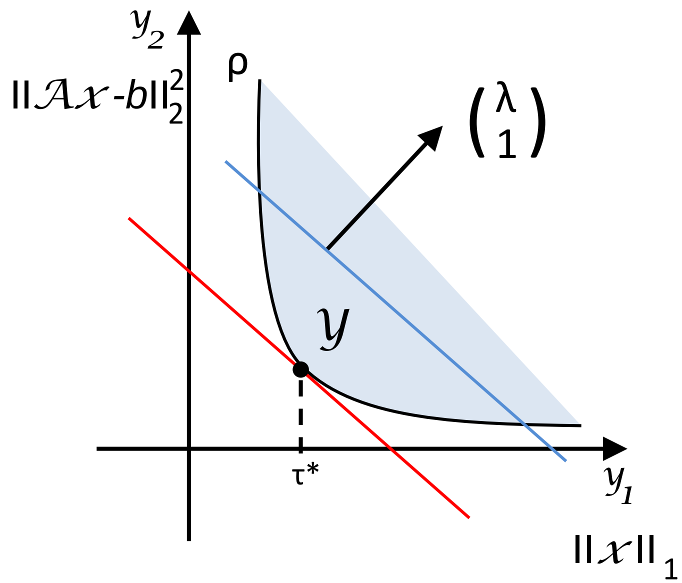
$$(P_{1,\tau}) \qquad \min_{x\in\mathbb{R}^n}\{\Phi(x) + \iota_{\mathrm{lev}_\tau\Psi}(x)\} = \min_{y\in Y}\{y_2 + \iota_\tau(y_1)\}$$

$$(P_{2,\lambda}) \qquad \min_{x\in\mathbb{R}^n}\{\underbrace{\Phi(x)}_{y_2} + \lambda\underbrace{\Psi(x)}_{y_1}\} = \min_{y\in Y}\left\langle \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \begin{pmatrix} \lambda \\ 1 \end{pmatrix}\right\rangle$$

where

$$Y := \left\{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \Psi(x) \\ \Phi(x) \end{pmatrix} : x\in\mathbb{R}^n\right\} \subset \mathbb{R}^2$$

- Consider Pareto frontier ρ of $\min_{y\in Y} y_1$ and $\min_{y\in Y} y_2$



Let $y^* = (\tau^*, \rho(\tau^*)) \in \operatorname{gr}\rho.$ Then

$$y^* \in \underset{y \in Y}{\operatorname{argmin}} \left\langle \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \begin{pmatrix} \lambda^* \\ 1 \end{pmatrix} \right\rangle \Leftrightarrow \begin{pmatrix} \lambda^* \\ 1 \end{pmatrix} \perp \begin{pmatrix} 1 \\ \rho'(\tau^*) \end{pmatrix}$$

$$\Leftrightarrow \lambda^* = -\rho'(\tau^*)$$

$$\Rightarrow \mathbf{g}(\tau) = -\rho'(\tau)$$

Our $\tau - \lambda$ curve g is the negative derivative of the Pareto curve ρ.

Outline

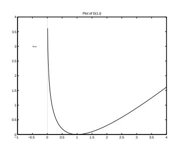
Problems

$$(P_{1,\tau}) \qquad \operatorname*{argmin}_{x \geq 0} \{\|Lx\| \quad \text{subject to} \quad D(b, Hx) \leq \tau\}$$

$$(P_{2,\lambda}) \qquad \operatorname*{argmin}_{x \geq 0} \{\|Lx\| + \lambda D(b, Hx)\}$$

with I-divergence

$$D(b,t) := \begin{cases} \langle 1_n, b \log \frac{b}{t} - b + t \rangle & \text{if } t > 0, \\ +\infty & \text{otherwise} \end{cases}$$



- Application in image restoration for images corrupted by Poisson noise (often medical images), (Bicous-Diaz, Figueiredo 2010, Brune et al. 2011) or multiplicative noise (SAR images) (Aubert,Aujol 2008, Sigelle et al. 2008, Shi,Osher 2011, St./Teuber 2010 ...)

Let

- $\mathcal{K} := \{x \in \mathbb{R}^n_{\geq 0} : Hx > 0\} \neq \emptyset$
- $\tau_0 = \min\limits_{x \geq 0} D(b, Hx)$
- $\tau_e = \min\limits_{x \geq 0, x \in \mathcal{N}(L)} D(b, Hx)$

Theorem

Let $\mathcal{N}(H) \cap \mathcal{N}(L) = \{0\}$ and $\mathcal{N}(L) \cap \operatorname{argmin}_{x \geq 0} D(b, H\cdot) = \emptyset$.
If $\hat{x}$ is a solution of $(P_{1,\tau})$ with $\tau \in (\tau_0, \tau_e)$, then there exists a unique $\lambda > 0$ such that $\hat{x}$ is also a solution of $(P_{2,\lambda})$.

Problems:

$$(P_{1,\tau}) \qquad \operatorname*{argmin}_{x\geq 0} \{\|Lx\| \quad \text{subject to} \quad D(b,Hx) \leq \tau\}$$

$$(P_{2,\lambda}) \qquad \operatorname*{argmin}_{x\geq 0} \{\|Lx\| + \lambda D(b,Hx)\}$$

Reformulation of the problems by **splitting technique**

$$(P_{1,\tau}) \quad \operatorname*{argmin}_{x,y_1,y_2,y_3} \left\{\iota_{\mathrm{lev}_\tau D(b,\cdot)}(y_1) + \|y_2\| + \iota_{y_3\geq 0}(y_3) \text{ s.t. } \begin{pmatrix} H \\ L \\ I \end{pmatrix} x = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}\right\}$$

$$(P_{2,\lambda}) \quad \operatorname*{argmin}_{x,y_1,y_2,y_3} \left\{\lambda D(b,y_1) + \|y_2\| + \iota_{y_3\geq 0}(y_3) \text{ s.t. } \begin{pmatrix} H \\ L \\ I \end{pmatrix} x = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}\right\}$$

References: Bioucas-Dias/Figueiredo 2009, Setzer/St./Teuber 2010 ...

Algorithm (ADMM for solving $(P_{1,\tau})$)
Initialization: $q_1^{(0)} = q_2^{(0)} = q_3^{(0)} = 0$, $y_1^{(0)} = Hb$, $y_2^{(0)} = Lb$, $y_3^{(0)} = b$ and $\gamma > 0$.
For $k = 0, 1, \ldots$ repeat until a stopping criterion is reached:

$$
\begin{aligned}
x^{(k+1)} &= \operatorname*{argmin}_{x \in \mathbb{R}^n} \{ \|q_1^{(k)} + Hx - y_1^{(k)}\|_2^2 + \|q_2^{(k)} + Lx - y_2^{(k)}\|_2^2 + \|q_3^{(k)} + x - y_3^{(k)}\|_2^2 \} \\
y_1^{(k+1)} &= \operatorname*{argmin}_{y_1 \in \mathbb{R}^n} \{ \iota_{\mathrm{lev}_\tau D(b,\cdot)}(y_1) + \frac{\gamma}{2} \| \underbrace{q_1^{(k)} + Hx^{(k+1)}}_{a^{(k+1)}} - y_1 \|_2^2 \} \\
y_2^{(k+1)} &= \operatorname*{argmin}_{y_2 \in \mathbb{R}^m} \{ \|y_2\| + \frac{\gamma}{2} \|q_2^{(k)} + Lx^{(k+1)} - y_2\|_2^2 \} \\
y_3^{(k+1)} &= \operatorname*{argmin}_{y_3 \in \mathbb{R}^n} \{ \iota_{y_3 \geq 0}(y_3) + \frac{\gamma}{2} \|q_3^{(k)} + x^{(k+1)} - y_3\|_2^2 \} \\
q_1^{(k+1)} &= q_1^{(k)} + Hx^{(k+1)} - y_1^{(k+1)} \\
q_2^{(k+1)} &= q_2^{(k)} + Lx^{(k+1)} - y_2^{(k+1)} \\
q_3^{(k+1)} &= q_3^{(k)} + x^{(k+1)} - y_3^{(k+1)}
\end{aligned}
$$

- Subproblem 2 is an I-divergence constrained least squares problem

Subproblem 2 reads

$$\operatorname*{argmin}_{y\in\mathbb{R}^n}\{\frac{1}{2}\|y-a\|_2^2 \quad \text{subject to} \quad D(b,y)\le\tau\}$$

Solution via corresponding penalized problem

$$\operatorname*{argmin}_{y\in\mathbb{R}^n}\{\frac{1}{2}\|y-a\|_2^2+\lambda D(b,y)\}$$

Theorem

For $\lambda>0$ and $a\neq b$ the penalized problem has the analytical solution

$$\hat{y}(\lambda) = \frac{1}{2}\left(y-\lambda+\sqrt{(a-\lambda)^2+4\lambda b}\right).$$

The function

$$f(\lambda) := D(b,\hat{y}(\lambda))(=\tau)$$

is strictly decreasing and convex.

Theorem (Convergence of the Algorithm)

Let $\mathcal{N}(L) \cap \mathcal{N}(H) = \{0\}$ and $\operatorname{argmin}_{x \geq 0} D(b, Hx) \cap \mathcal{N}(L) = \emptyset$. Let $\tau \in (\tau_0, \tau_e)$.

Then the sequence $\{(x^{(k)}, y^{(k)}, q^{(k)}, \lambda_k)\}_k$ generated by our ADMM algorithm with inner Newton step converges to $(\hat{x}, \hat{y}, \hat{q}, \hat{\lambda})$, where

- $\hat{x}$ is a solution of $(P_{1,\tau})$ and $(P_{2,\hat{\lambda}})$, where $\lim_{k \to \infty} \lambda_k = \hat{\lambda}$,
- $\hat{p} = \gamma \hat{q}$ is a solution of the dual problems $(D_{1,\tau})$ and $(D_{2,\hat{\lambda}})$,
- $\hat{y} = (H^\mathsf{T}\, L^\mathsf{T}\, I)^\mathsf{T} \hat{x}$.

Theorem Let $B = (B_1, \dots, B_n)$ be a random vector and $t = (t_1, \dots, t_n) \in \mathbb{R}^n_{>0}$.

i) If each B_i is Poisson distributed with expectation value t_i for $i = 1, \dots, n$, then it holds that

$$\mathbb{E}(\, D(B,t)\,) \;=\; \frac{1}{2}n + \sum_{i=1}^{n} O\Big(\frac{1}{t_i}\Big).$$

ii) If all V_i are iid Gamma distributed with probability density function

$$p_V(v) = \frac{K^K}{\Gamma(K)} v^{K-1} \exp(-K\,v)\, 1_{v\geq 0}(v), \qquad K \geq 1$$

and $B_i := t_i\, V_i$ for $i = 1, \dots, n$, we have

$$\mathbb{E}(\, D(B,t)\,) \;=\; \Big(\sum_{i=1}^{n} t_i\Big)\big(\psi(K{+}1){-}\log(K)\big) \;=\; \Big(\sum_{i=1}^{n} \mathbb{E}(B_i)\Big)\big(\psi(K{+}1){-}\log(K)\big),$$

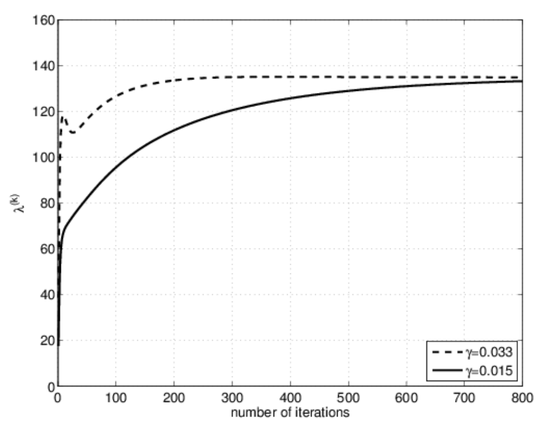
where $\psi(x) := \frac{\partial}{\partial x}\log\Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ represents the digamma function.

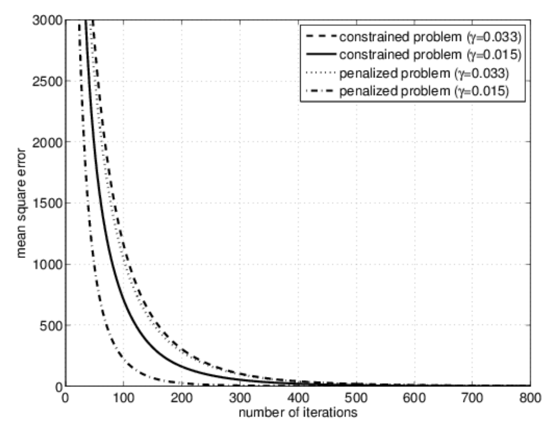
Reference for i): Bardesley et al. 2009, Bertero et al. 2010,

Poisson Noise



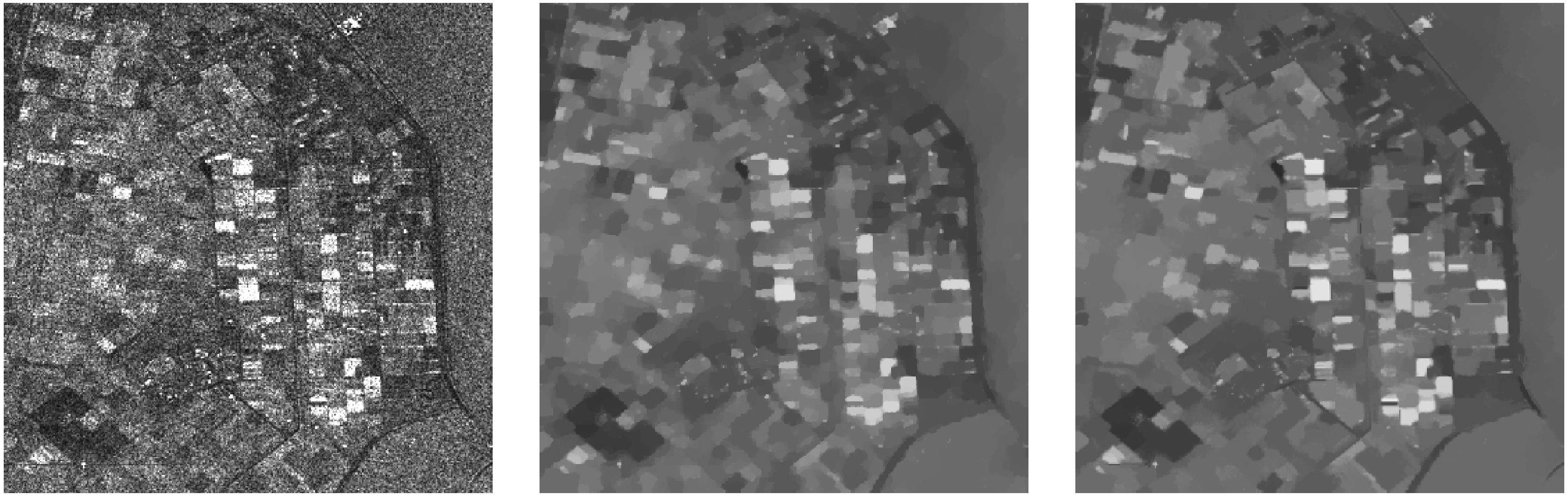
Left: Original image with values scaled to $[0, 3000]$. *Middle:* Corrupted image. *Right:* Restoration result by the I-divergence - TV model





Left: Iterates $\lambda^{(k)}$ for different parameters γ. *Right:* Mean square errors $\frac{1}{N}\|x^{(k)} - x^*\|_2^2$.

Multiplicative Gamma Noise $(H = I)$



Left: Real multi-look SAR image of size 512×512 with values in $[0, 255]$ $(K \approx 2.6)$.

Middle: Restoration result by the I-divergence -TV model.

Right: Result by the I-divergence - NL-means model (see Gilboa/Osher 2008).

5.4 Anscombe Transform Constraints

- Poisson distribution exhibits a mean/variance relationship
- Variance-stabilizing transformations can remove this relation. One of these transforms is the Anscombe transform:

$$T : u \to 2\sqrt{u + \frac{3}{8}}$$

- Transforms Poisson noise to approximately Gaussian noise with zero-mean and unit variance

Consider

$$\arg\min_{u \geq 0} \|Lu\| \quad \text{subject to} \quad \|T(Hu) - T(f)\|_2^2 \leq \tau.$$

Good choice: $\tau = n$

References: Mikkitalo, Foi 2011, Dupe, Starck, Fadili 2009 ...

Solution via epigraphical projection

- $\operatorname{epi}\varphi := \{(u,\zeta) \in \mathbb{R}^n \times \mathbb{R} : \varphi(u) \le \zeta\}$
- P_C the *orthogonal projector* onto C
- $V_\tau := \{u \in \mathbb{R}^n : \langle 1_n, u\rangle \le \tau\}$

Our constrained problem can be rewritten as

$$\operatorname{argmin}\left\{\iota_{\mathbb{R}_{\ge 0}}(u) + \|Lu\| + \sum_{i=1}^n \iota_{\operatorname{epi}\varphi_i}\left((Hu)_i + \frac{3}{8}, \zeta_i\right) + \iota_{V_\tau}(\zeta)\right\}$$

where

$$\varphi_i(s) := (2\sqrt{s + \frac{3}{8}} - (T(f))_i)^2, \quad i = 1, \ldots, n$$

Reformulation as

$$\operatorname*{argmin}_{(u,\zeta),(v_1,v_2,\eta)} \iota_C(u) + \iota_{V_\tau}(\zeta) + \|v_2\| + \sum_{i=1}^n \iota_{\operatorname{epi}\varphi_i}(v_{1,i}, \eta_i)$$

$$\text{subject to} \quad \left(\begin{array}{c|c} H & 0 \\ L & 0 \\ \hline 0 & I \end{array}\right) \begin{pmatrix} u \\ \zeta \end{pmatrix} + \begin{pmatrix} 3/8 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \eta \end{pmatrix}$$

Algorithm (PDHGMp for solving the Anscombe constrained problem)
Initialization: $u^{(0)}$, $\zeta^{(0)}$, $(p_j^{(0)})_{1\le j\le 3} = (\bar p_j^{(0)})_{1\le j\le 3}$, $\theta \in (0,1]$, $(\rho,\sigma) \in (0,+\infty)^2$ with $\rho\sigma < 1/\max\{1, \|H^*H + L^*L\|_2\}$
For $k = 0, 1, \ldots$ repeat until a stopping criterion is reached

1. $u^{(k+1)} = P_{\ge 0}\left(u^{(k)} - \sigma\rho\Big(H^*\bar p_1^{(k)} + L^*\bar p_2^{(k)}\Big)\right)$
2. $\zeta^{(k+1)} = P_{V_\tau}\left(\zeta^{(k)} - \sigma\rho\bar p_3^{(k)}\right)$
3. $(v_{1,i}^{(k+1)}, \eta_i^{(k+1)}) = P_{\mathrm{epi}\,\varphi_i}\left(p_{1,i}^{(k)} + (Hu^{(k+1)})_i + 3/8\,,\, p_{3,i}^{(k)} + \zeta_i^{(k+1)}\right), \quad i = 1,\ldots,n$
4. $v_2^{(k+1)} = \mathrm{prox}_{\sigma^{-1}\|\cdot\|}(p_2^{(k)} + Lu^{(k+1)})$
5. $p_1^{(k+1)} = p_1^{(k)} + Hu^{(k+1)} + 3/8 - v_1^{(k+1)}$
6. $p_2^{(k+1)} = p_2^{(k)} + Lu^{(k+1)} - v_2^{(k+1)}$
7. $p_3^{(k+1)} = p_3^{(k)} + \zeta^{(k+1)} - \eta^{(k+1)}$
8. $\bar p_j^{(k+1)} = p_j^{(k+1)} + \theta(p_j^{(k+1)} - p_j^{(k)}), \quad j = 1,2,3.$

- $\{p_3^{(k)}\}_k$ converges to $\lambda 1$

Proposition

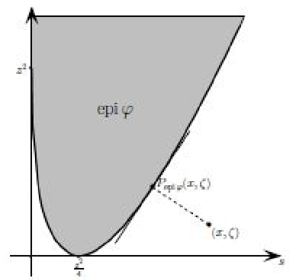
The epigraphical projection of $(x,\zeta) \in \mathbb{R}^2$ is given by

$$P_{\operatorname{epi}\varphi}(x,\zeta) = \begin{cases} (\max\{x,0\},\zeta) & \text{if } (\max\{x,0\},\zeta) \in \operatorname{epi}\varphi, \\ \left(\left(\frac{t_+ + z}{2}\right)^2, t_+^2\right) & \text{if } 4x \geq z^2, \\ \left(\left(\frac{t_- + z}{2}\right)^2, t_-^2\right) & \text{if } 4x < z^2, \end{cases}$$

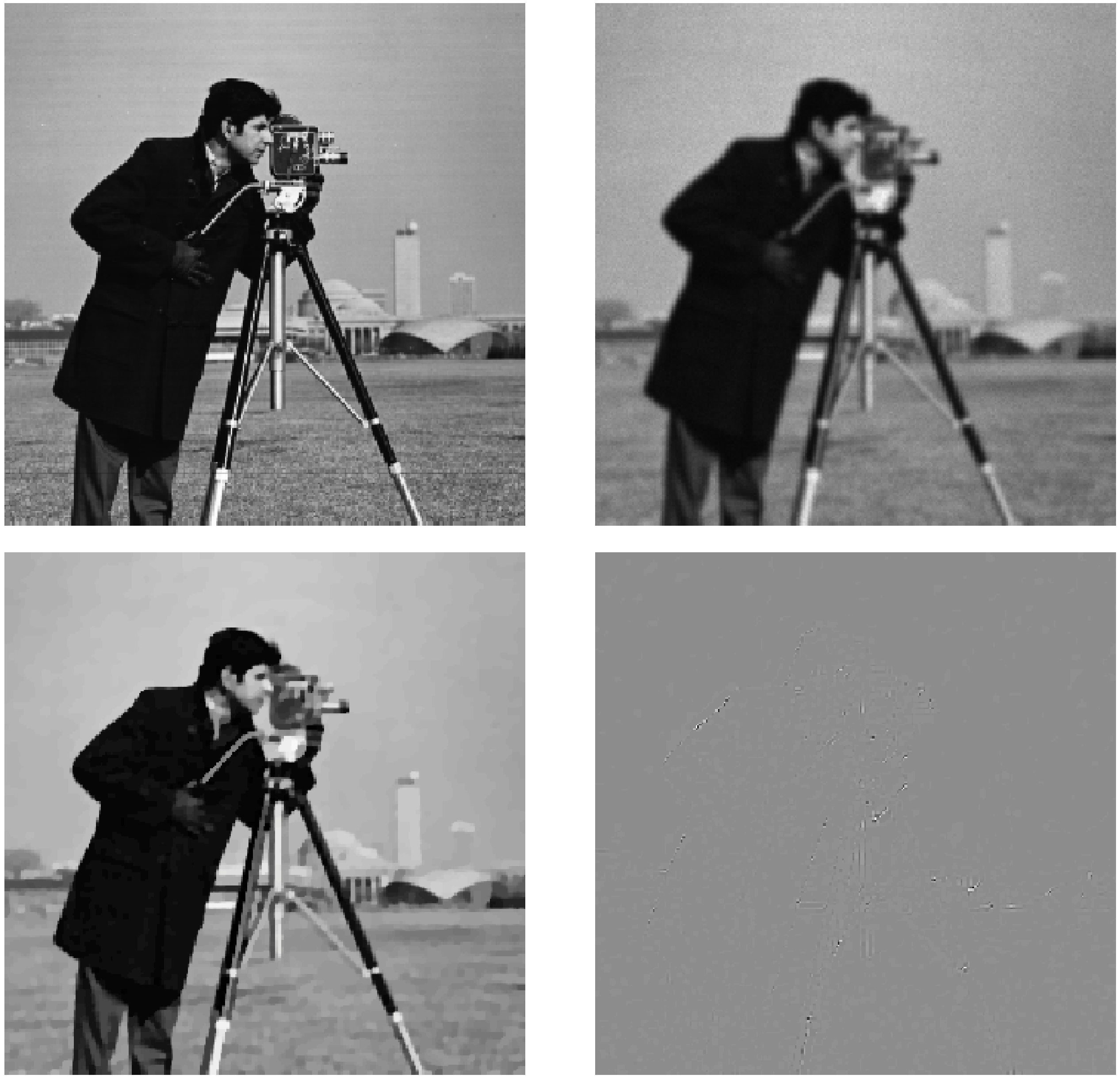
where t_+, resp. t_- is the unique root in $[0,+\infty)$, resp. in $[-z,0)$ of the cubic polynomial

$$p(t) := 17t^3 + 3zt^2 + (3z^2 - 16\zeta - 4x)t + z(z^2 - 4x).$$

- roots of p can be efficiently computed by Newton's algorithm



5.4 Anscombe Transform Constraints



Top: Original, Degraded image, Gaussian blur, $\sigma = 0.3$, noise scale $[0, 3000]$, Bottom: deblurred image: Anscombe constraint n, difference image to I-div constraint $n/2$

To take home:

- Solution of constrained problems (e.g. Morozov principle in inner iteration step, epigraphical projection)
- Relations between parameters for constrained and penalized problems
- Relations to Pareto frontiers

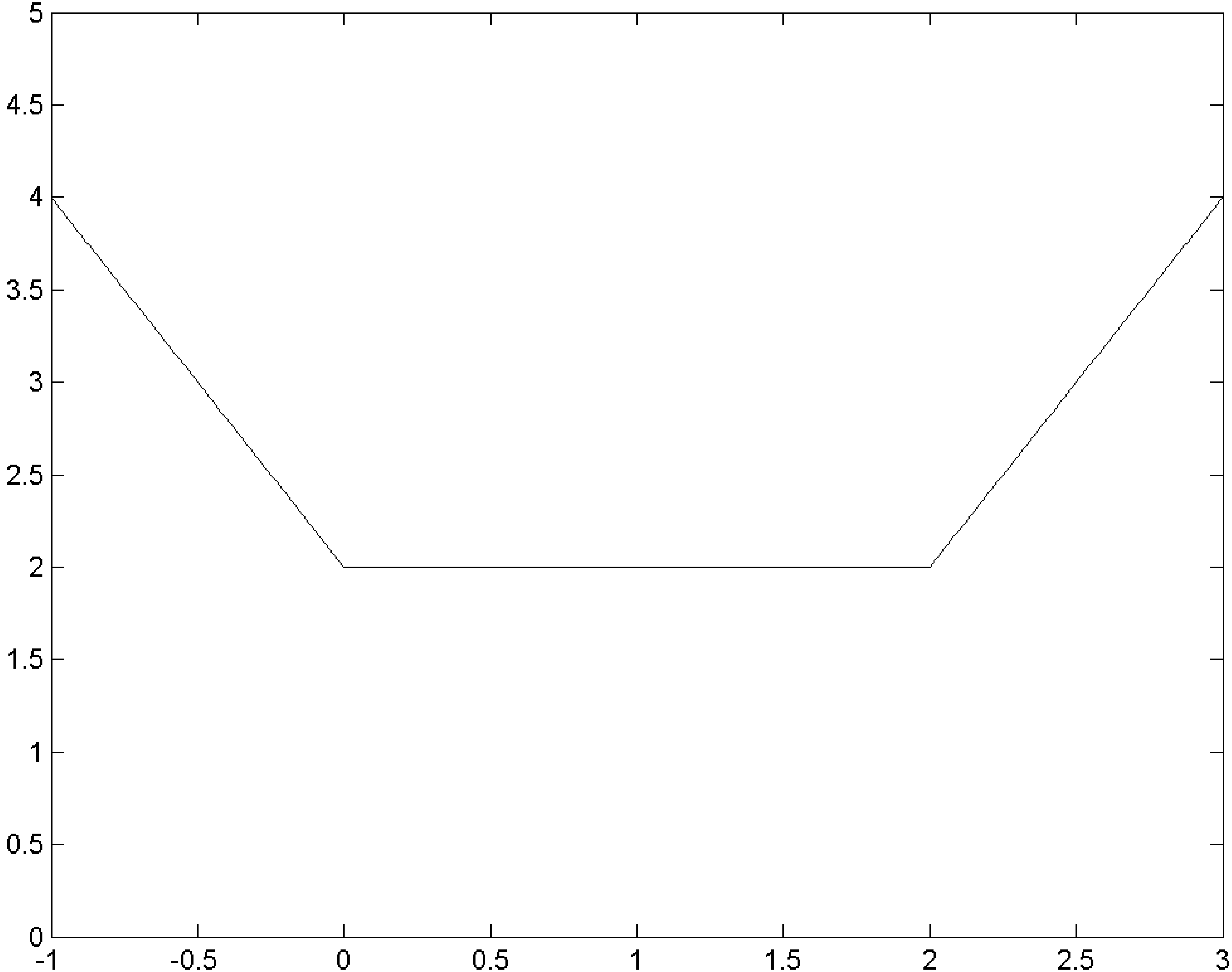
Future Tasks:

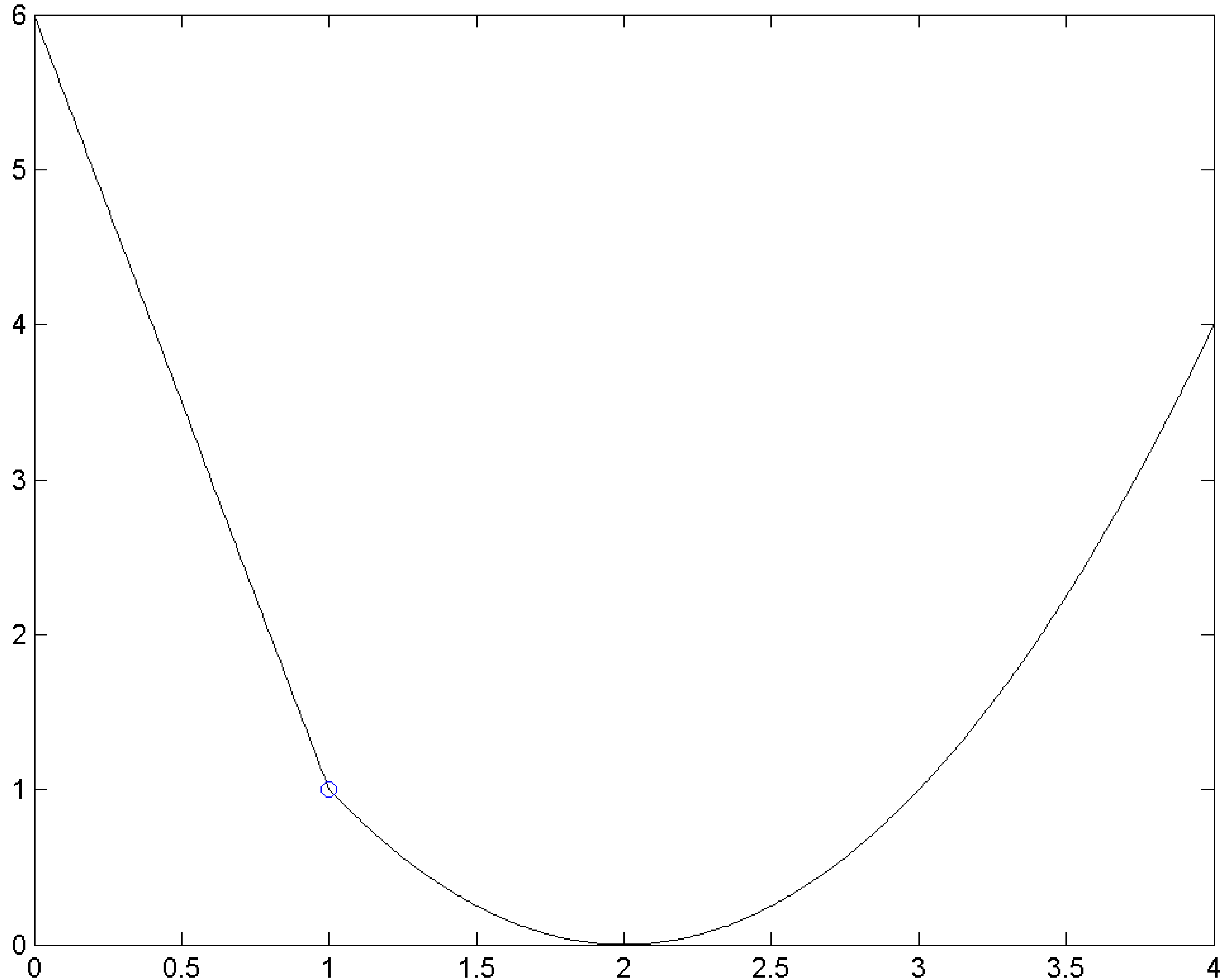
- Estimation of τ, λ
- Spatially adapted regularization parameter selection (see Dong/Hintermüller/Rincon-Camacho 2011, Chen/Cheng 2012)
- I-divergence with $b \geq 0$
- Poisson+Gaussian noise (see talk of X. Zhang)
- Nonconvex constraints ...

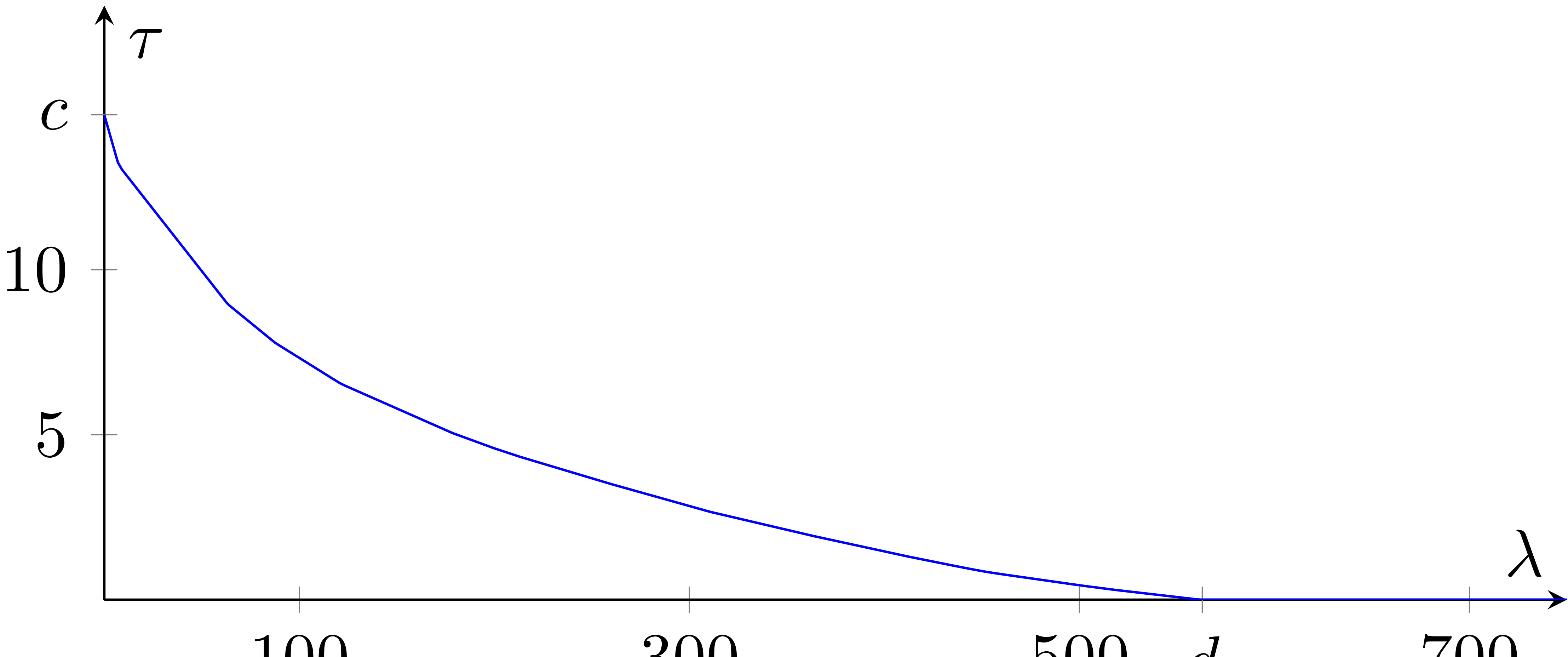
THANK YOU FOR YOUR ATTENTION!

References (steidlATmathematik.uni-kl.de):

- T. Teuber, G. Steidl and R. H. Chan,
  Minimization and parameter estimation for seminorm regularization models with I-divergence constraints,
  Inverse Problems, accepted
- R. Ciak, B. Shafei and G. Steidl,
  Homogeneous penalizers and constraints in convex image restoration,
  Journal of Mathematical Imaging and Vision, accepted
- T. Teuber and G. Steidl,
  Removing multiplicative noise by Douglas-Rachford splitting methods,
  Journal of Mathematical Imaging and Vision, 36(3), 168 - 184, 2010
- S. Setzer, G. Steidl and T. Teuber,
  Deblurring Poissonian images by split Bregman techniques,
  J. Visual Communication and Image Representation, 21, 193 - 199, 2010

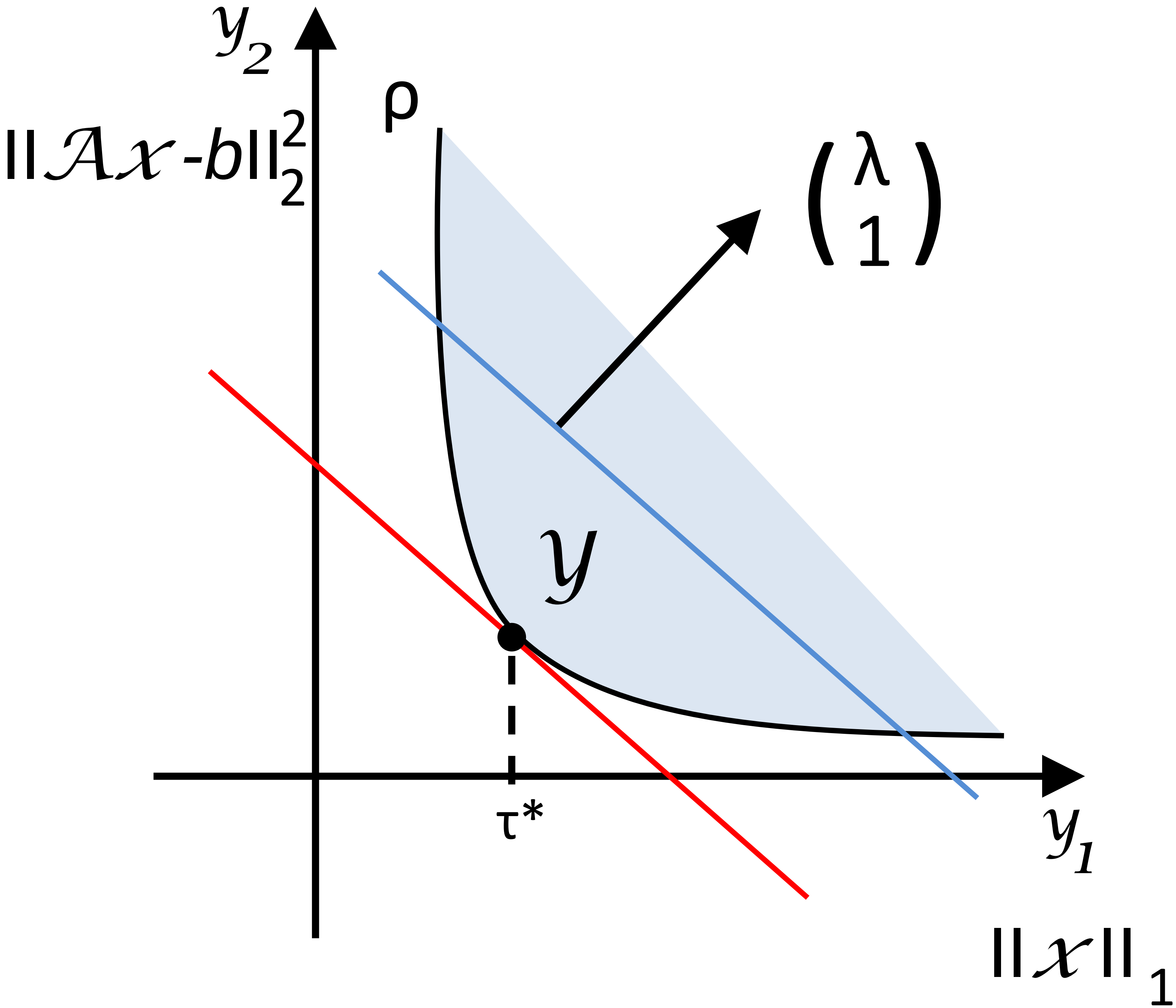

5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
-1
-0.5
0
0.5
1
1.5
2
2.5
3


6
5
4
3
2
1
0
0
0.5
1
1.5
2
2.5
3
3.5
4

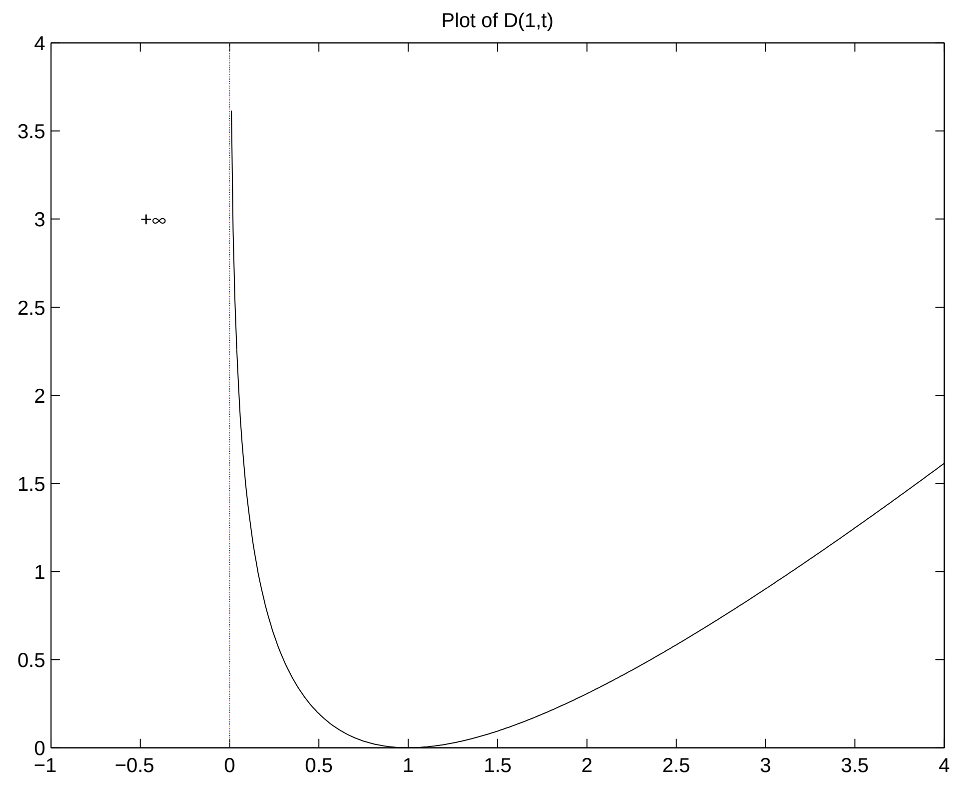


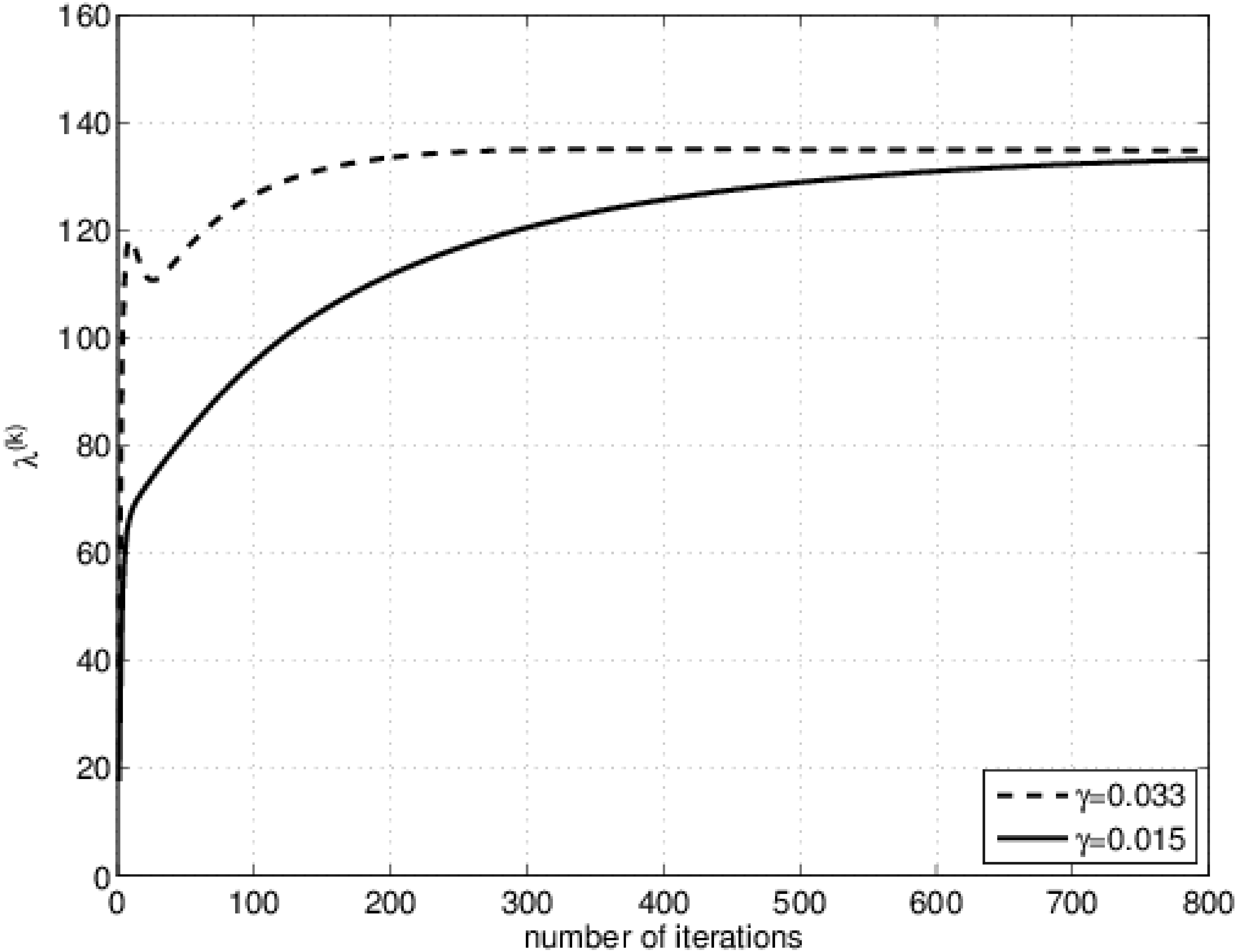
τ
c
10
5
λ
100
300
500
d
700


Take five times which plus half
of what and make the square of
what you've got. Divide by one
and thirty square to get just
four – that's right, it's there.
Now two more points I must
impress: Both of which and
what are fractionless. And what
less which is not a lot, just two
or three. So now, what's what?

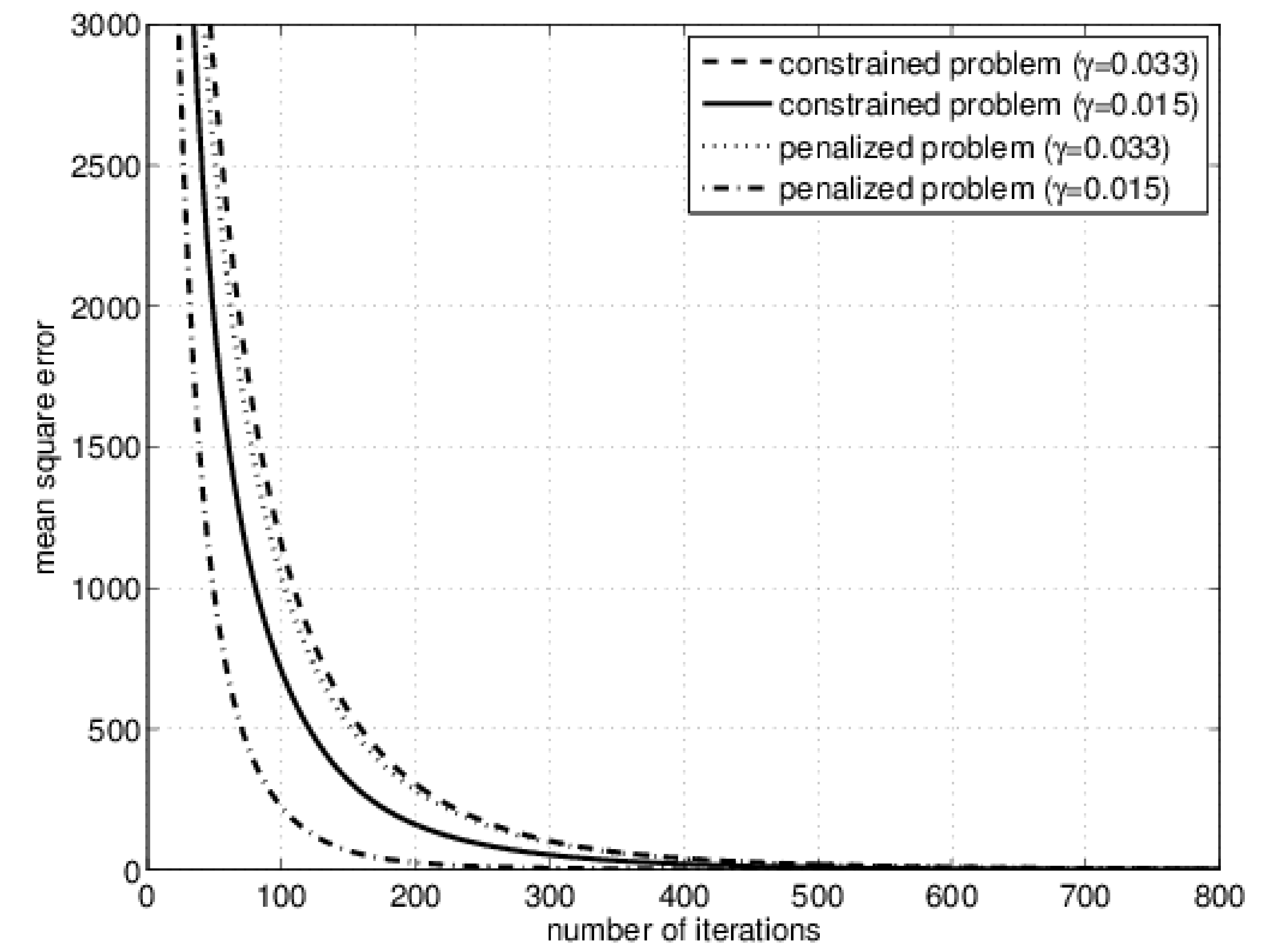


y_2
$\|\mathcal{A}x - b\|_2^2$
ρ
$\begin{pmatrix}\lambda \\ 1\end{pmatrix}$
y
τ*
y_1
$\|x\|_1$

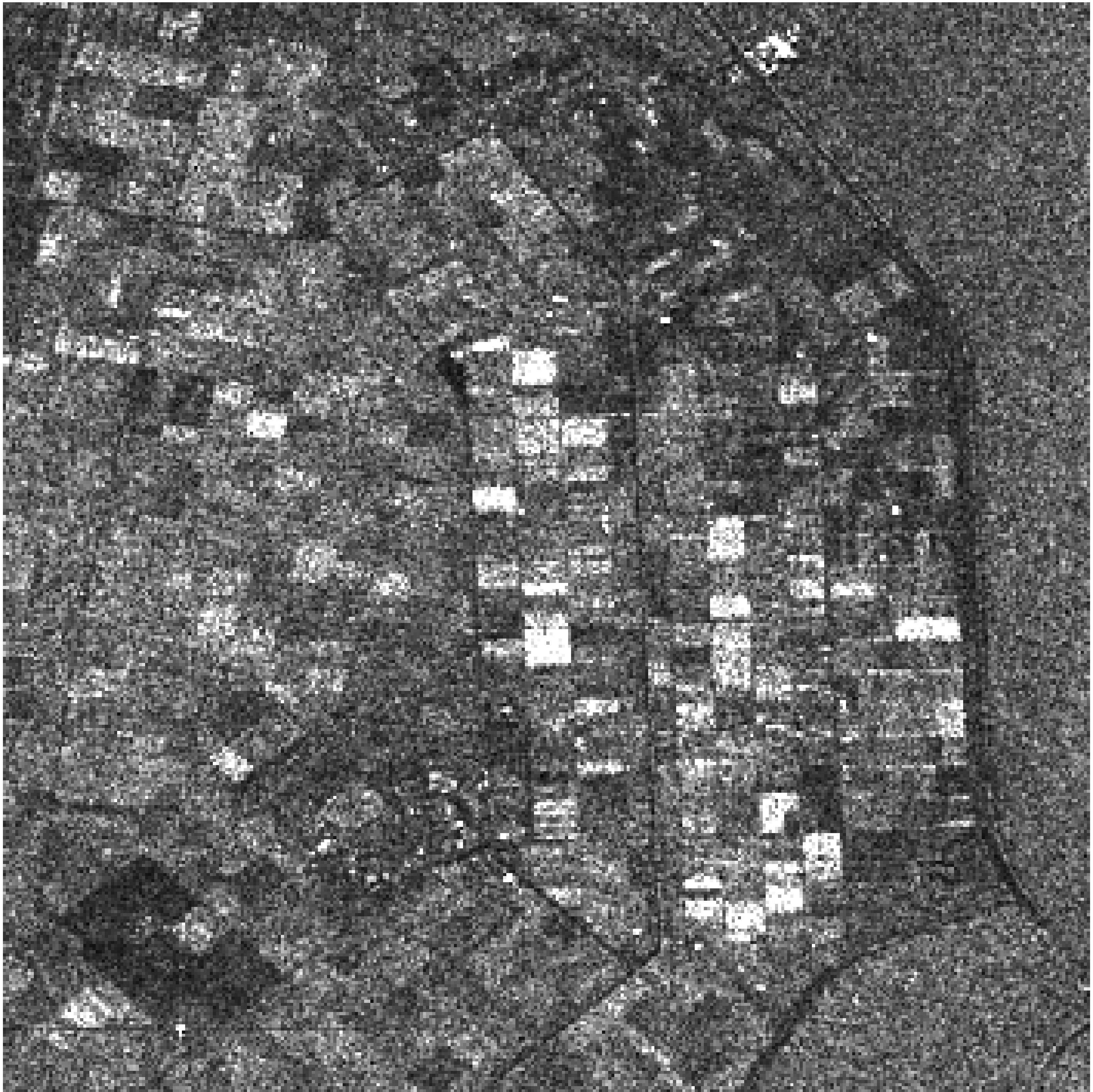

Plot of D(1,t)
+∞
4
3.5
3
2.5
2
1.5
1
0.5
0
−1
−0.5
0
0.5
1
1.5
2
2.5
3
3.5
4

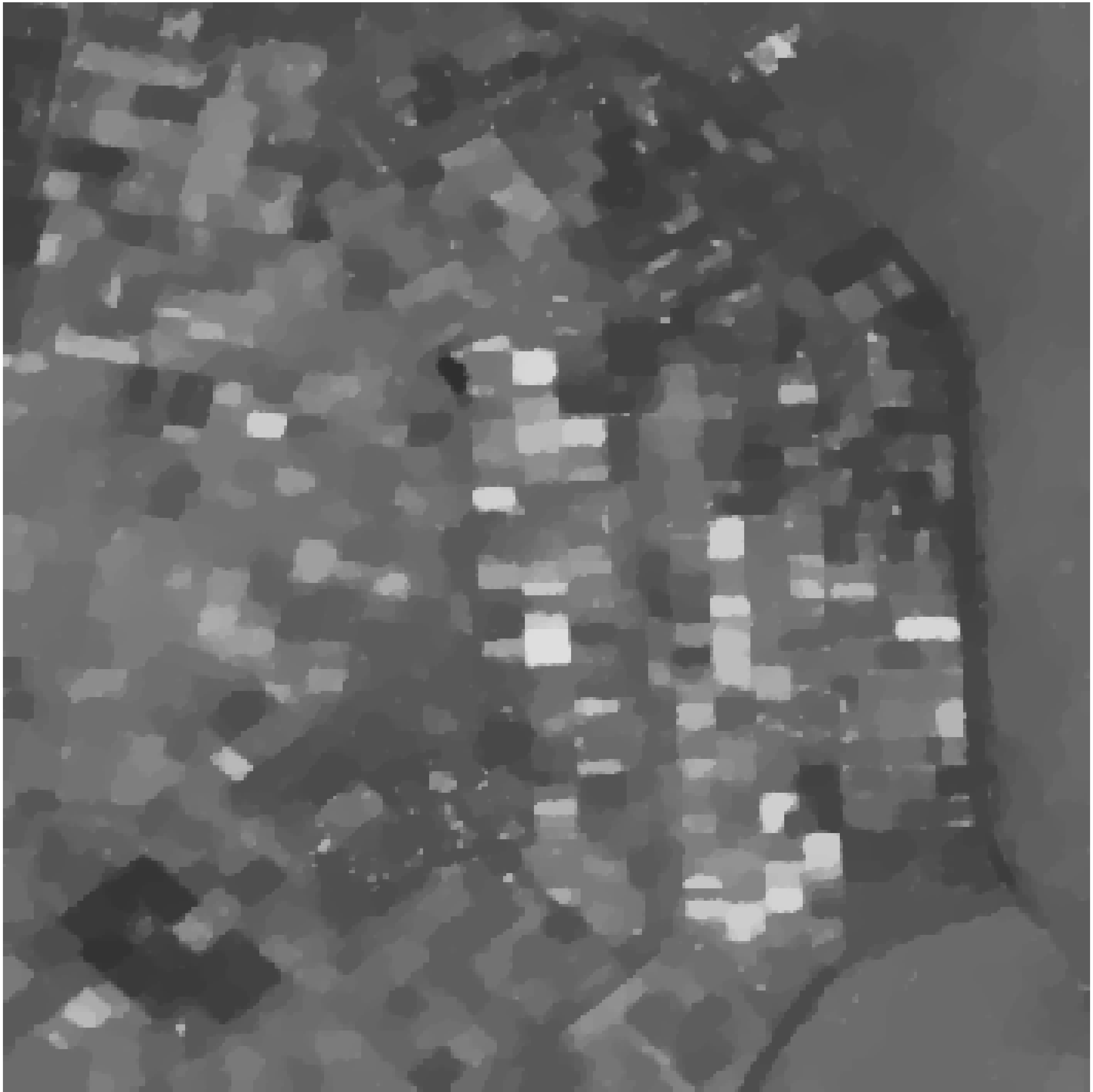


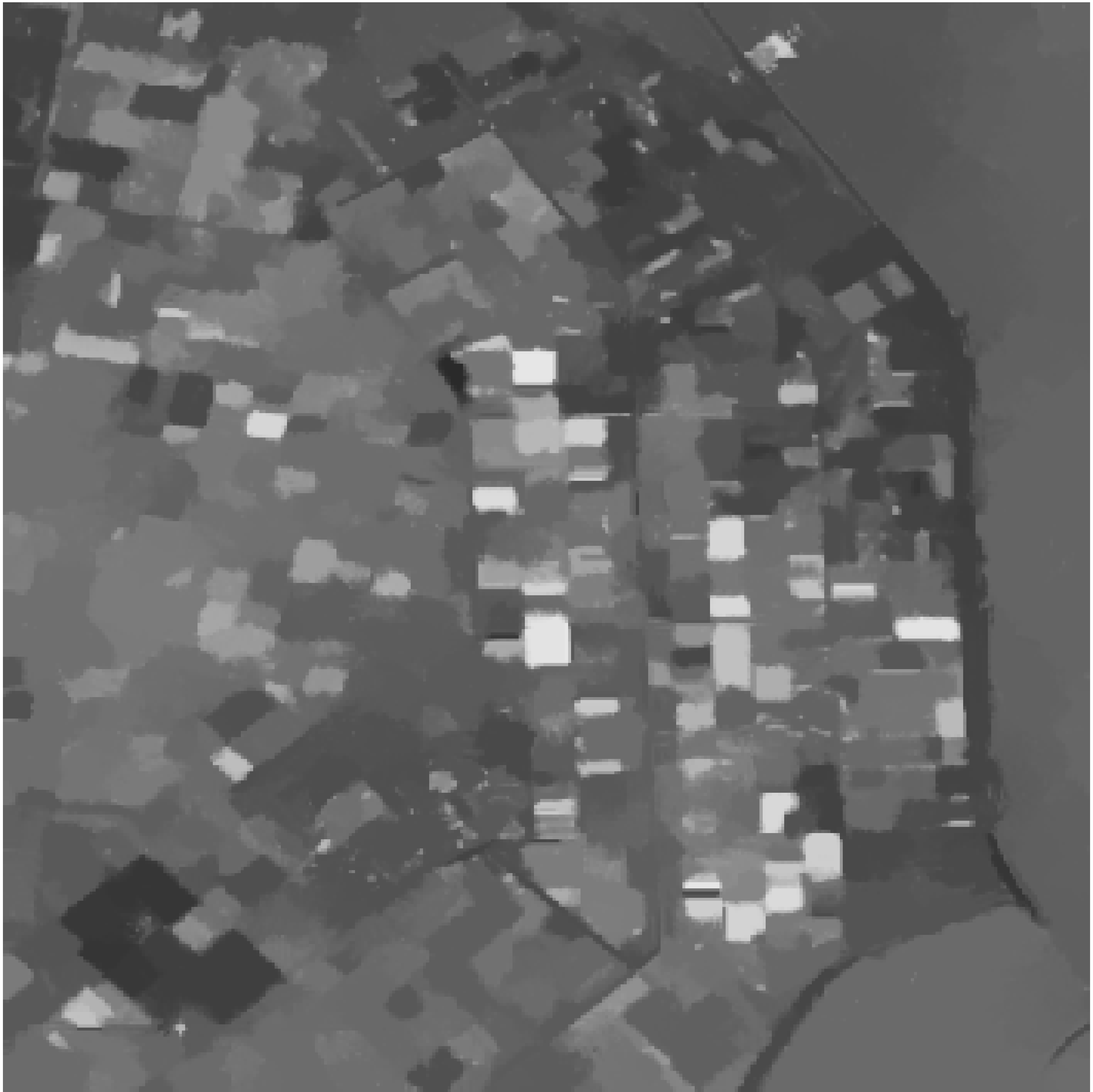
160
140
120
100
80
60
40
20
0
0
100
200
300
400
500
600
700
800
number of iterations
$\lambda^{(k)}$
γ=0.033
γ=0.015

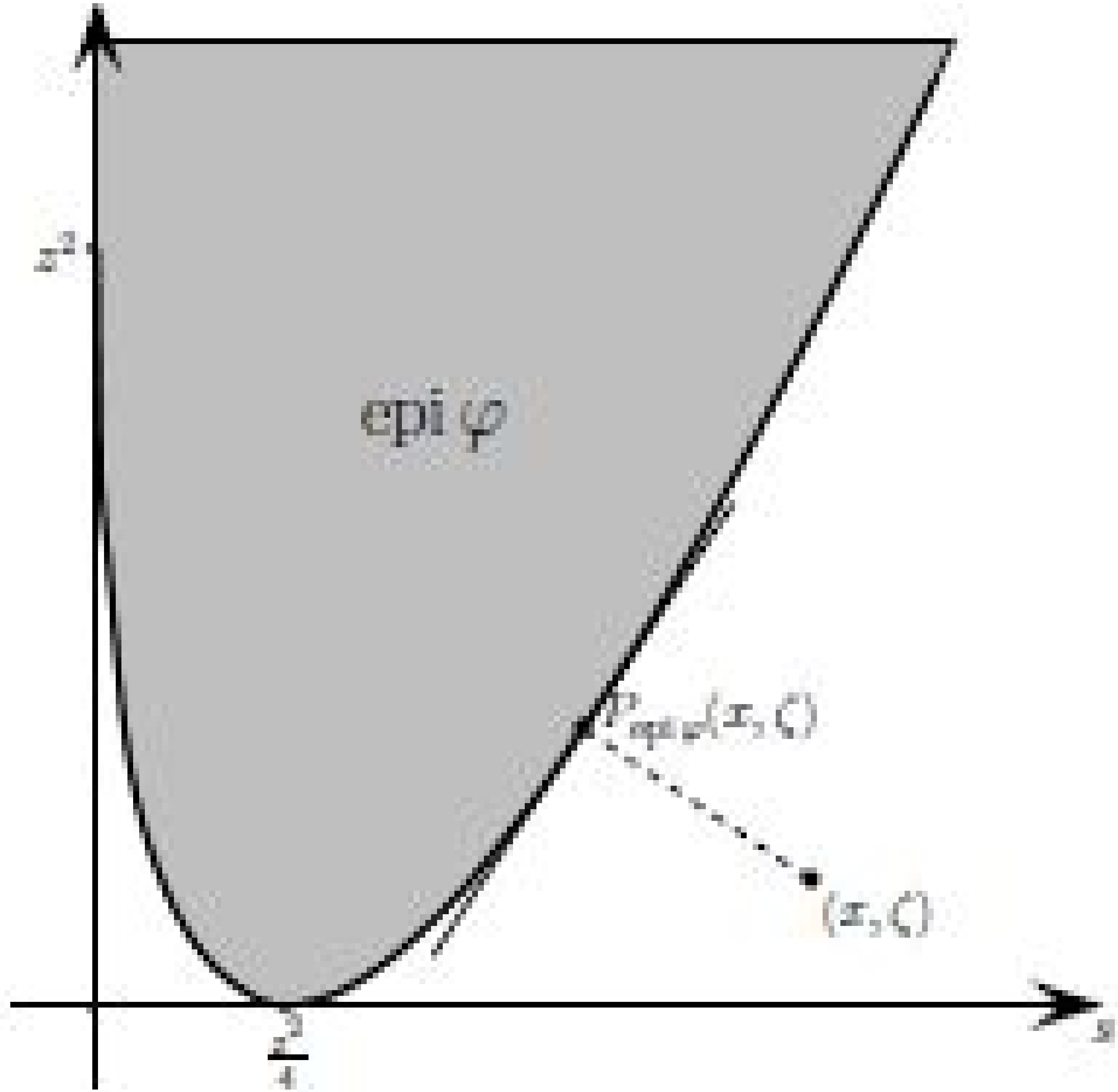


constrained problem (γ=0.033)
constrained problem (γ=0.015)
penalized problem (γ=0.033)
penalized problem (γ=0.015)
mean square error
number of iterations
0
500
1000
1500
2000
2500
3000
0
100
200
300
400
500
600
700
800








epi φ
$P_{\text{epi}\,\varphi}(x,\zeta)$
(x,ζ)