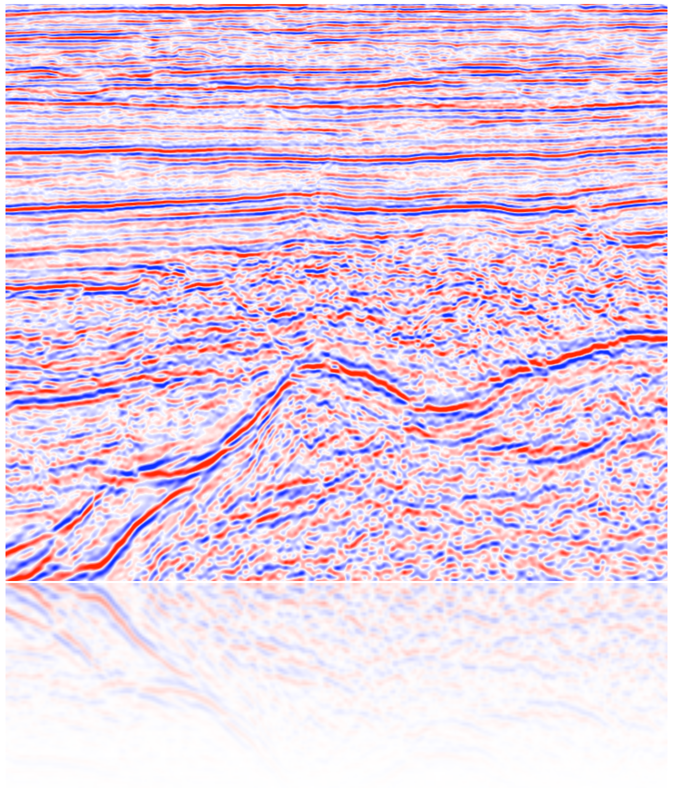
New insights into one-norm solvers from the Pareto curve

Gilles Hennenfent
ghennenfent@eos.ubc.ca
http://wigner.eos.ubc.ca/~hegilles

Ewout van den Berg
Michael P. Friedlander
Felix J. Herrmann

Seismic Laboratory for Imaging & Modeling
Department of Earth & Ocean Sciences
The University of British Columbia

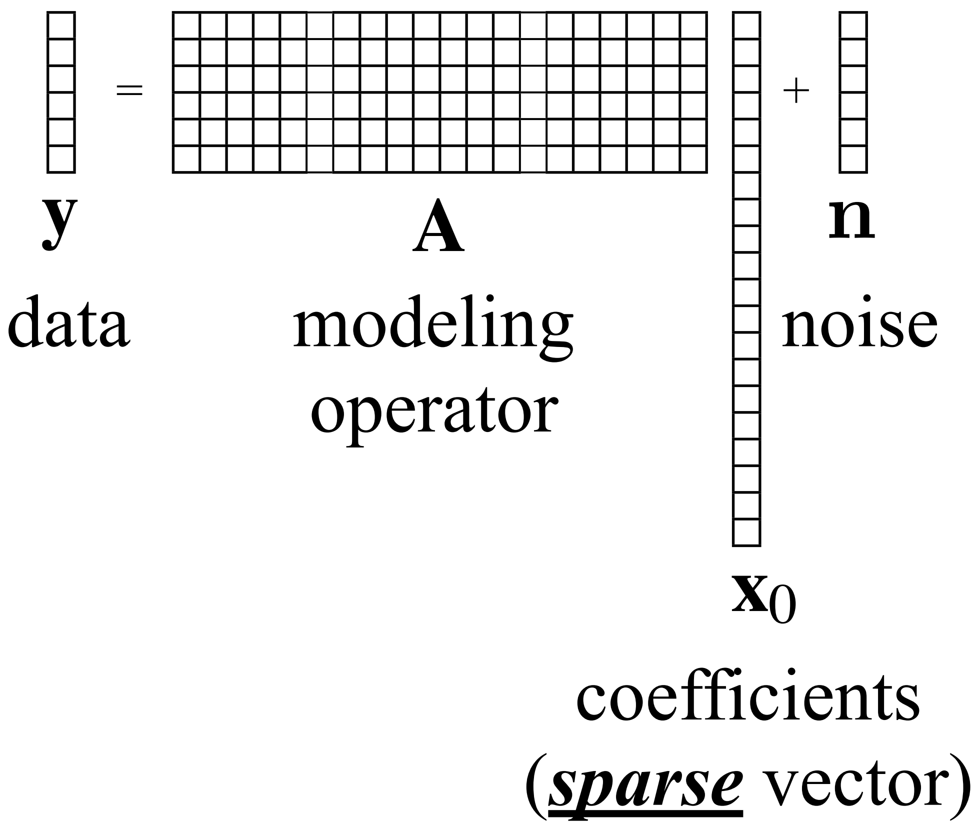


SLIM consortium meeting
Background on sparsity-promoting solvers
Wednesday, February 20th, 2008 - 10:00 AM

[Hennenfent et al., 2007]

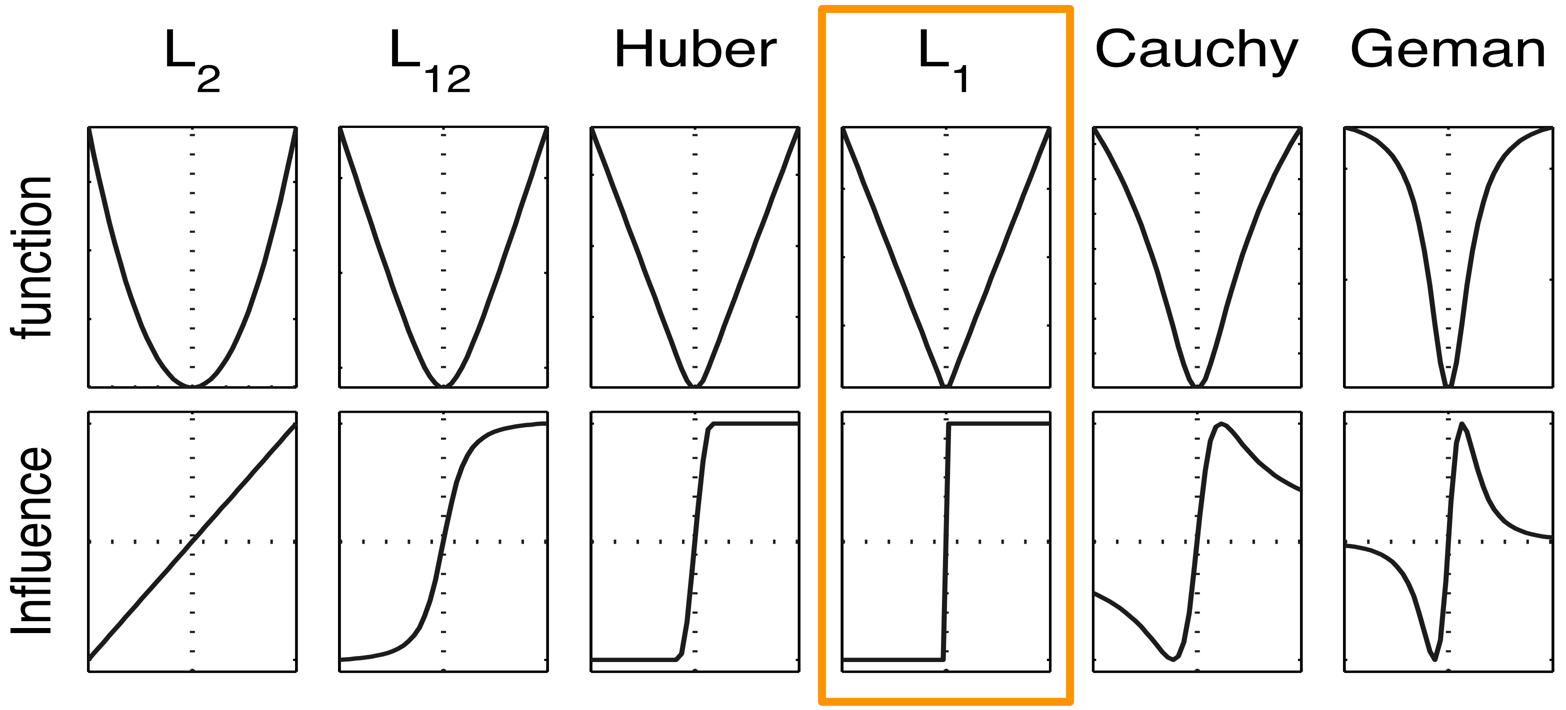
Problem statement

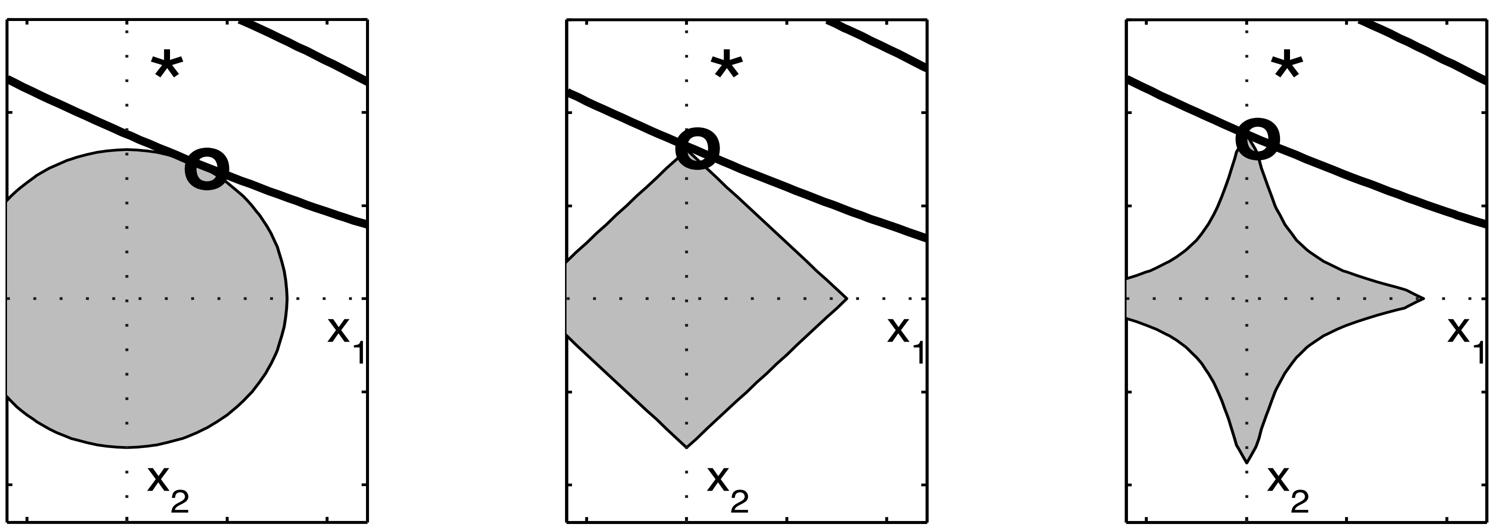
- underdetermined system of linear equations
  - setup



  - examples
    - wavefield reconstruction with **A=RS**^{H} [Sacchi et al ‘98], [Xu et al ‘05], [Zwartjes and Sacchi, 2007], [Herrmann and Hennenfent, 2007] & talks by J. Johnson and J. Yan
    - denoising with **A=S**^{H} [many references!] & talks by R. Neelamani and V. Kumar
    - (deconvolution with **A=KS**^{H} [Hennenfent et al., 2005] & talk by V. Kumar)

Sparse solution via one-norm





Approaches

- quadratic programming [many references!]

$$\mathrm{QP}_\lambda : \quad \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1$$

- basis pursuit denoise [Chen et al., 1995]

$$\mathrm{BP}_\sigma : \quad \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \sigma$$

- LASSO [Tibshirani, 1996]

$$\mathrm{LS}_\tau : \quad \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq \tau$$

Approaches

- quadratic programming [many references!]

$$\mathrm{QP}_\lambda : \quad \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1$$

- basis pursuit denoise [Chen et al., 1995]

$$\mathrm{BP}_\sigma : \quad \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \sigma$$

- LASSO [Tibshirani, 1996]

$$\mathrm{LS}_\tau : \quad \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq \tau$$

Approaches

- quadratic programming [many references!]

$$\mathrm{QP}_\lambda : \quad \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2 + \boxed{\lambda}\|\mathbf{x}\|_1$$

- basis pursuit denoise [Chen et al., 1995]

$$\mathrm{BP}_\sigma : \quad \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2 \le \boxed{\sigma}$$

- LASSO [Tibshirani, 1996]

$$\mathrm{LS}_\tau : \quad \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2 \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \le \tau$$

Approaches

- quadratic programming [many references!]

$$\mathrm{QP}_\lambda : \quad \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \boxed{\lambda}\|\mathbf{x}\|_1$$

- basis pursuit denoise [Chen et al., 1995]

$$\mathrm{BP}_\sigma : \quad \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \boxed{\sigma}$$

- LASSO [Tibshirani, 1996]

$$\mathrm{LS}_\tau : \quad \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq \boxed{\tau}$$

Solvers

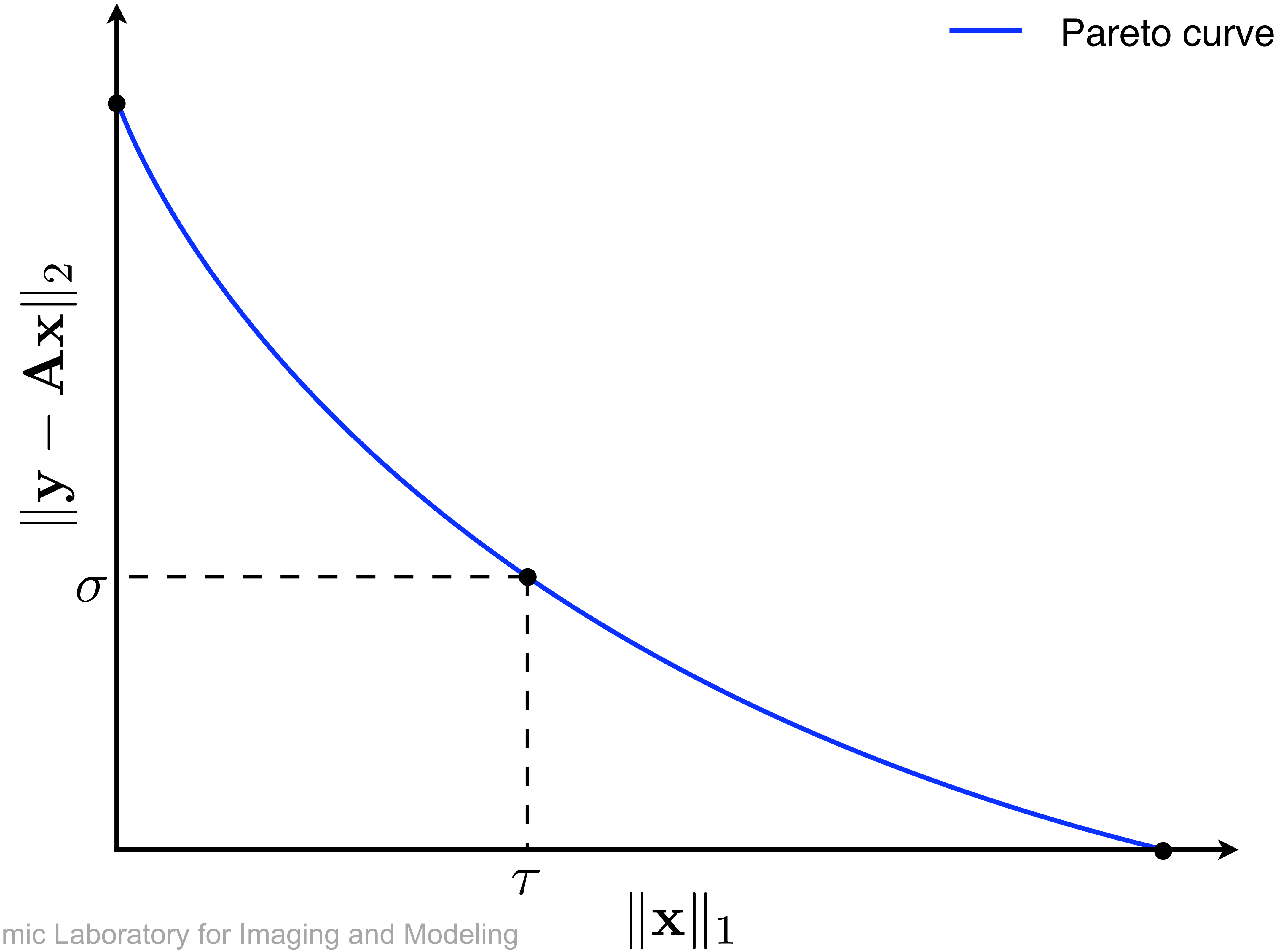
- quadratic programming
  - iteratively re-weighted least-squares (IRLS) [Gersztenkorn et al., 1986]
  - iterative soft thresholding (IST) [Daubechies et al., 2004]
  - extensions of IST [Figueiredo and Nowak, 2003], [Herrmann and Hennenfent, 2008]
  - primal-dual interior method for convex objectives (PDCO)
- basis pursuit denoise
  - iterative soft thresholding with cooling (ISTc) [Herrmann and Hennenfent, 2008]
  - spectral projected-gradient for ℓ_1-norm (SPGℓ_1) [van den Berg and Friedlander, 2007]
  - log-barrier methods for second-order cone programming
  - homotopy methods [Osborne et al., 2000], [Donoho and Tsaig, 2006]
- LASSO
  - projected gradient [Daubechies et al., 2007], [van den Berg and Friedlander, 2007]

Connection & comparison

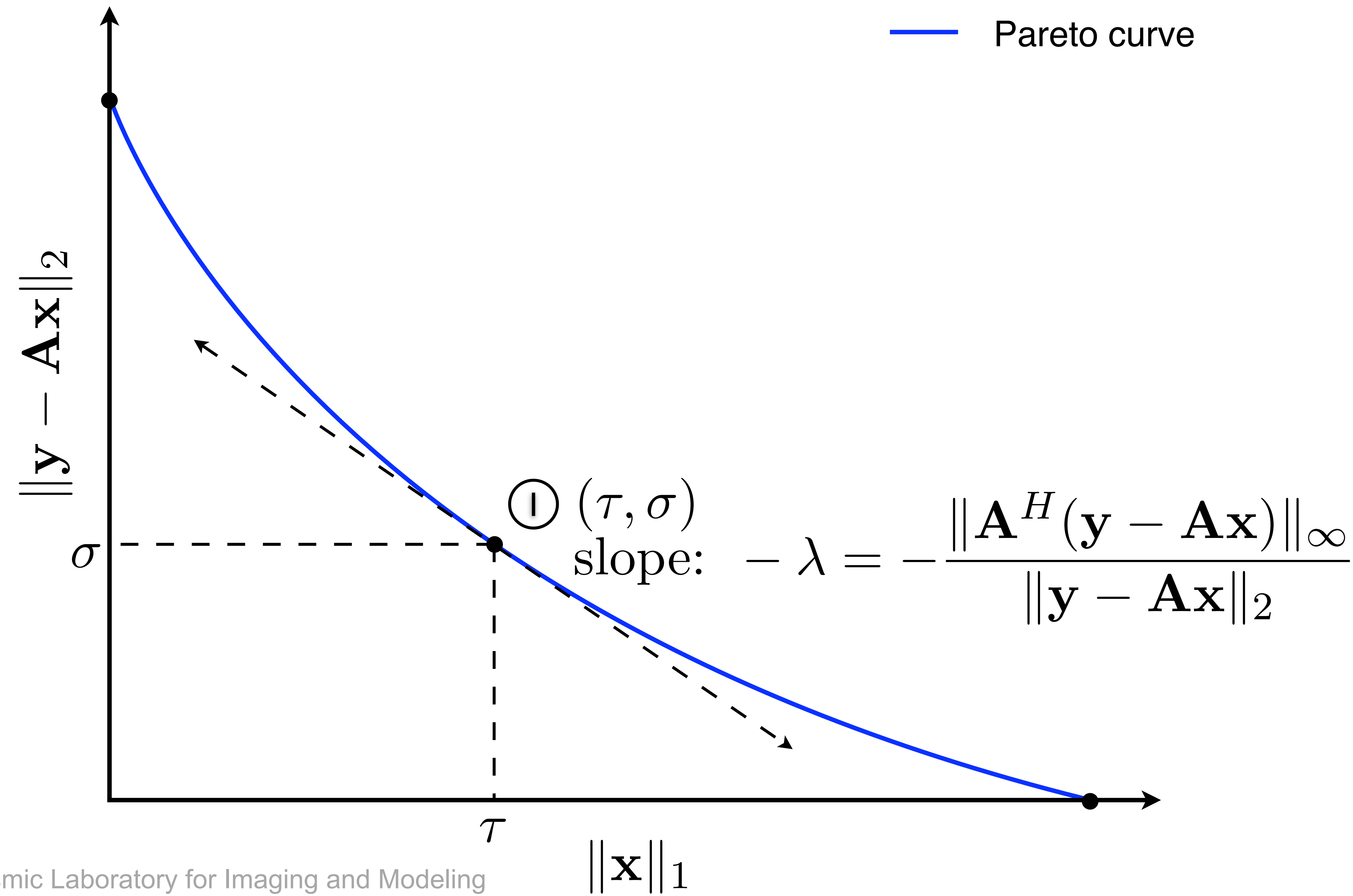
- context
  - ONE problem
  - THREE approaches
  - MANY solvers…

- Pareto curve
  - defined as the **optimal tradeoff** between $||\mathbf{y}-\mathbf{Ax}||_2$ and $||\mathbf{x}||_1$
  - **establishes the connection** between the three approaches
  - **exposes the behavior** of one-norm solvers
  - used to **evaluate the performance** of one-norm solvers

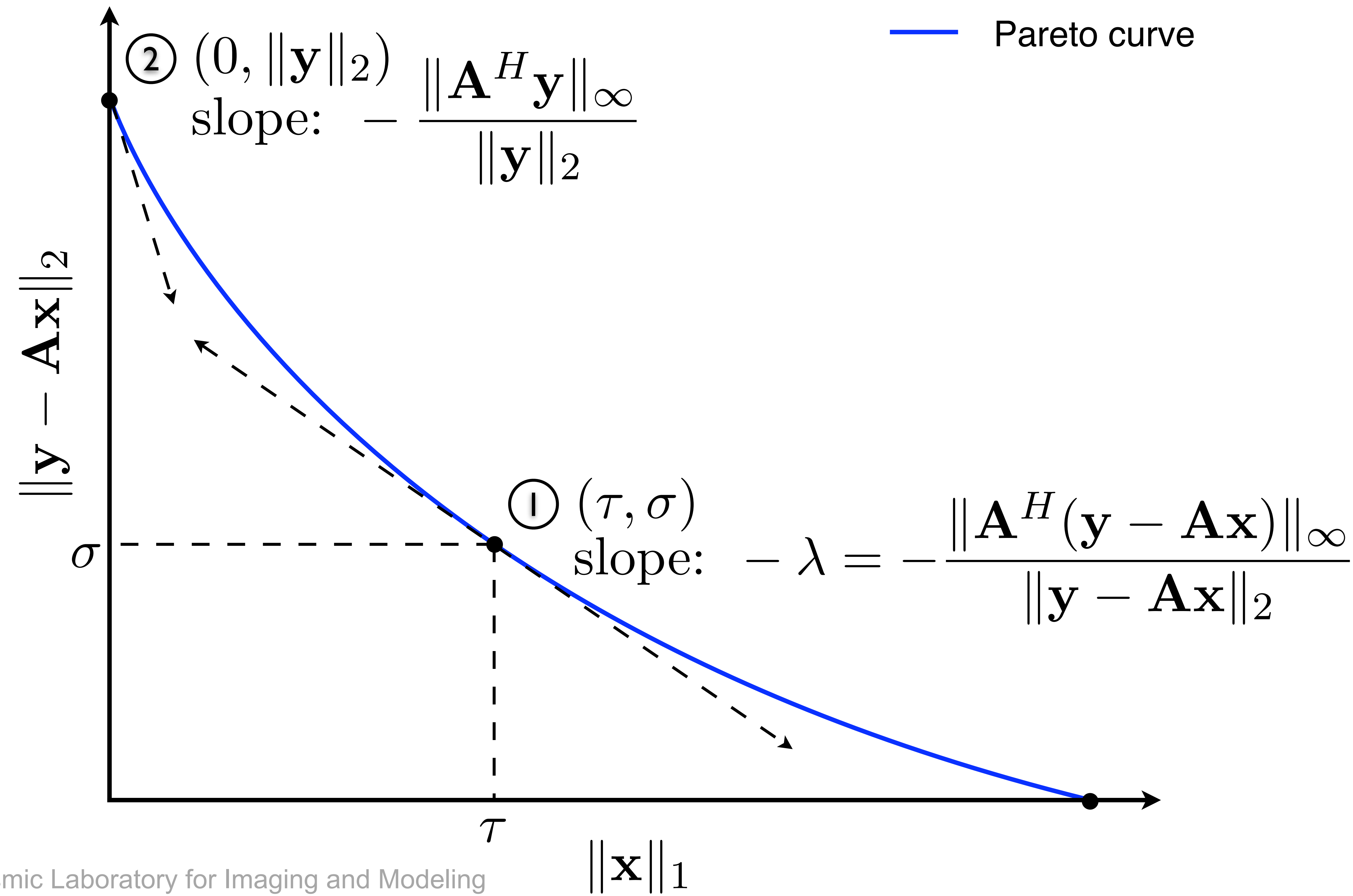
Pareto curve



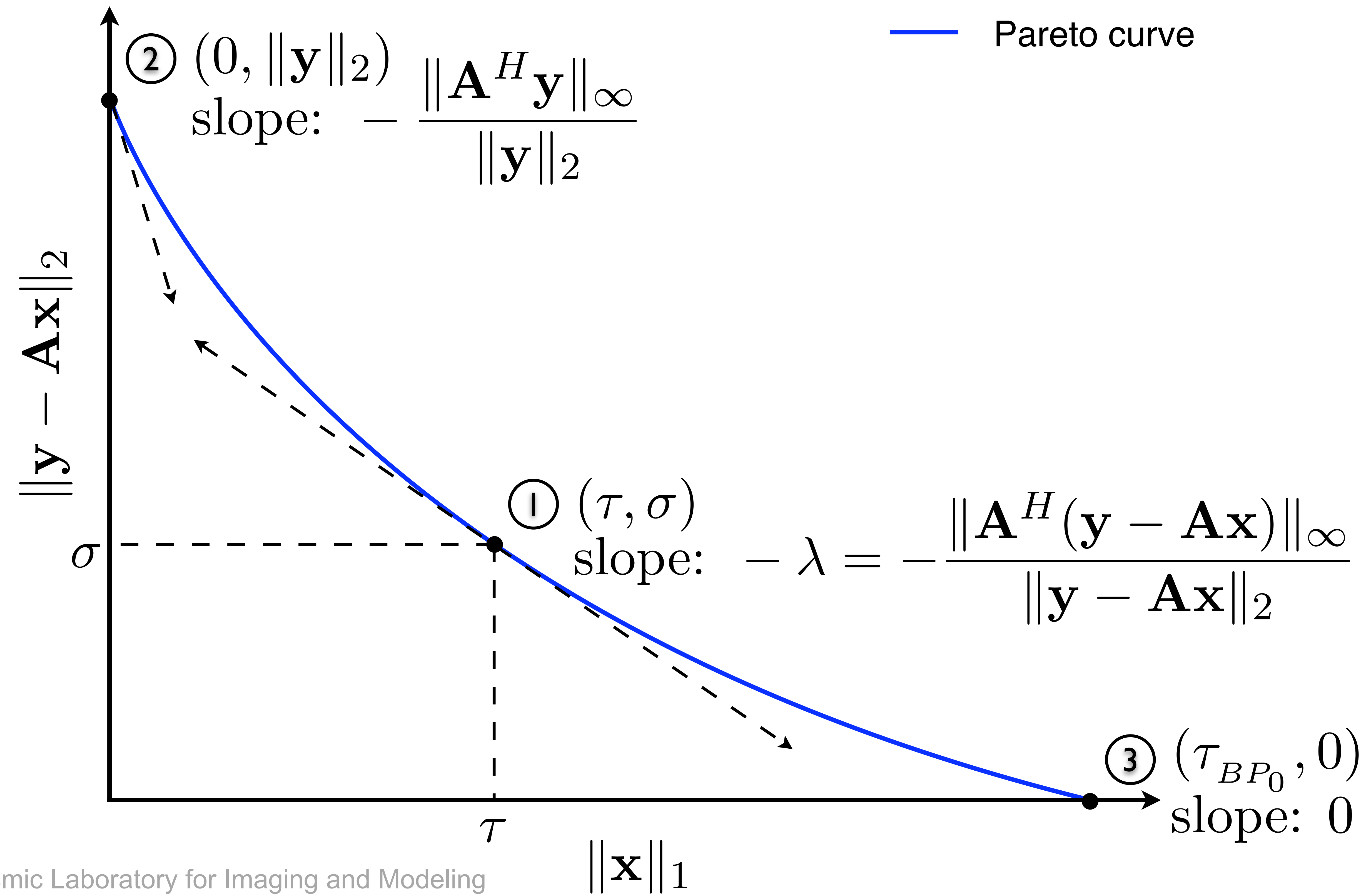
Pareto curve



Pareto curve



Pareto curve



Pareto curve

- useful properties [van den Berg and Friedlander, 2007]
  - convex & decreasing
  - continuously differentiable
  - negative slope given by

$$\lambda = \frac{\|\mathbf{A}^H(\mathbf{y} - \mathbf{A}\mathbf{x})\|_\infty}{\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2}$$

- consequence
  - good approximation to the Pareto curve obtained with VERY few interpolating points

Pareto curve

- useful properties [van den Berg and Friedlander, 2007]
  - convex & decreasing
  - continuously differentiable
  - negative slope given by

$$\lambda = \frac{\|\mathbf{A}^H(\mathbf{y} - \mathbf{A}\mathbf{x})\|_\infty}{\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2}$$

- consequence
  - good approximation to the Pareto curve obtained with VERY few interpolating points

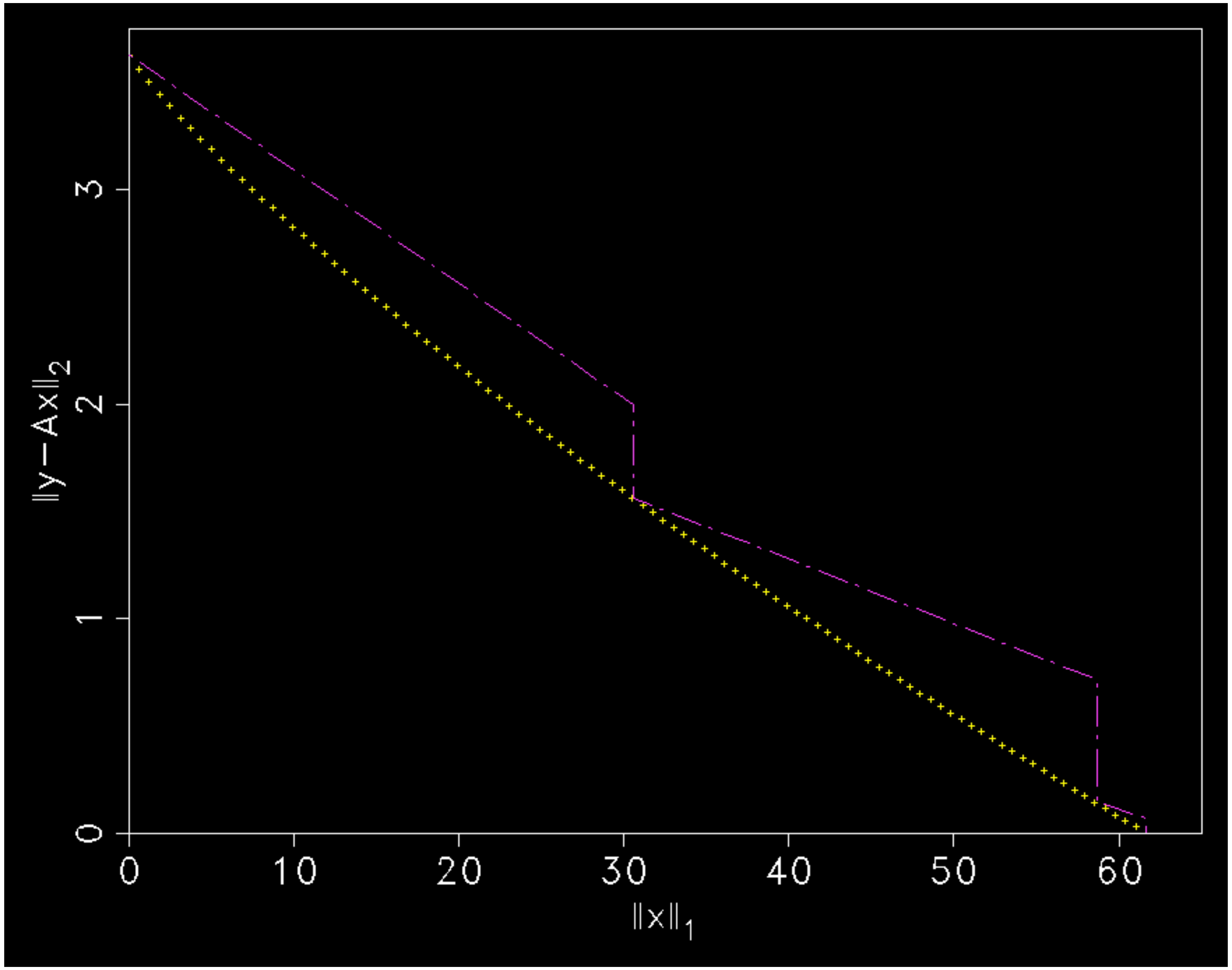
key point for large-scale geophysical applications

Comparison of one-norm solvers

- solution path
  - track the evolution of the data misfit versus the one norm of successive solver iterates

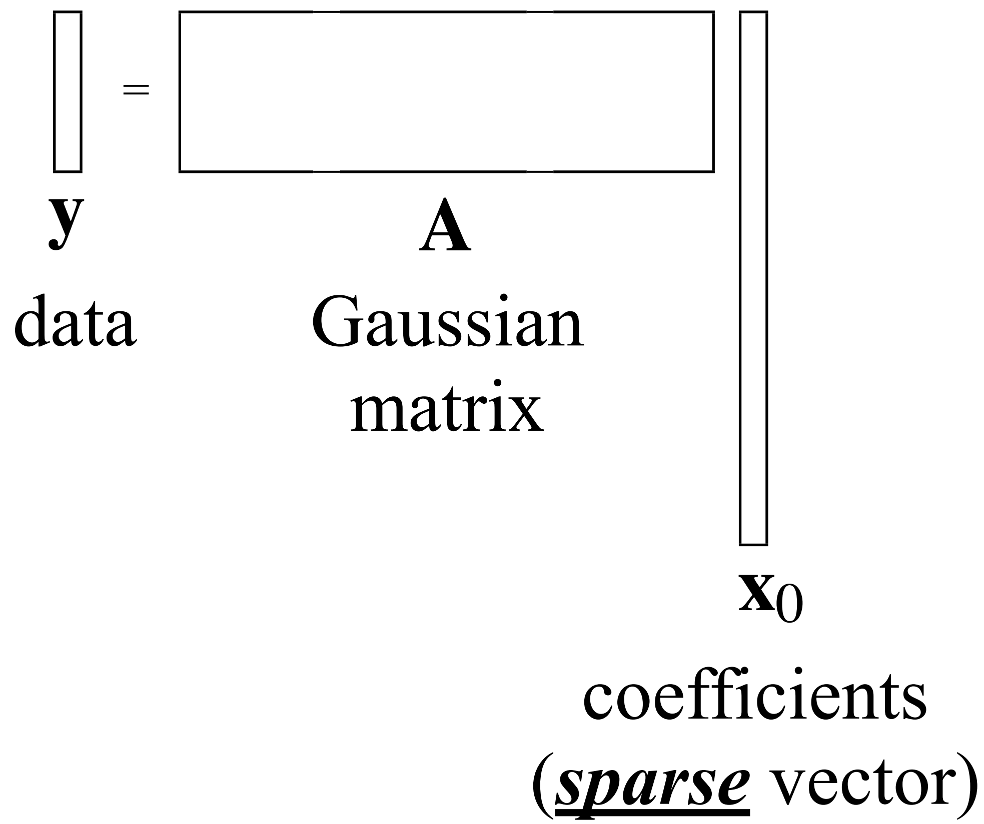
Comparison of one-norm solvers

- solution path
  - track the evolution of the data misfit versus the one norm of successive solver iterates



Benchmark problem

- setup



- objective
  - basis pursuit (BP) solution

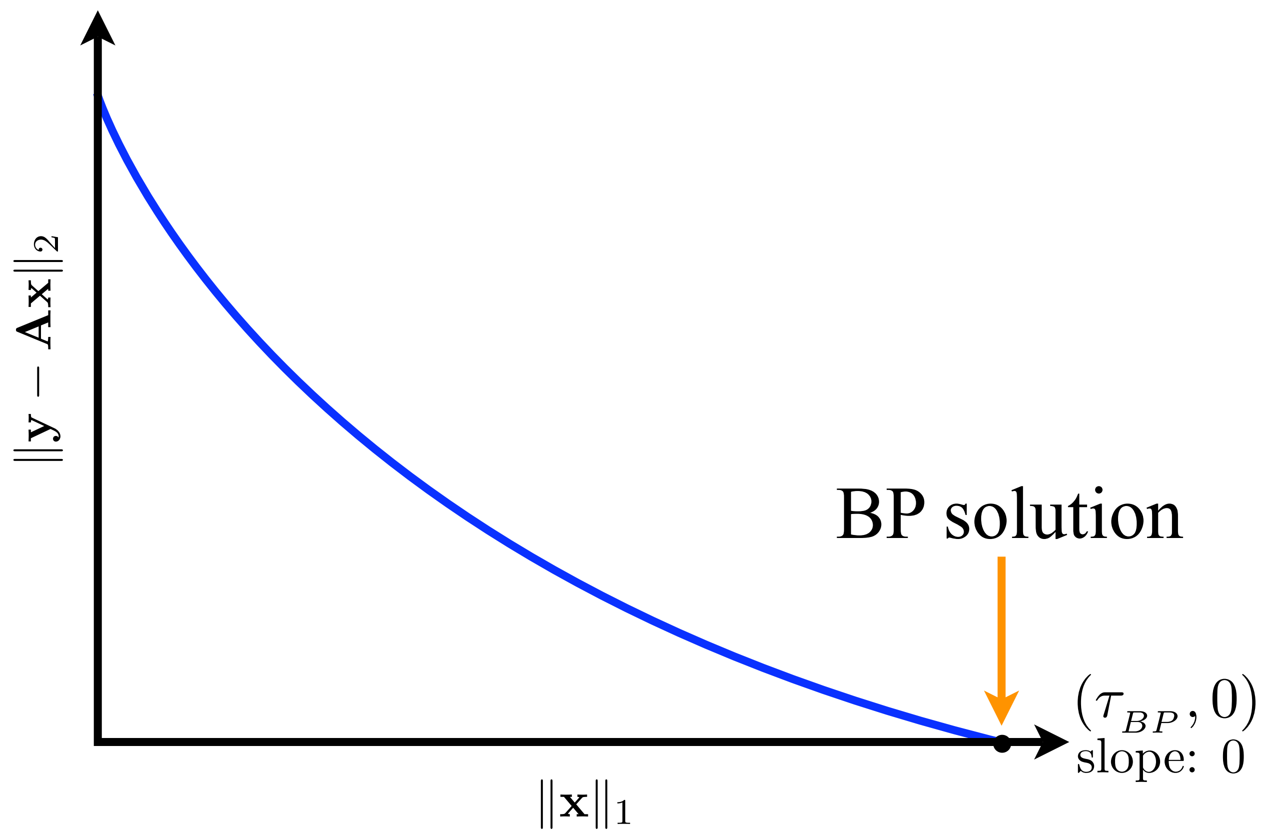
$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\mathbf{x}$$

[Candès et al., 2006]
[Donoho, 2006]

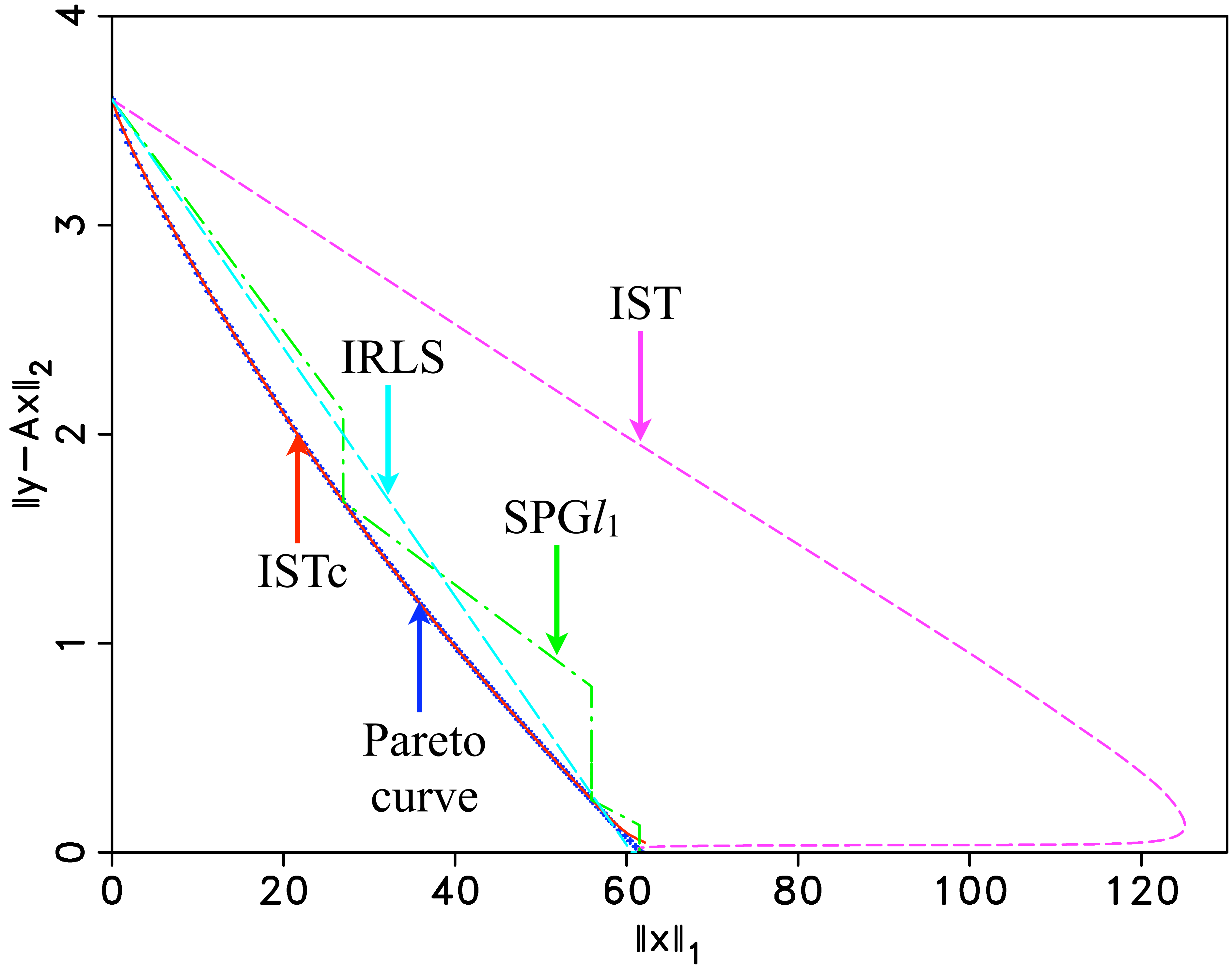
Basis pursuit

$$\text{BP}: \quad \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\mathbf{x}$$

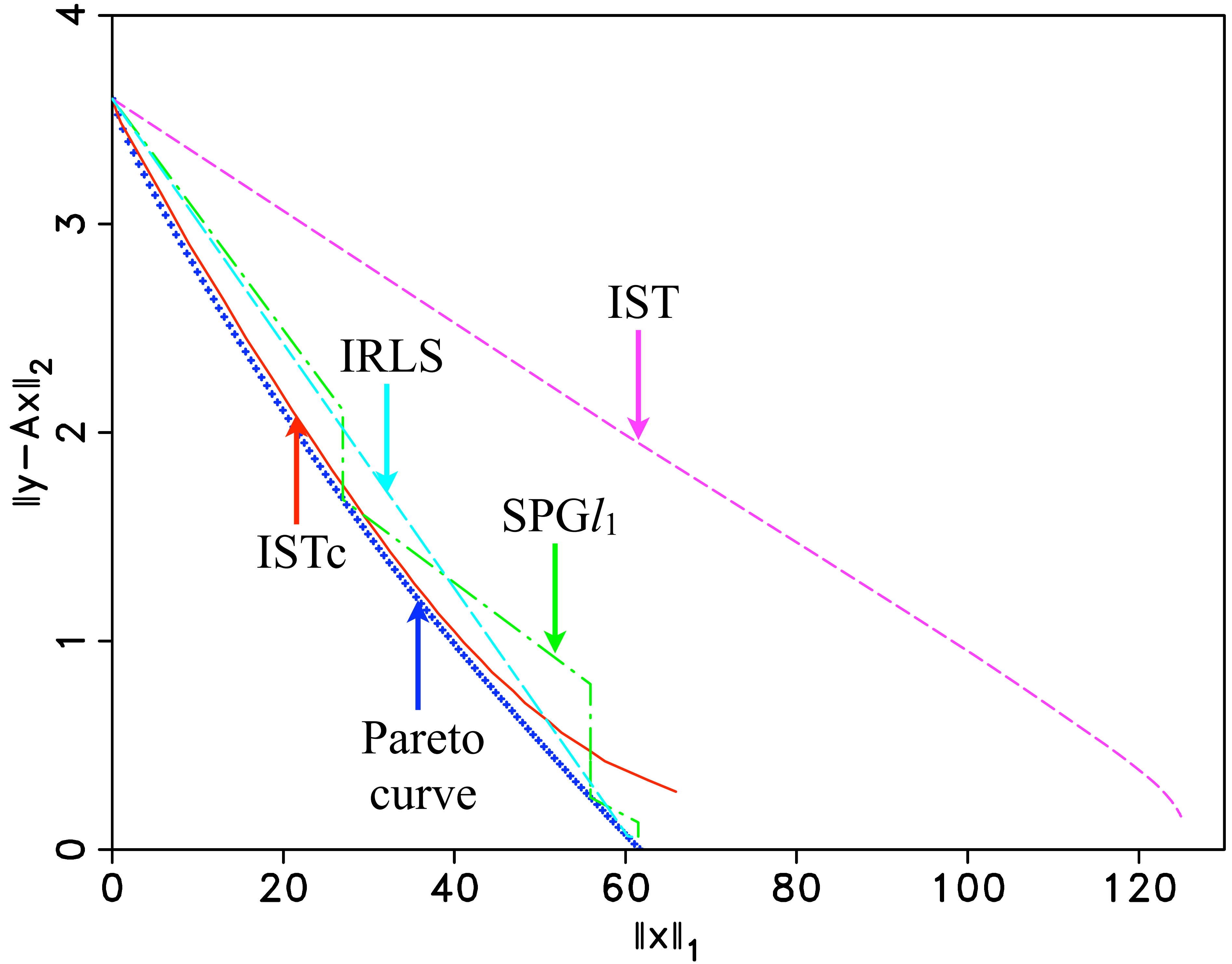
- challenging problem for solvers that attack QP_λ
  - solution only attained in the limit as $\lambda \to 0$
- applications
  - interpolation of noise-free wavefields
  - noise-free deconvolution



Solution paths: large-enough # of iterations



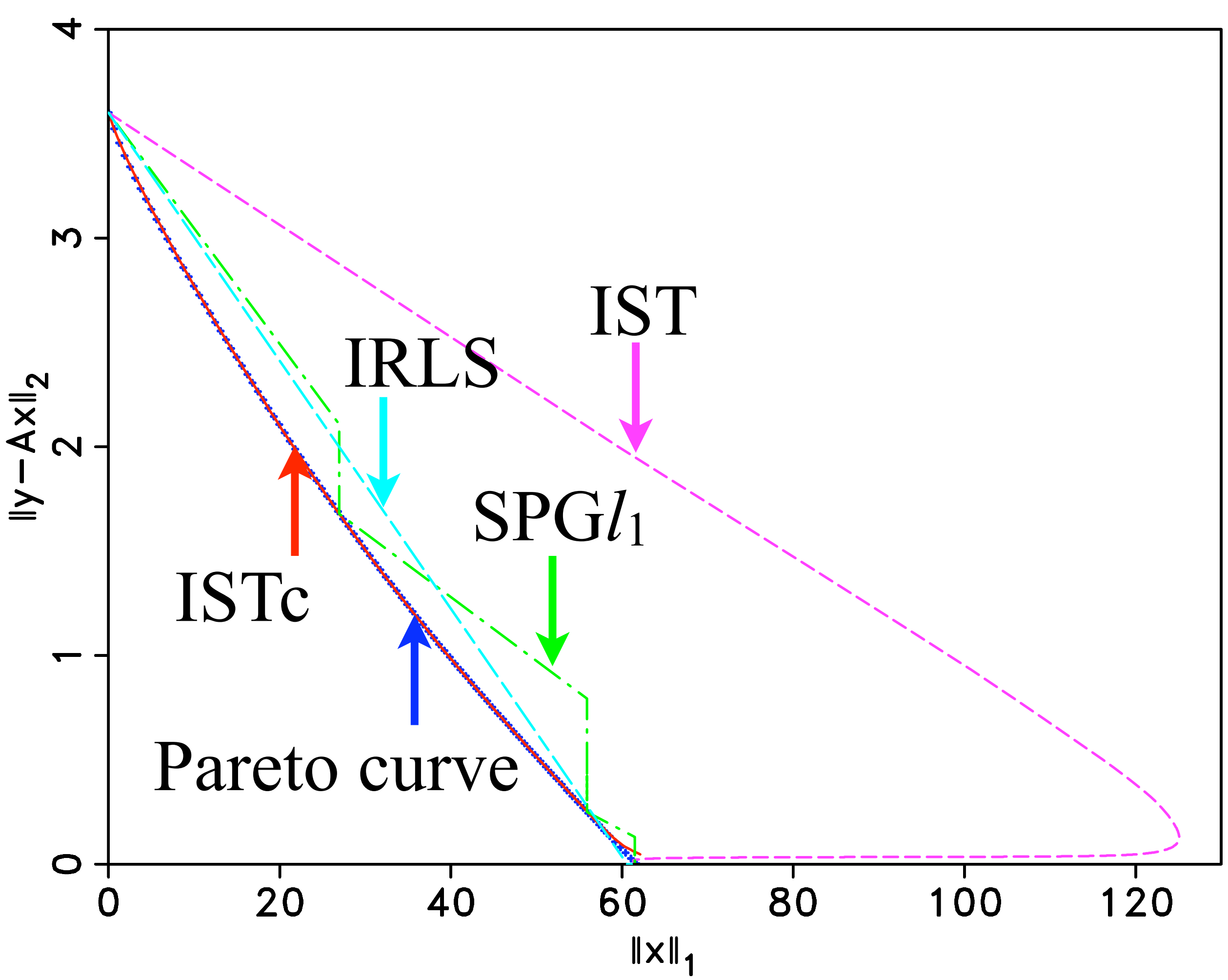
Solution paths: (very) limited # of iterations



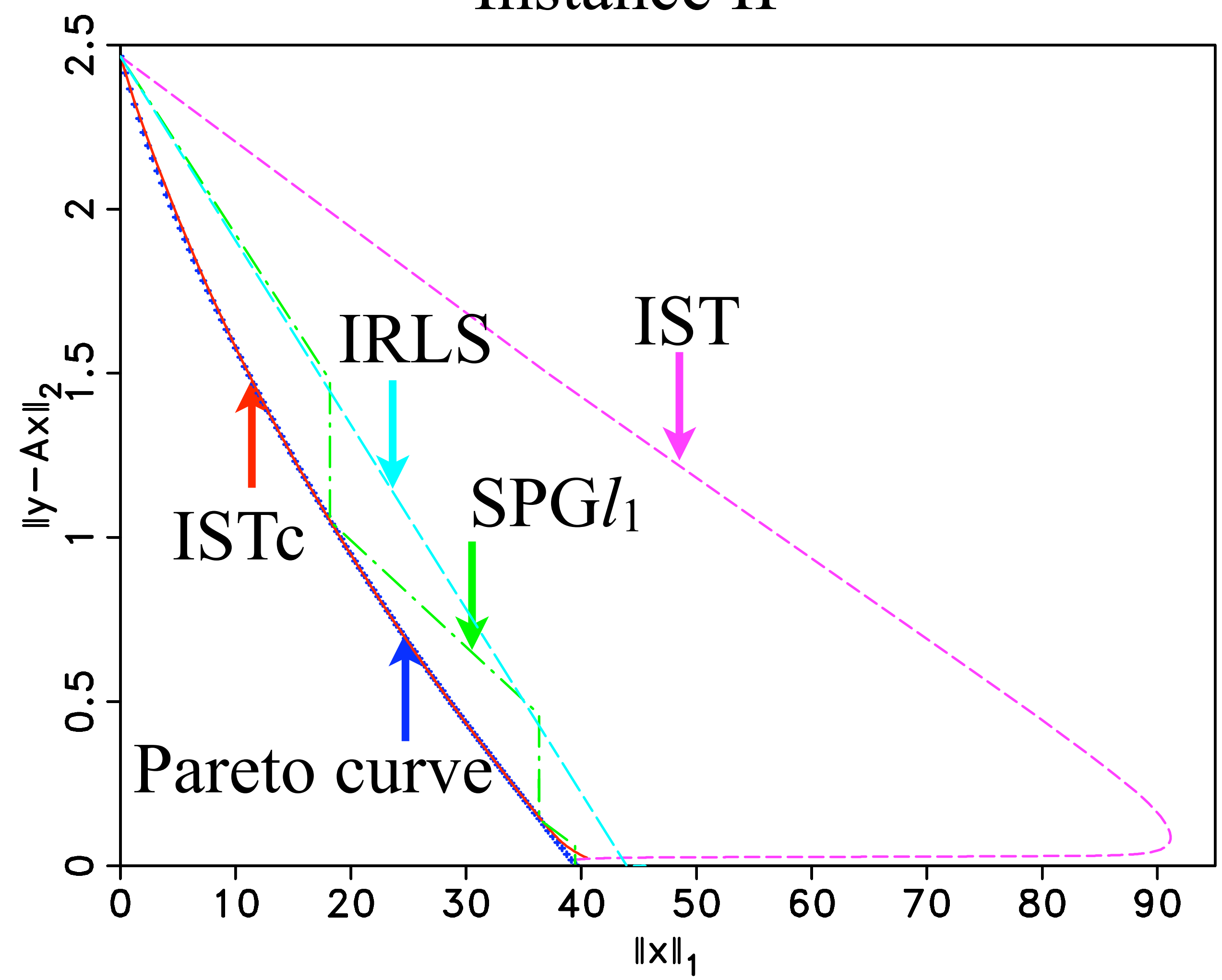
Exposing solvers' weakness

- sensitivity test
  - new instance of benchmark problem
  - same solvers tuning parameters as previous instance

Instance I

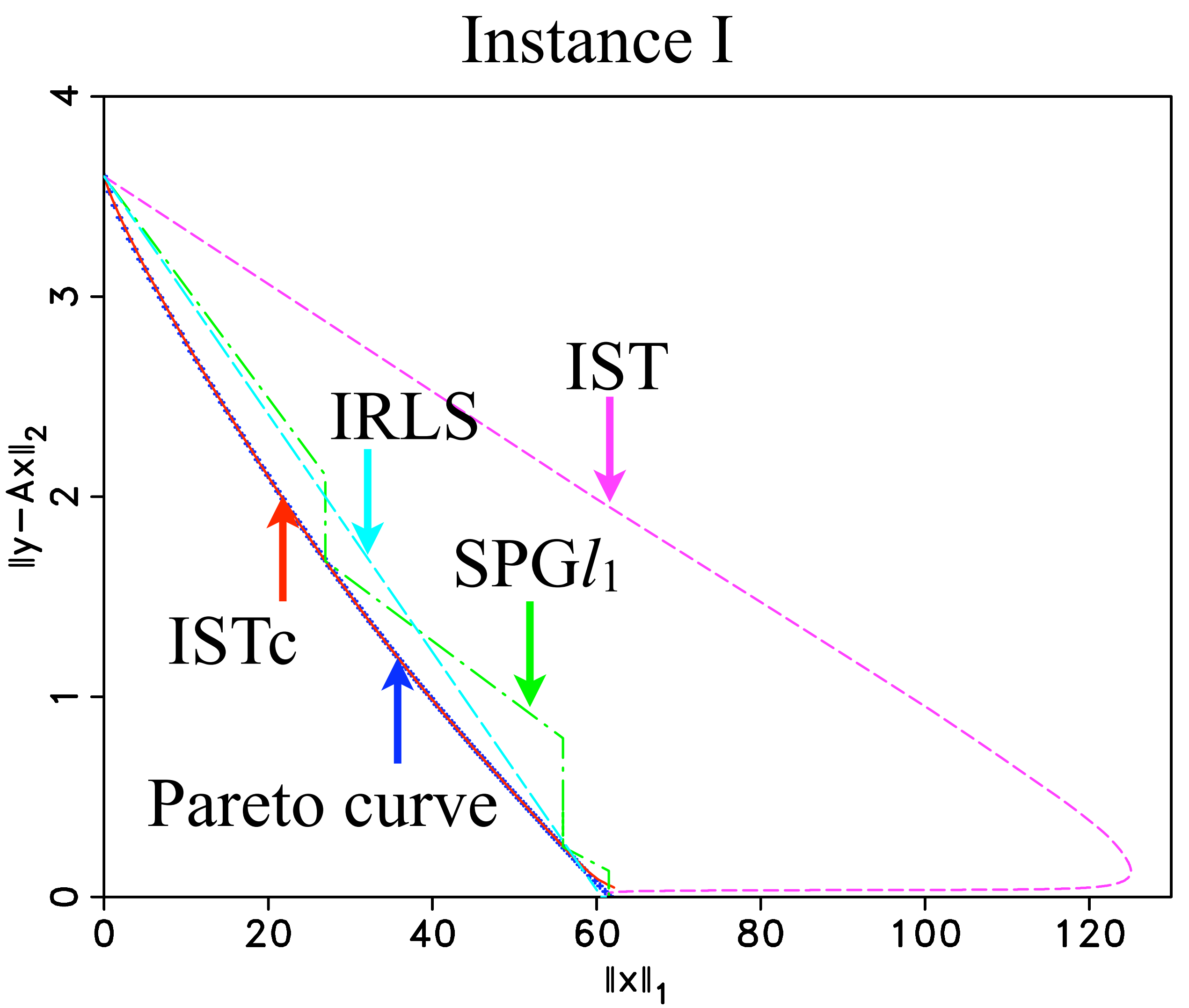


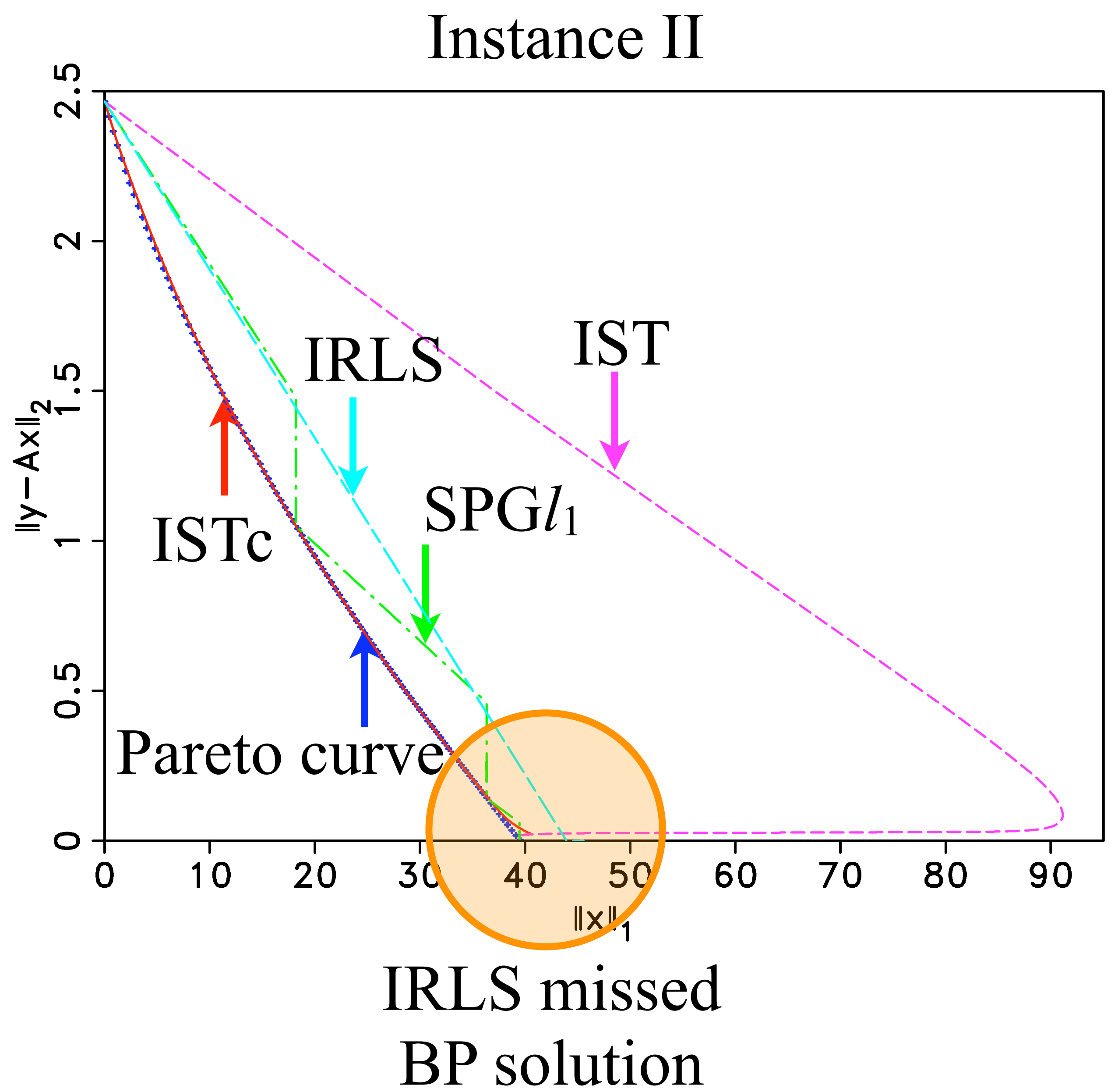
Instance II



Exposing solvers' weakness

- sensitivity test
  - new instance of benchmark problem
  - same solvers tuning parameters as previous instance





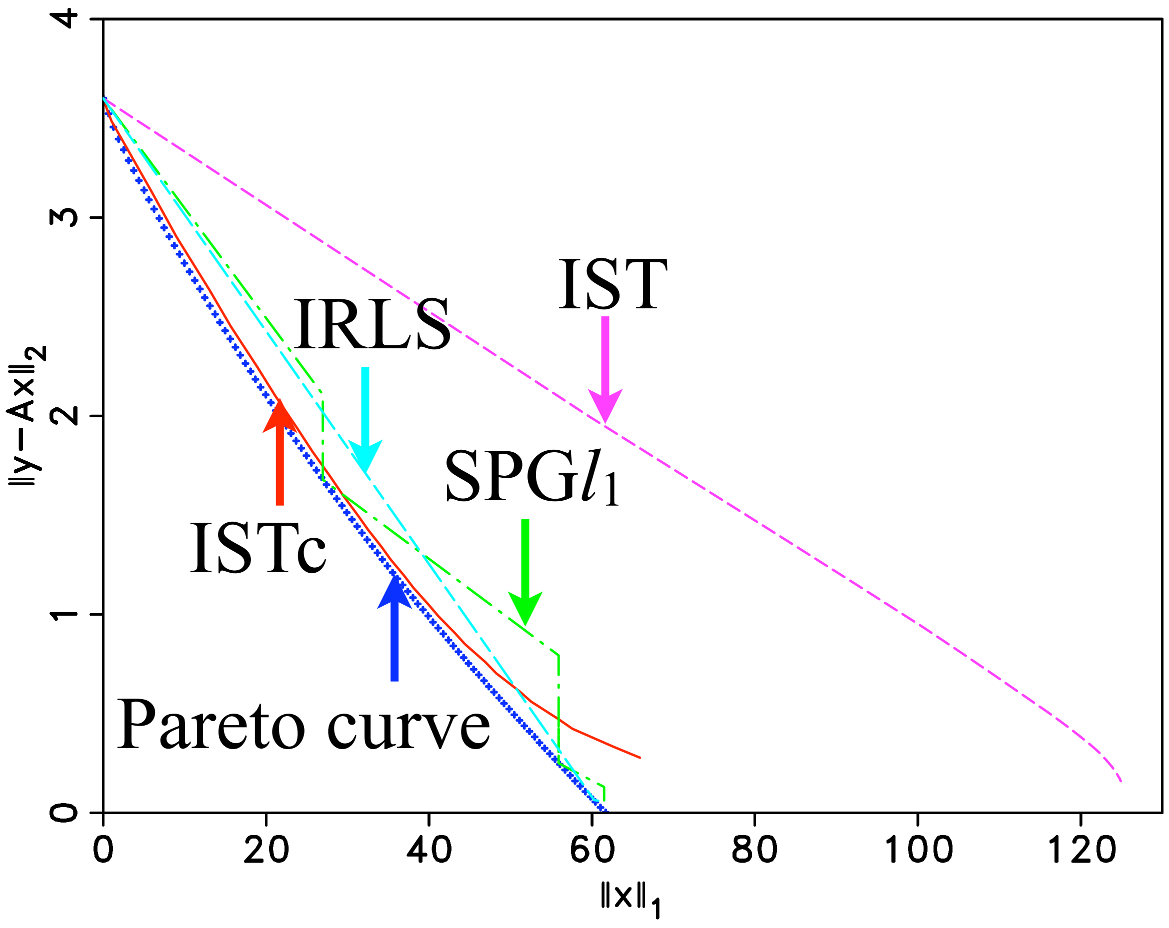
Conclusions

- Pareto curve
  - **optimal tradeoff** between $||\mathbf{y}-\mathbf{Ax}||_2$ and $||\mathbf{x}||_1$
  - **establishes the connection** between QP_λ, BP_σ, and LS_τ
  - **smooth**
    - good approximation to the curve obtained with VERY few interpolating points

- usage
  - explore the nature of a solver's iterations
    - informed decision on how to truncate solution process
    - **safely trade computational cost against solution accuracy**
  - evaluate the performance of one-norm solvers

Future work

- use new insights into one-norm solvers to improve them!!!
  - make SPGl_1 more agressive, yet avoiding overshooting beyond the BP solution
  - keep ISTc closer the Pareto curve towards the BP solution
    - i.e., better usage of the last few iterations
- apply analysis to geophysical problems



Acknowledgments

- S. Fomel, P. Sava, and the other developers of Madagascar (rsf.sourceforge.net)
- E. van den berg and M. P. Friedlander for SPGl_1 (http://www.cs.ubc.ca/labs/scl/spgl1)
- E. van den Berg, M. P. Friedlander, G. Hennenfent, F. J. Herrmann, R. Saab, and O. Yilmaz for Sparco (www.cs.ubc.ca/labs/scl/sparco/)

This presentation was carried out as part of the SINBAD project with financial support, secured through ITF, from the following organizations: BG, BP, Chevron, ExxonMobil, and Shell. SINBAD is part of the collaborative research & development (CRD) grant number 334810-05 funded by the Natural Science and Engineering Research Council (NSERC).