Geometric Multilevel Optimization For Discrete Tomography

Stefania Petra
Mathematical Imaging Group
Heidelberg University

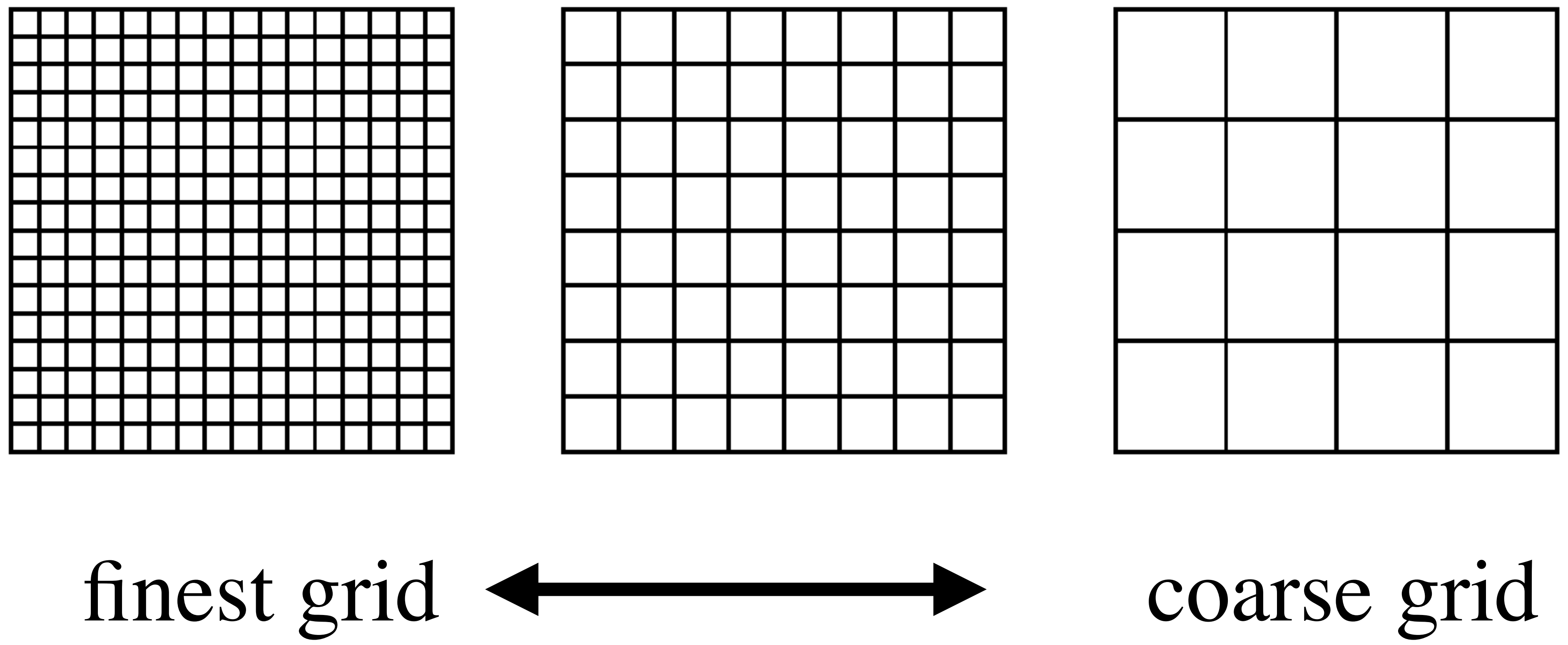
IN MEMORIAM OF CARLA PERI

TAIR 2022
Politecnico di Milano, May 2, 2022 – May 4, 2022

Motivation: Multigrid for PDEs

2D discrete Poisson equation

$$-\Delta u = f$$



$$u \in \mathbf{R}^n$$

Poisson solver complexity

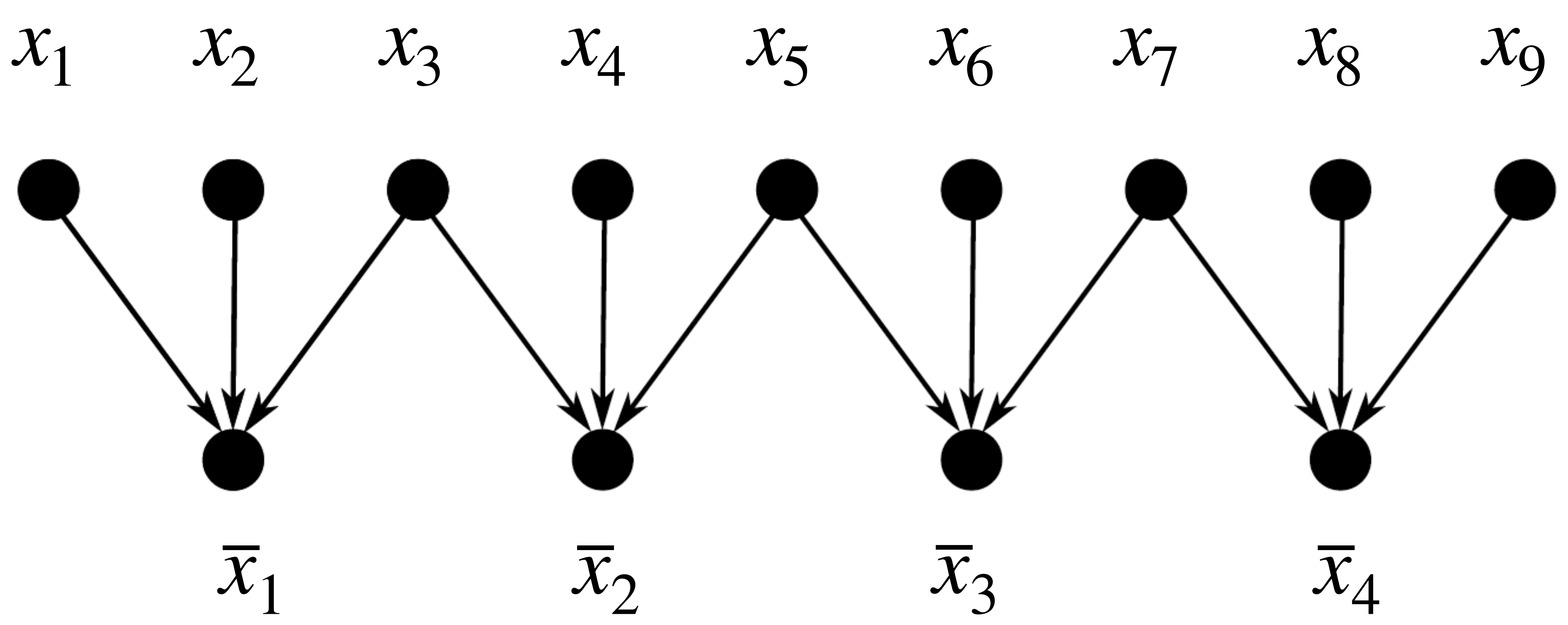
Gaussian elimination	$O(n^2)$
Jacobi iteration	$O(n^2 \log \varepsilon)$
Gauss–Seidel iteration	$O(n^2 \log \varepsilon)$
Conjugate gradient (CG)	$O(n^{3/2} \log \varepsilon)$
Fast Fourier transform (FFT)	$O(n \log n)$
Multigrid (iterative)	$O(n \log \varepsilon)$
Multigrid (FMG)	$O(n)$

This Talk: Geometric Multilevel Optimization

problem discretization on a 3D grid graph

Information transfer between levels

1. **Geometric** multigrid for PDEs



2. Algebraic

using the "natural" geometry (positivity, box constraints)

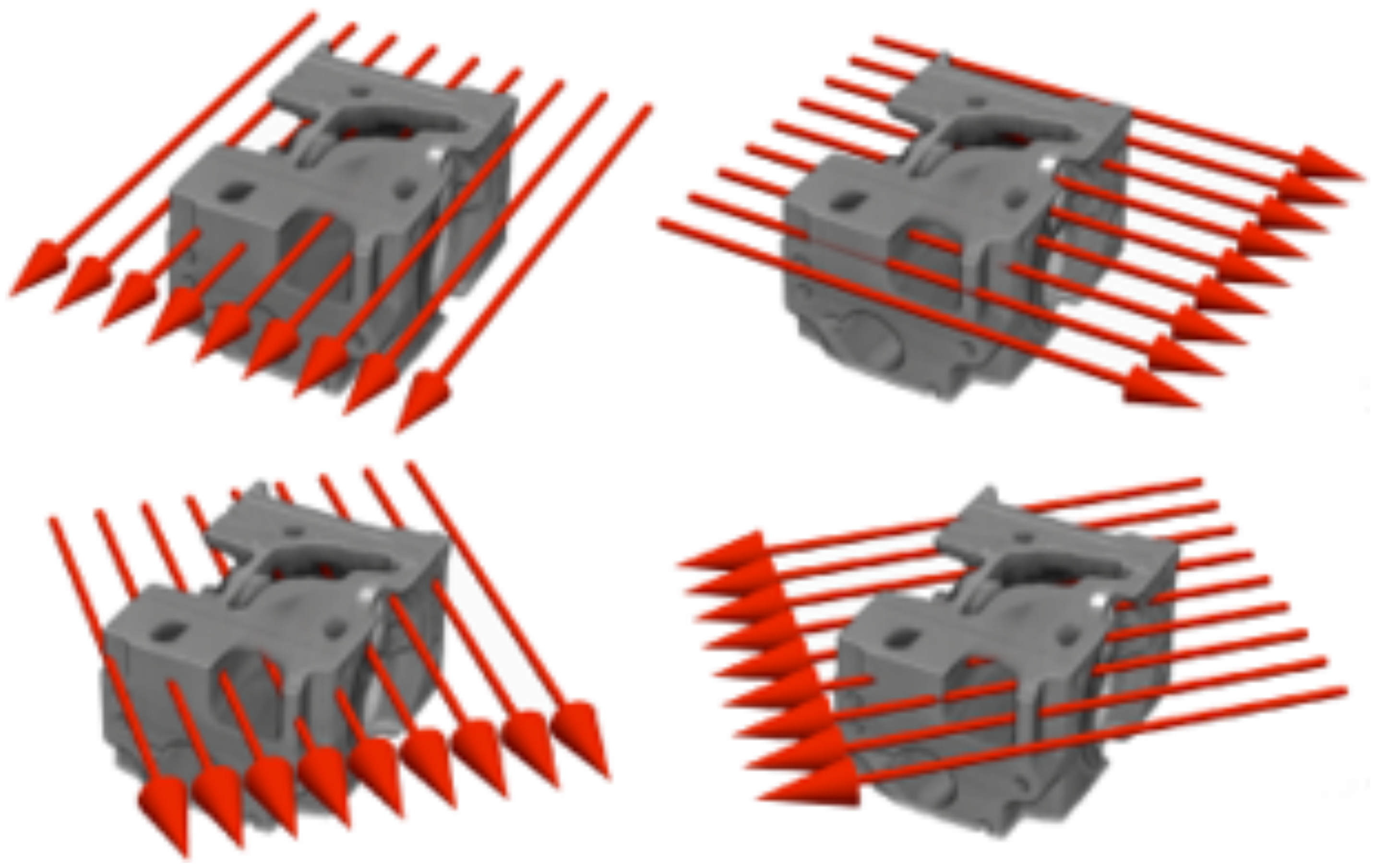
$$\min_{x \in C} f(x)$$

Challenges

- f convex but not quadratic
- non-smoothness due to constraints

Aplication Scenario: Discrete Tomography

Kuske, P., ***Performance Bounds For Co-/Sparse Box Constrained Signal Recovery***, 2019



$$\min_{[0,1]^n} \mathrm{KL}(Ax, b) + \lambda \|\nabla x\|_{1,\rho}$$

problem

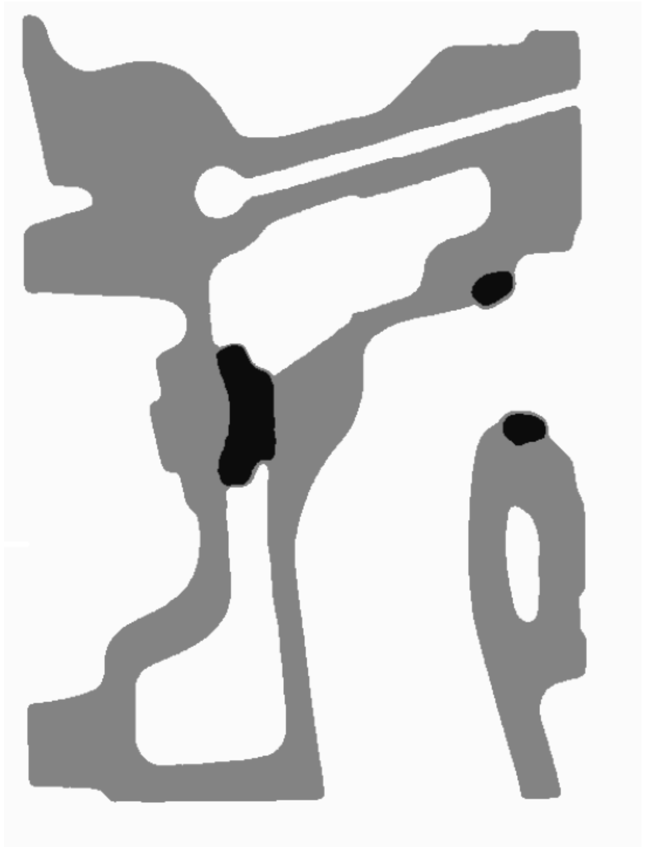
- limited angle/data: severely ill-posed problem
- integrality constraints

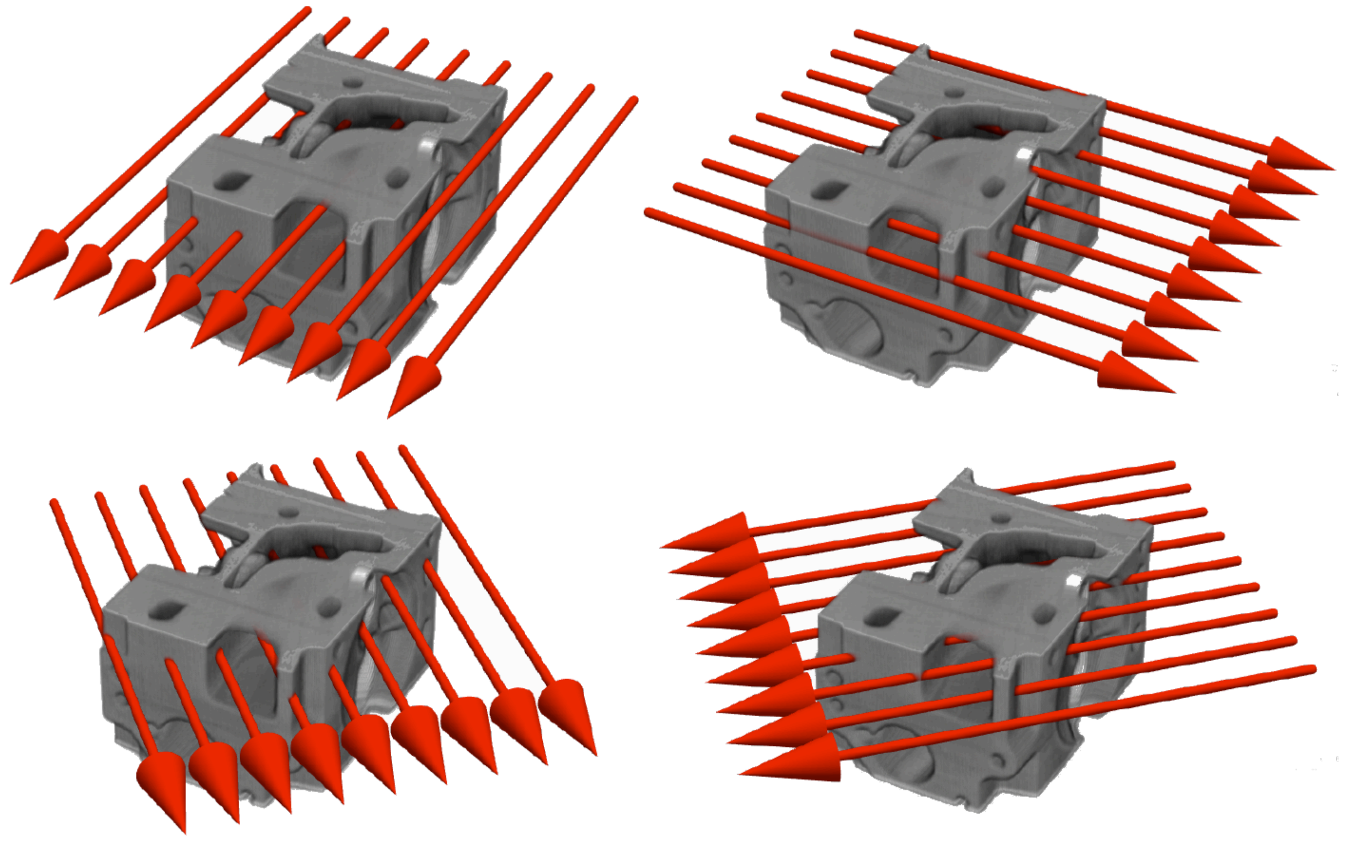
regularization

- constraints (few materials)
- sparse image gradient

Application Scenario: Discrete Tomography

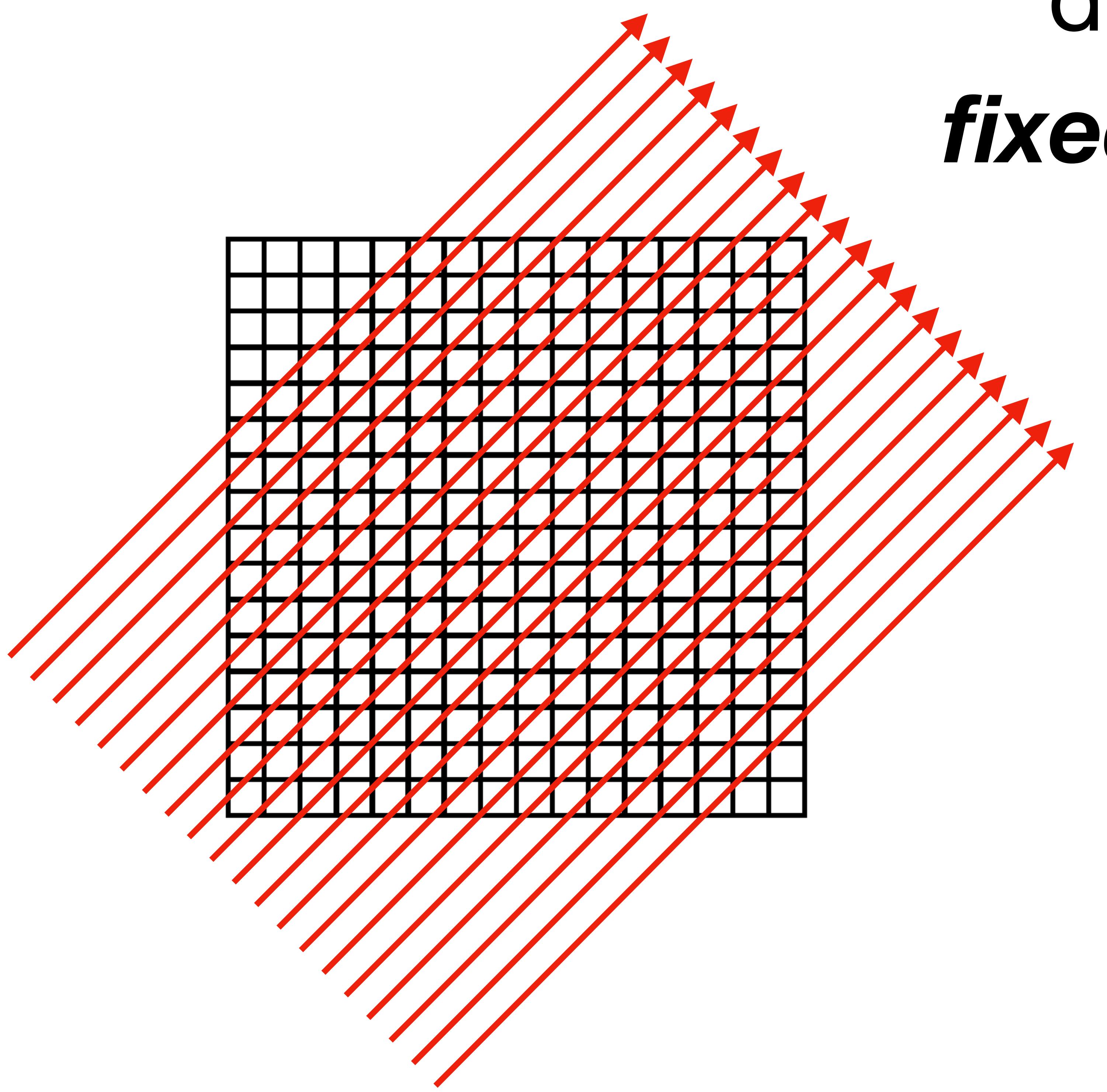
3D image
with finite range





few projection angles

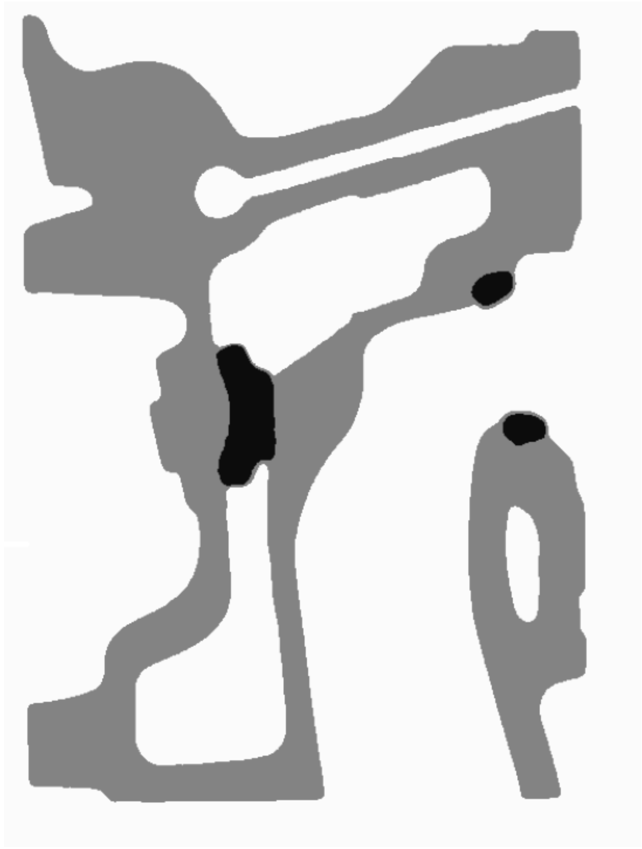
detector with
fixed resolution m

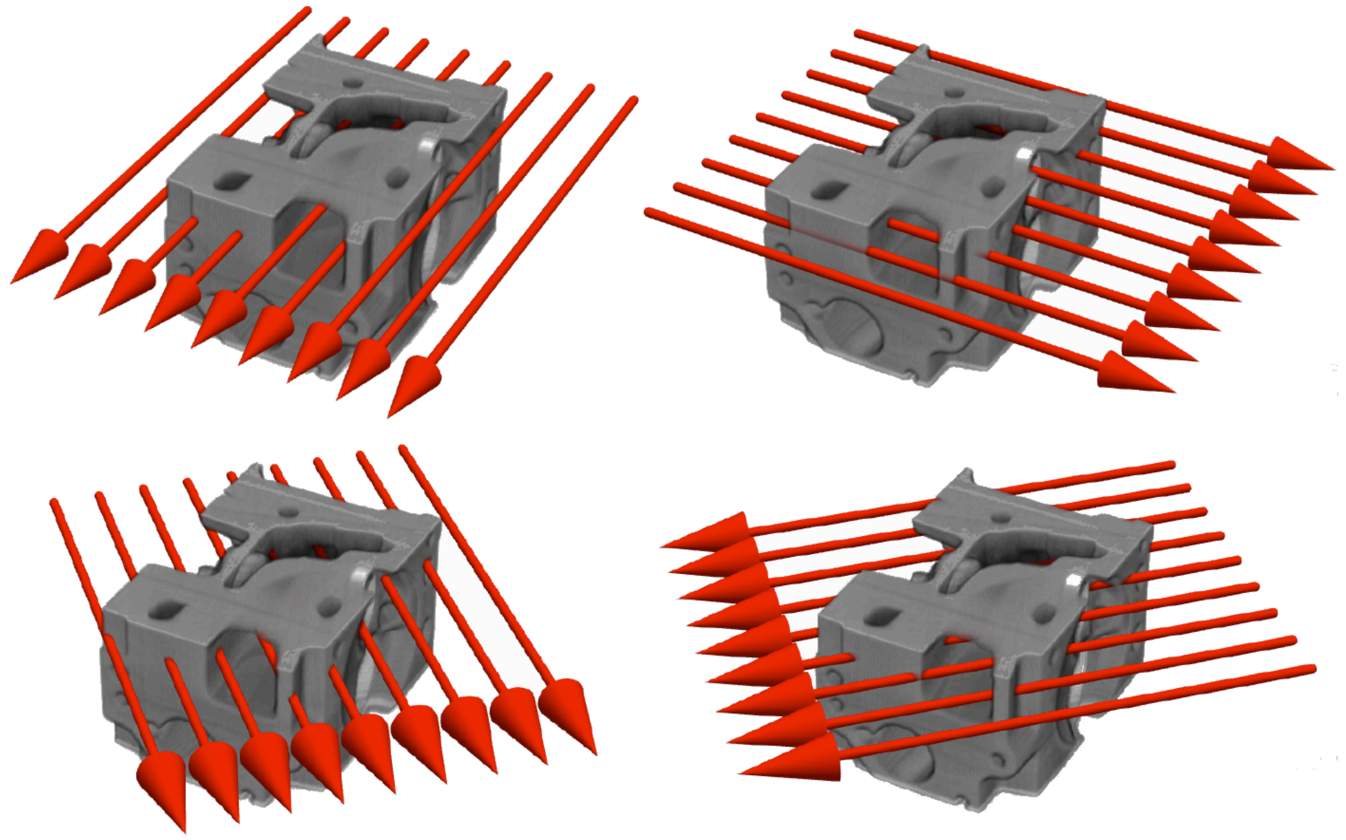


$$Ax = b, \quad A \in \mathbb{R}^{m \times n}, \quad m \ll n$$

Application Scenario: Discrete Tomography

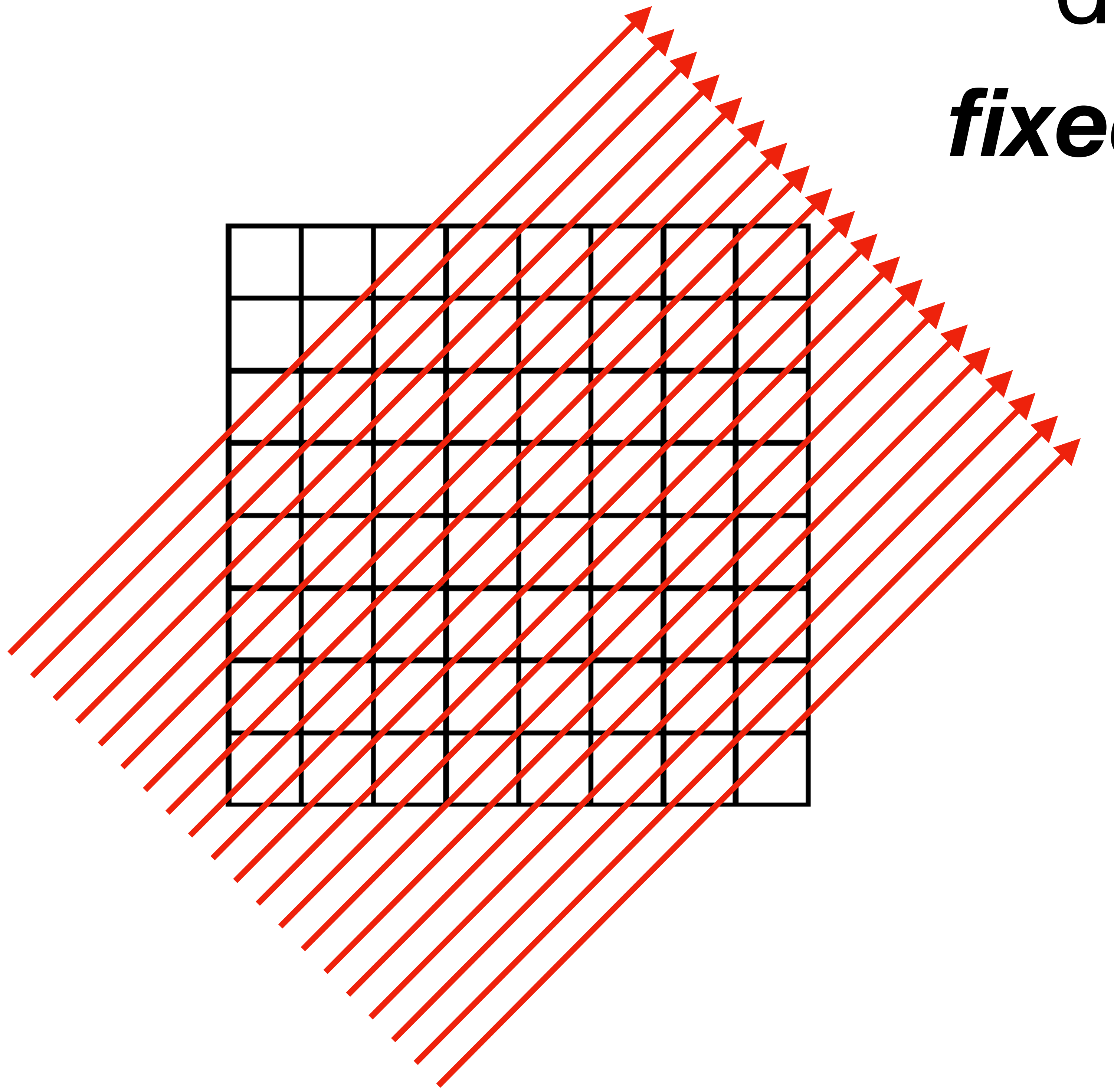
3D image
with finite range





few projection angles

detector with
fixed resolution m



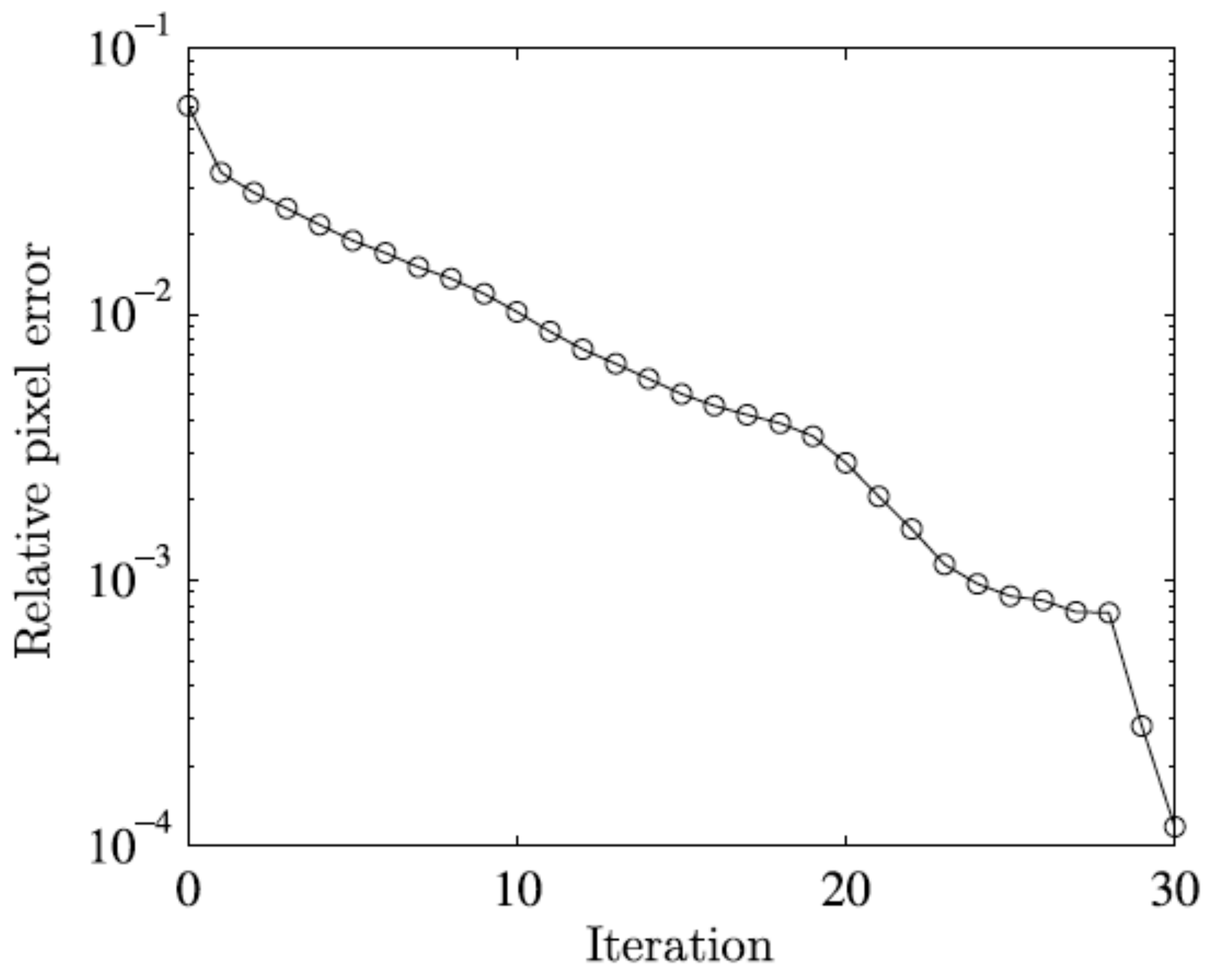
$$\bar{A}\bar{x} = b, \quad \bar{A} \in \mathbb{R}^{m \times \bar{n}},$$

$m/\bar{n}$
gets larger

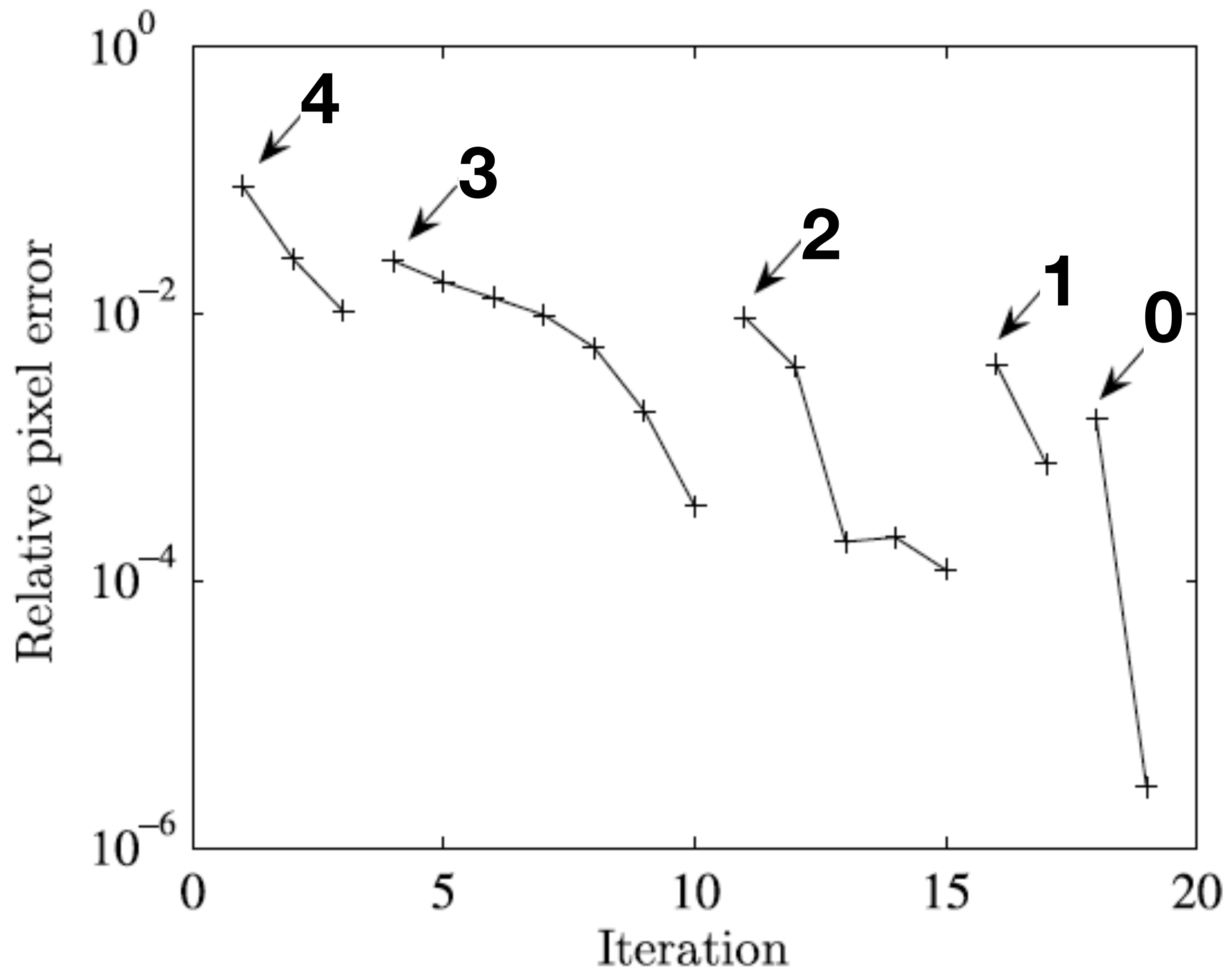
Related Work: Multiscale Acceleration for Discrete Tomography

S. Roux, H. Leclerc, F. Hild, **Efficient Binary Tomographic Reconstruction,** J Math Imaging Vis, 2014

A. Dabravolski, K. J. Batenburg, J. Sijbers, **A Multiresolution Approach to Discrete Tomography Using DART,** PLoS One. 2014



single scale



multiscale

0: fine scale *1024* x *1024*

1: coarse scale *512* x *512*

2: coarse scale *256* x *256*

3: coarse scale *128* x *128*

4: coarse scale *64* x *64*

Challenges

- principled information transfer between levels

Multilevel Optimization

smooth optimization

S. G. *Nash,* **A multigrid approach to discretized optimization problems.** Optimization Methods and Software, 2000

S. *Gratton, A. Sartenaer, P. L. Toint,* **Recursive trust-region methods for multiscale nonlinear optimization.** SIAM Journal on Optimization, 2008

W. Zaiwen, and D. Goldfarb, **A line search multigrid method for large-scale nonlinear optimization.** SIAM Journal on Optimization, 2009

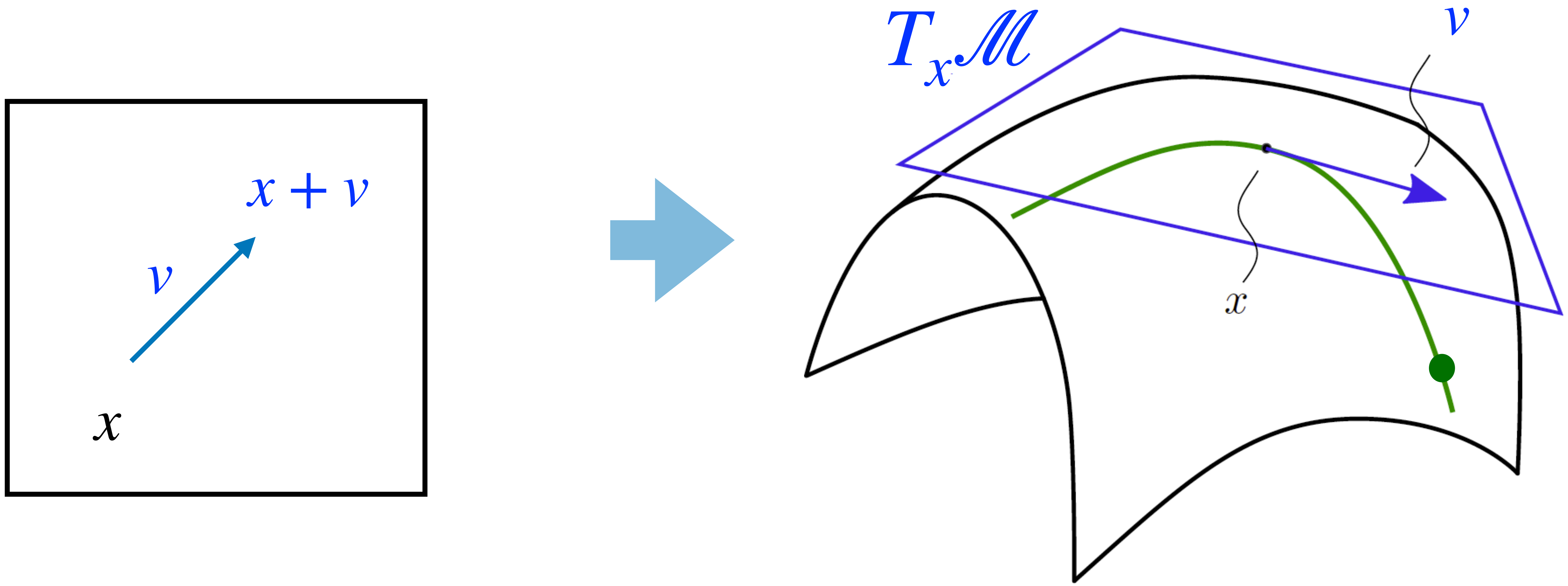
C. P. Ho, M. Kocvara, P. Parpas, **Newton-type Multilevel Optimization Method.** Optimization Methods and Software, 2019

nonsmooth composite convex optimization

M. Kocvara and S. Mohammed, **A first-order multigrid method for bound-constrained convex optimization.** Optimization Methods and Software, 2016

P. Parpas, **A multilevel proximal gradient algorithm for a class of composite optimization problems.** *SIAM J. Sci. Comput.*, 2017

Contribution



We incorporate constraints **<u>smoothly</u>** into MLO
by changing the geometry of the box

Outline

Optimization: local approximation

Two-grid optimization

Connection to nonlinear multigrid (FAS)

Connection to standard local approximation

Geometric two-grid optimization

Application to discrete tomography

Local Approximation: Smooth Unconstrained Optimization

$$\min_{x \in V} f(x)$$

At a current guess x select h by minimizing local approximation

$$f(x+h) \approx \underbrace{f(x) + \langle \nabla f(x), h \rangle}_{\text{linear}} + \underbrace{\frac{1}{2}\langle h, B_x h \rangle}_{\text{quadratic}}, \qquad B_x \succ 0$$

Update x by e.g. line search along h

$$x^+ = x + \alpha h$$

to improve $f(x^+) < f(x)$

Success Story: Smoothness & Convexity

$$\min_{x \in V} f(x)$$

with Lipschitz continuous gradient

At a current guess x select h by minimizing local approximation

$$f(x+h) \approx \underbrace{f(x) + \langle \nabla f(x), h \rangle}_{\text{linear}} + \underbrace{\frac{1}{2} L_f \|h\|^2}_{\text{quadratic}},$$

Update x by e.g. line search along h

$$x^+ = x + \frac{1}{L_f} h$$

no line search!

global convergence rate

$$f(x^k) - f^* = O\left(\frac{L_f}{k}\right)$$

Success Story: Smoothness, Convexity and Acceleration

$$\min_{x \in V} f(x)$$

with Lipschitz continuous gradient

Nesterov '83

accelerated convergence rate

$$f(x^k) - f^* = O\left(\frac{L_f}{k^2}\right)$$

At a current guess x select h by minimizing

$$f(x+h) \approx \underbrace{f(x) + \langle \nabla f(x), h \rangle}_{\text{linear}} + \underbrace{\frac{1}{2} L_f \|h\|^2}_{\text{quadratic}},$$

$$\min \; f + g$$

Update y by along d

$$y^+ = x + \frac{1}{L_f} h$$

nonsmooth composite convex

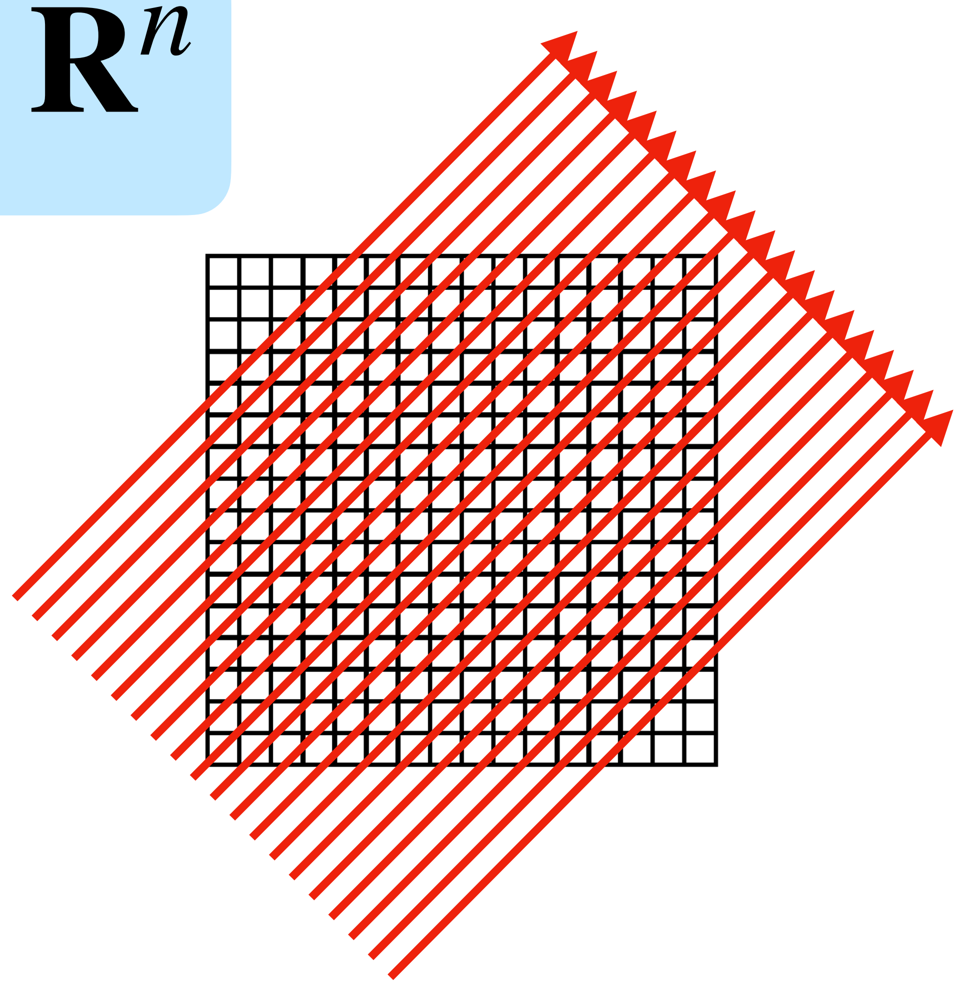
$$x^+ = (1-\gamma) y^+ + \gamma y$$

state of the art e.g. FISTA $O(1/k^2)$

no line search!

Example: Grid Dependent Smoothness

fine grid $x \in \mathbf{R}^n$



Least-squares with smoothed sparsity prior

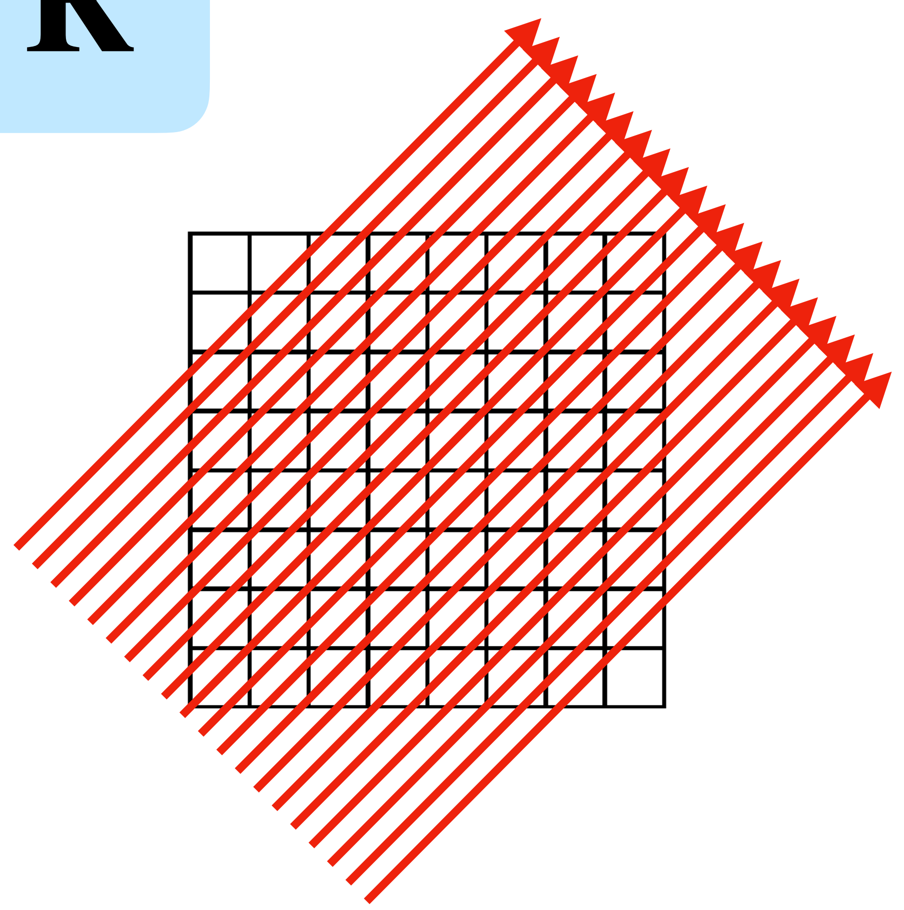
$$f(x) = \frac{1}{2}\|Ax - b\|^2 + \lambda \, \|\nabla x\|_{1,\rho}$$

Lipschitz constant

$$L_f = \|A\|^2 + 8\lambda/\rho$$

---

coarse grid $\bar{x} \in \mathbf{R}^{\bar{n}}$



$$\bar{f}(\bar{x}) = \frac{1}{2}\|\bar{A}\bar{x} - b\|^2 + \bar{\lambda} \, \|\bar{\nabla}\bar{x}\|_{1,\rho}$$

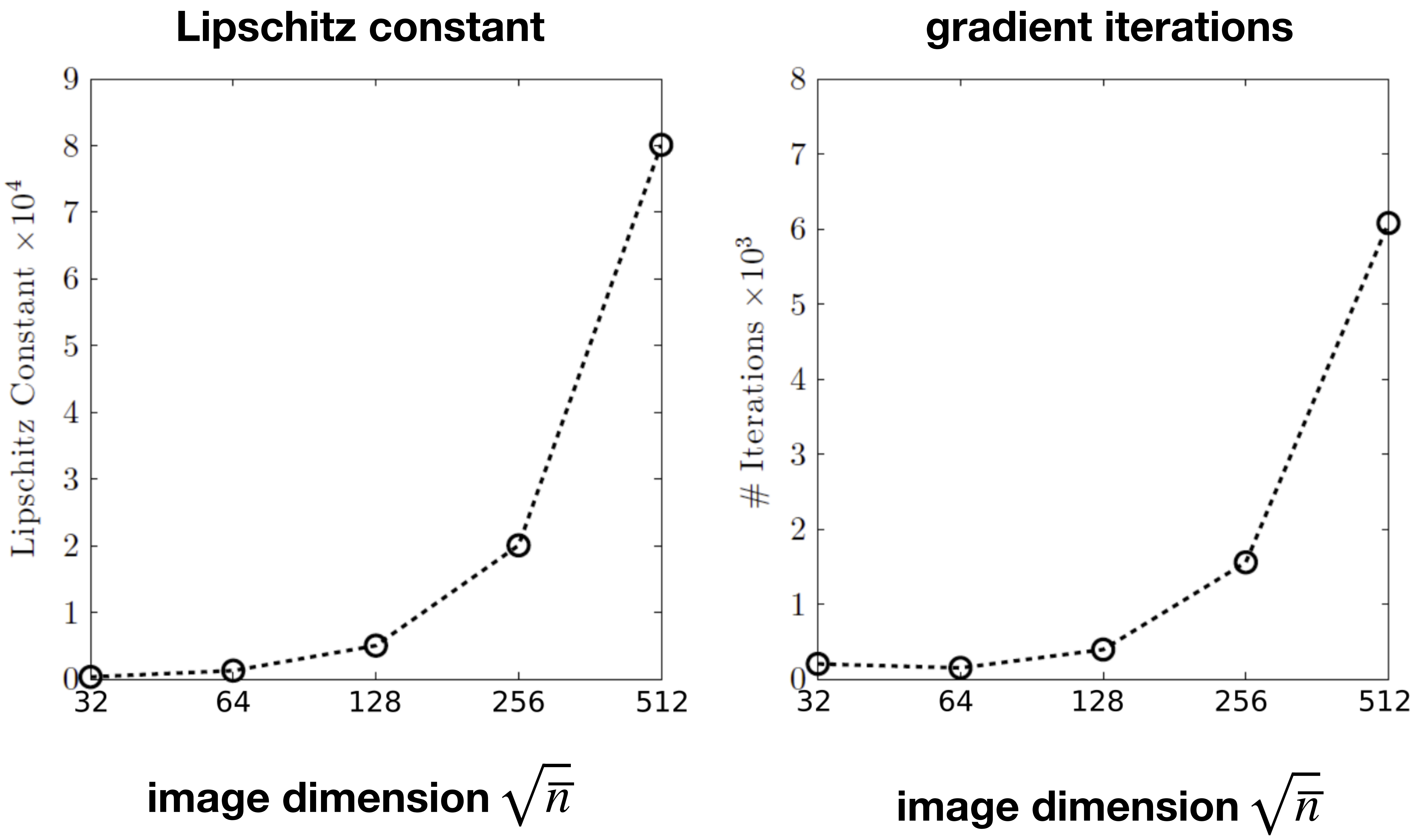
Lipschitz constant

$$L_{\bar{f}} = \|\bar{A}\|^2 + 8\bar{\lambda}/\rho$$

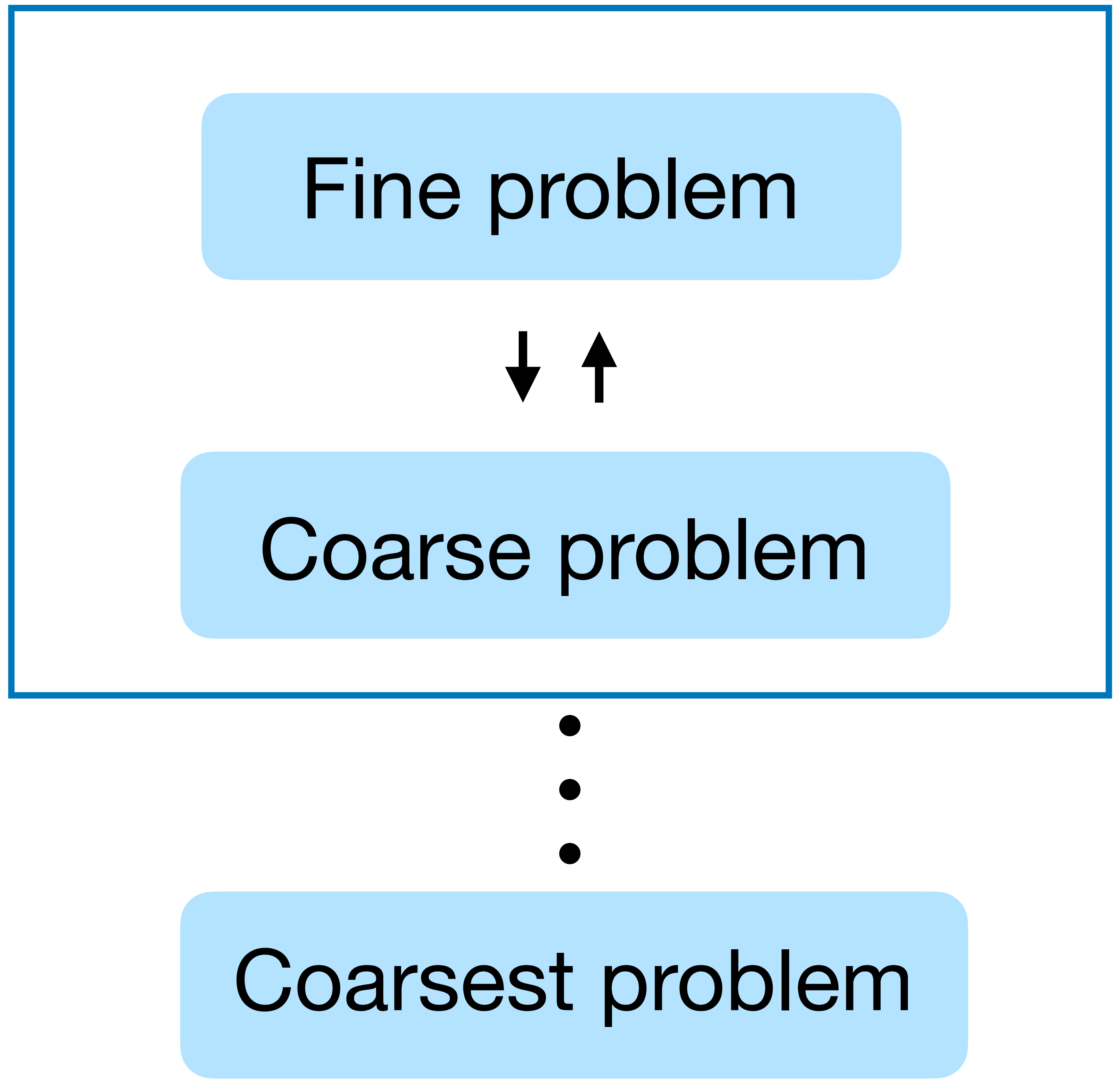
Example: Grid Dependent Smoothness

Least-squares with smoothed sparsity prior

$$\bar{f}(\bar{x}) = \frac{1}{2}\|\bar{A}\bar{x} - b\|^2 + \bar{\lambda}\,\|\bar{\nabla}\bar{x}\|_{1,\rho}$$



Structure: Hierarchy of Grid Dependent Problems



Outline

Optimization: local approximation

Two-grid optimization

Connection to nonlinear multigrid (FAS)

Connection to standard local approximation

Geometric two-grid optimization

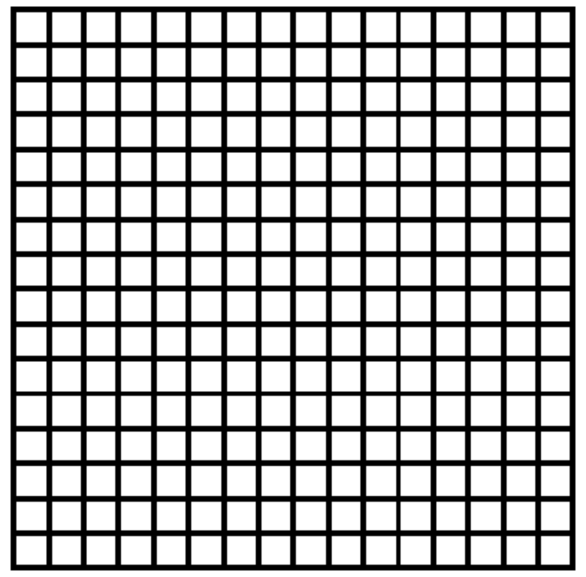
Application to discrete tomography

Hierarchical Representations on Two Grids

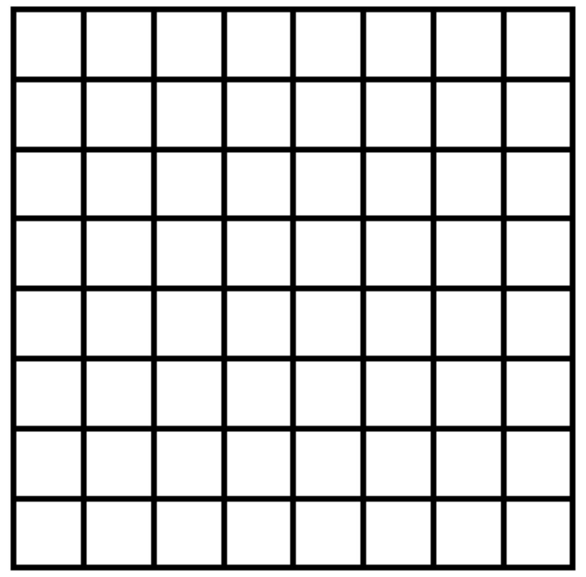
fine grid variable $x \in \mathbf{R}^n$

fine objective $f \in C^1(\mathbf{R}^n, \mathbf{R})$

fine



$$\min_{[0,1]^n} f(x) := \mathrm{KL}(Ax, b) + \lambda \|\nabla x\|_{1,\rho}$$



$$\min_{[0,1]^{\bar{n}}} \bar{f}(\bar{x}) := \mathrm{KL}(\bar{A}\bar{x}, \bar{b}) + \bar{\lambda} \|\bar{\nabla} \bar{x}\|_{1,\rho}$$

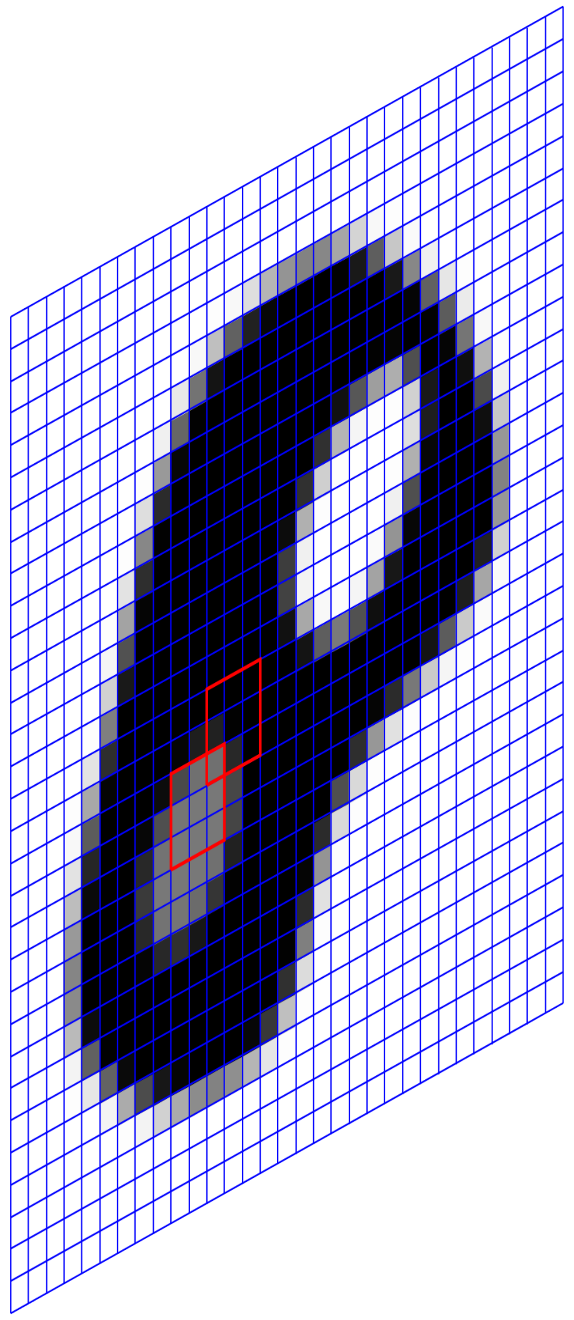
coarse

coarse objective $\bar{f} \in C^1(\mathbf{R}^{\bar{n}}, \mathbf{R})$

coarse grid variable $\bar{x} \in \mathbf{R}^{\bar{n}}$,

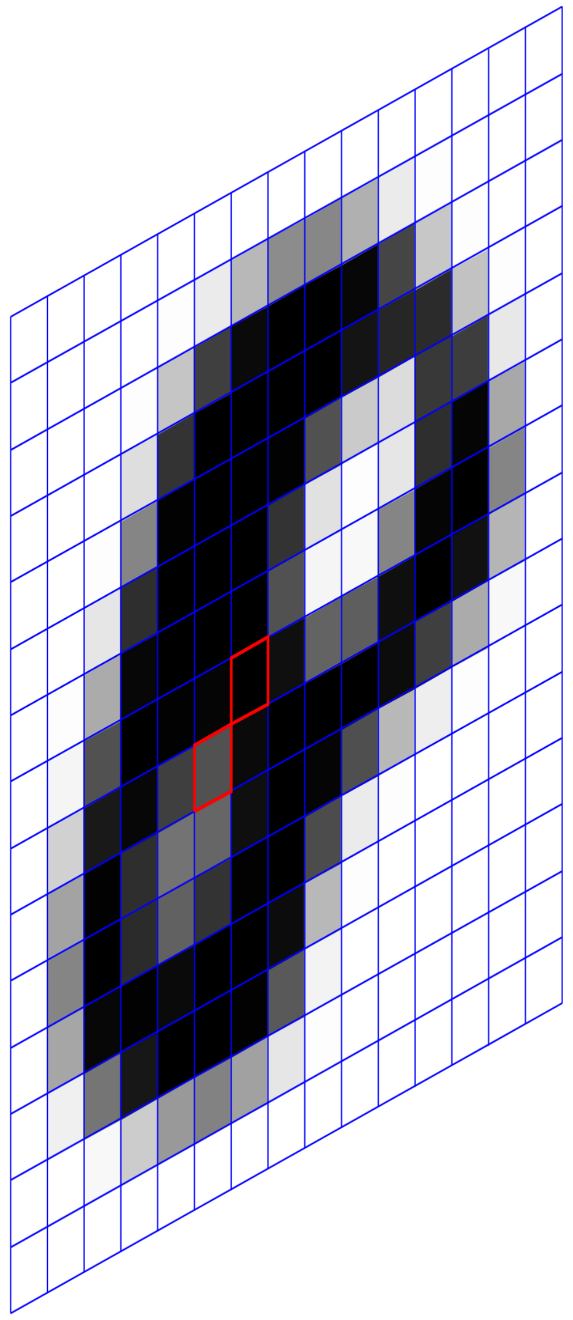
Intergrid Transfer Operators

x



$$x \xrightarrow[\text{restriction}]{R} \bar{x}$$

$$x \xleftarrow[P]{\text{prolongation}} \bar{x}$$

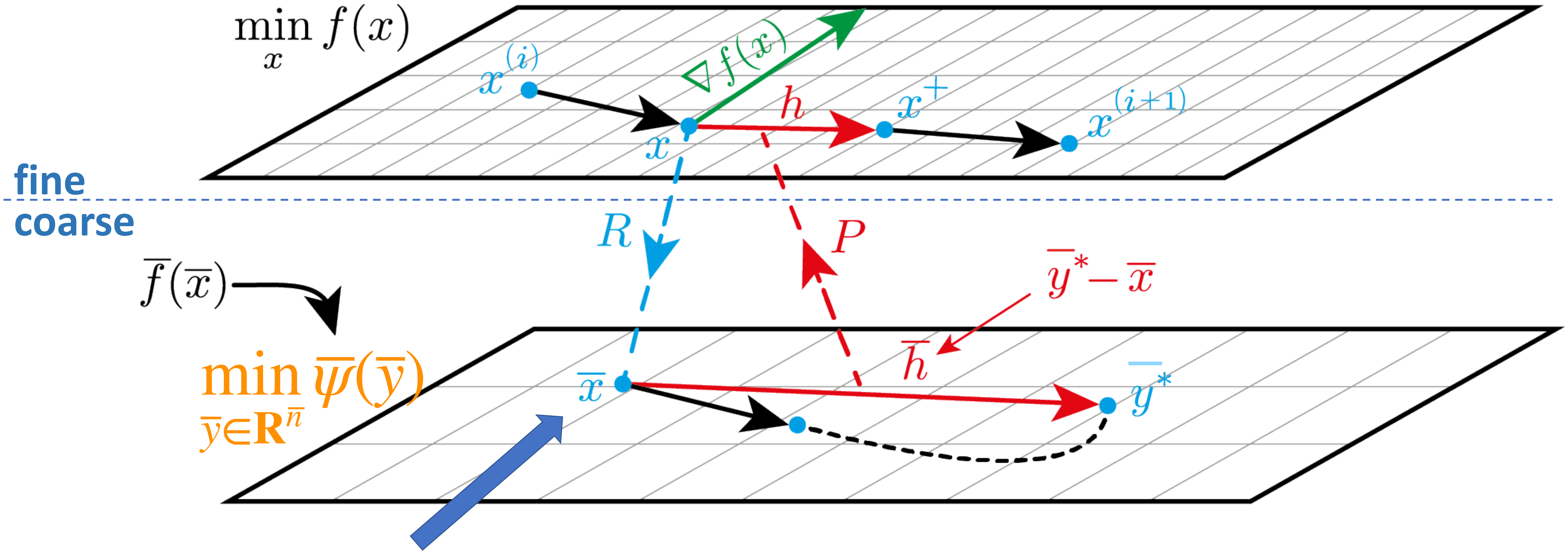


$\bar{x} = Rx$

$n > \bar{n}$, $\quad R : \mathbb{R}^n \to \mathbb{R}^{\bar{n}}$, $\quad P : \mathbb{R}^{\bar{n}} \to \mathbb{R}^n$, $\quad R = P^\top$

Euclidean Multilevel Optimization

Nash: *A multigrid approach to discretized optimization problems, 2000*



$$\overline{x} = Rx$$

Two Grid Approach, Coarse Model

for *current* fine grid variable $x \in \mathbf{R}^n$ define coarse grid model

$$\overline{\psi}(\overline{y}) = \bar{f}(\overline{y}) - \langle \overline{v}_x, \overline{y} - \overline{x} \rangle$$

with $\overline{v}_x = \nabla \bar{f}(Rx) - R \nabla f(x)$ and $\overline{x} = Rx$

(Nash, 2000, Gratton et al., 2008, Wen and Goldfarb, 2009)

Two Grid Approach, Coarse Model

for *current* fine grid variable $x \in \mathbf{R}^n$ define coarse grid model

$$\overline{\psi}(\overline{y}) = \bar{f}(\overline{y}) - \langle \overline{v}_x, \overline{y} - \overline{x} \rangle$$

with $\overline{v}_x = \nabla \bar{f}(Rx) - R\,\nabla f(x)$

first order coherence condition

$$\nabla \overline{\psi}(\overline{x}) = R\,\nabla f(x), \quad \boxed{\overline{x} = Rx}$$

starting iterate at coarse level

(Nash, 2000, Gratton et al., 2008, Wen and Goldfarb, 2009)

Two Grid Approach, Coarse Model

for *current* fine grid variable $x \in \mathbf{R}^n$, set $\bar{x} = Rx$ and rewrite coarse grid model

$$
\begin{aligned}
\bar{\psi}(\bar{y}) &= \bar{f}\bar{y}) - \langle \nabla \bar{f}(\bar{x}) - R\nabla f(x), \bar{y} - \bar{x} \rangle, \\
&= \bar{f}(\bar{y}) - \bar{f}(\bar{x}) - \langle \nabla \bar{f}(\bar{x}) - R\nabla f(x), \bar{y} - \bar{x} \rangle + \bar{f}(\bar{x}), \\
&= D_{\bar{f}}(\bar{y}, \bar{x}) + \langle R\nabla f(x), \bar{y} - \bar{x} \rangle + \cancel{const},
\end{aligned}
$$

with Bregman distance

$$
D_{\bar{f}}(\bar{y}, \bar{x}) = \bar{f}(\bar{y}) - \bar{f}(\bar{x}) - \langle \nabla \bar{f}(\bar{x}), \bar{y} - \bar{x} \rangle
$$

Two Grid Approach, Coarse Model

for *current* fine grid variable $x \in \mathbf{R}^n$, set $\bar{x} = Rx$ and rewrite coarse grid model

$$\begin{aligned} \bar{\psi}(\bar{y}) &= D_{\bar{f}}(\bar{y}, \bar{x}) + \langle R \nabla f(x), \bar{y} - \bar{x} \rangle \\ &= D_{\bar{f}}(\bar{y}, \bar{x}) + \langle \nabla f(x), P(\bar{y} - \bar{x}) \rangle \end{aligned}$$

$$\boxed{R = P^\top}$$

Galerkin condition

with Bregman distance

$$D_{\bar{f}}(\bar{y}, \bar{x}) = \bar{f}(\bar{y}) - \bar{f}(\bar{x}) - \langle \nabla \bar{f}(\bar{x}), \bar{y} - \bar{x} \rangle$$

Two Grid Approach, Coarse Model

for *current* fine grid variable $x \in \mathbf{R}^n$, set $\bar{x} = Rx$

$$\bar{\psi}(\bar{y}) = \underbrace{D_{\bar{f}}(\bar{y}, \bar{x})}_{\geq 0} + \langle \nabla f(x), P(\bar{y} - \bar{x}) \rangle$$

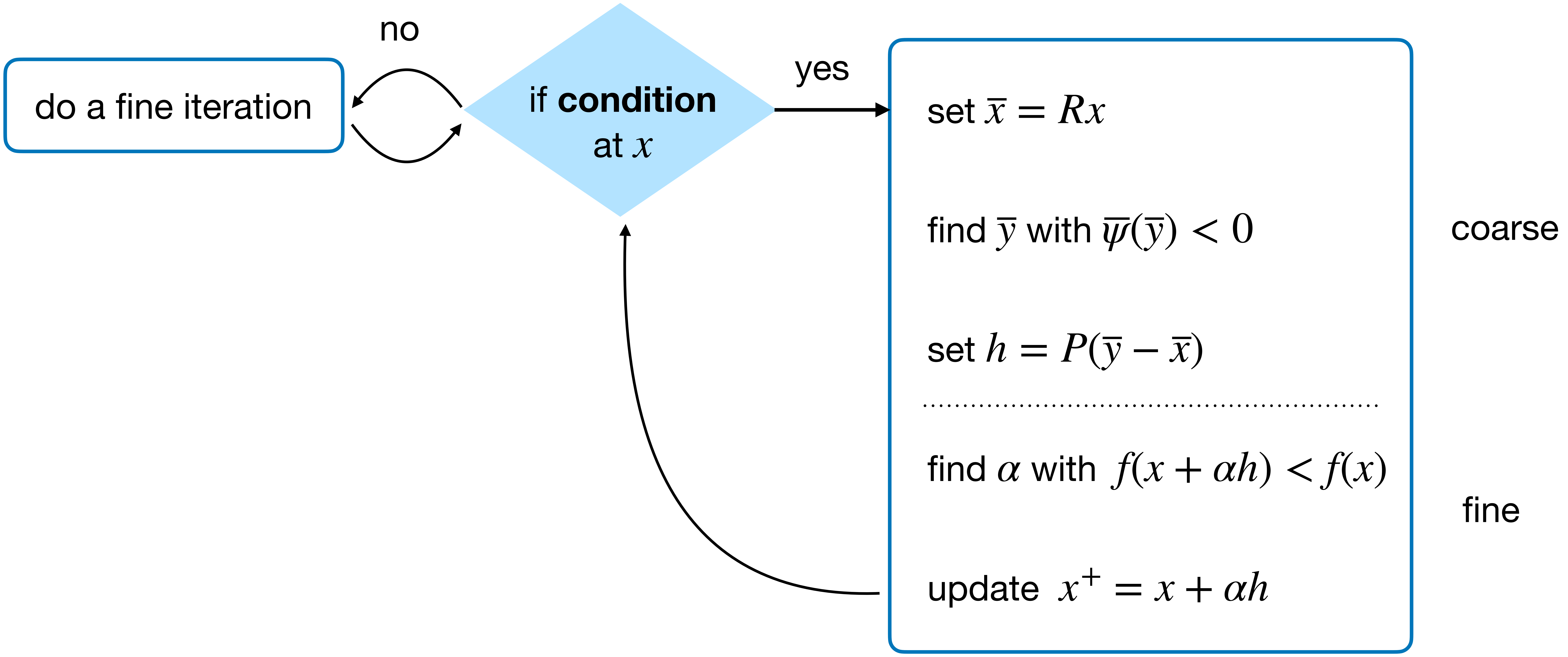
$$R = P^{\top}$$

if $\bar{f}$ convex

$$\bar{\psi}(\bar{y}) < 0 \quad \Longrightarrow \quad \langle \nabla f(x), h \rangle < 0, \quad h = P(\bar{y} - \bar{x})$$

descent direction for f

Two Grid Approach, Coarse Model



Two Grid Approach, Coarse Model

do a fine iteration

no

if **condition** at x

yes

set $\bar{x} = Rx$

find $\bar{y}$ with $\bar{\psi}(\bar{y}) < 0$

set $h = P(\bar{y} - \bar{x})$

coarse

find α with $f(x + \alpha h) < f(x)$

update $x^+ = x + \alpha h$

fine

Coarse correction if:

$\|R\nabla f(x)\| \geq \kappa \|\nabla f(x)\|$ **and** $\|R\nabla f(x)\| > \varepsilon$

$\kappa \in (0, \min(1, \|R\|)), \quad \varepsilon \in (0,1)$

Web, Goldfarb, *SIAM J Optim, 2009*

Outline

Optimization: local approximation

Two-grid optimization

Connection to nonlinear multigrid (FAS)

Connection to standard local approximation

Geometric two-grid optimization

Application to discrete tomography

Classical Multigrid Methods

Consider e.g. some elliptic PDE, discretize on some grid $Ax = b$

Discretisation $\bar{A}\bar{x} = \bar{b}$ of the same problem on coarser grid

$$
\begin{tikzcd}
{} \arrow[r] & x \arrow[r] \arrow[d, "R"] & r = b - Ax \arrow[d, "R"] & & h \arrow[r] & x + h \\
& \bar{x} \arrow[r] & \bar{r} = \bar{b} - \bar{A}\bar{x} \arrow[rr] & & \bar{A}\bar{h} = \bar{r} \arrow[u, "P"] & \text{coarse grid correction}
\end{tikzcd}
$$

- Relaxation methods no not eliminate smooth components of the error efficiently

- Smooth components projected on a coarser grid appear more oscillatory

Full Approximation Scheme

$\nabla f(x^*) = 0$ nonlinear equation

$$
\begin{tikzcd}
{} \arrow[r] & x \arrow[r] \arrow[d, "R"] & r = 0 - \nabla f(x) \arrow[d, "R"] & h \arrow[r] & \begin{array}{c}\text{line search} \\ x + \alpha h\end{array} \arrow[r] & {} \\
& \bar{x} & \bar{r} = 0 - R \nabla f(x) \arrow[r] & \nabla \bar{f}(\bar{x} + \bar{h}) - \nabla \bar{f}(\bar{x}) = \bar{r} \quad \text{coarse grid correction} \arrow[u, "P"] & &
\end{tikzcd}
$$

$$\Updownarrow$$

inexact optimization $0 \overset{!}{>} \bar{\psi}(\bar{x} + \bar{h}) = \bar{f}(\bar{x} + \bar{h}) - \langle \nabla \bar{f}(\bar{x}) - R \nabla f(x), \bar{h} \rangle + C(\bar{x})$

$$= D_{\bar{f}}(\bar{x} + \bar{h}, \bar{x}) + \langle R \nabla f(x), \bar{h} \rangle + C'(\bar{x})$$

Outline

Optimization: local approximation

Two-grid optimization

Connection to nonlinear multigrid (FAS)

Connection to standard local approximation

Geometric two-grid optimization

Application to discrete tomography

Connection to Standard Local Approximation

``hierarchical approximation", multilevel optimization

$$\bar{\psi}(\bar{y}) = \langle \nabla f(x), P(\bar{y} - \bar{x}) \rangle + D_{\bar{f}}(\bar{y}, \bar{x})$$

$$D_{\bar{f}}(\bar{y}, \bar{x})$$

$$= \frac{1}{2} \langle \bar{y} - \bar{x}, \nabla^2 \bar{f}(z), \bar{y} - \bar{x} \rangle,$$

$$z \in \{(1-t)\bar{x} + t\bar{y}\}_{t \in [0,1]}$$

1st order approximation $\updownarrow$

$\updownarrow$ 2nd order approximation

$$q(y) = f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2} \langle y - x, B_x(y - x) \rangle, \qquad B_x \succ 0$$

quadratic approximation, quasi Newton

Outline

Optimization: local approximation

Two-grid optimization

Connection to nonlinear multigrid (FAS)

Connection to standard local approximation

Geometric two-grid optimization

Application to discrete tomography

Smooth Bound Constrained Convex Optimization

least-squares with sparsity prior and box constraints

$$\|Ax - b\|^2 + \lambda \, \|\nabla x\|_1, \quad x \geq \mathbf{0}, \quad x \leq \mathbf{1} \quad \lambda > 0$$

squared distance → Bregman divergence

non-smooth prior → Moreau env.

hypercube → Riemannian manifold

geometry?

geometry!

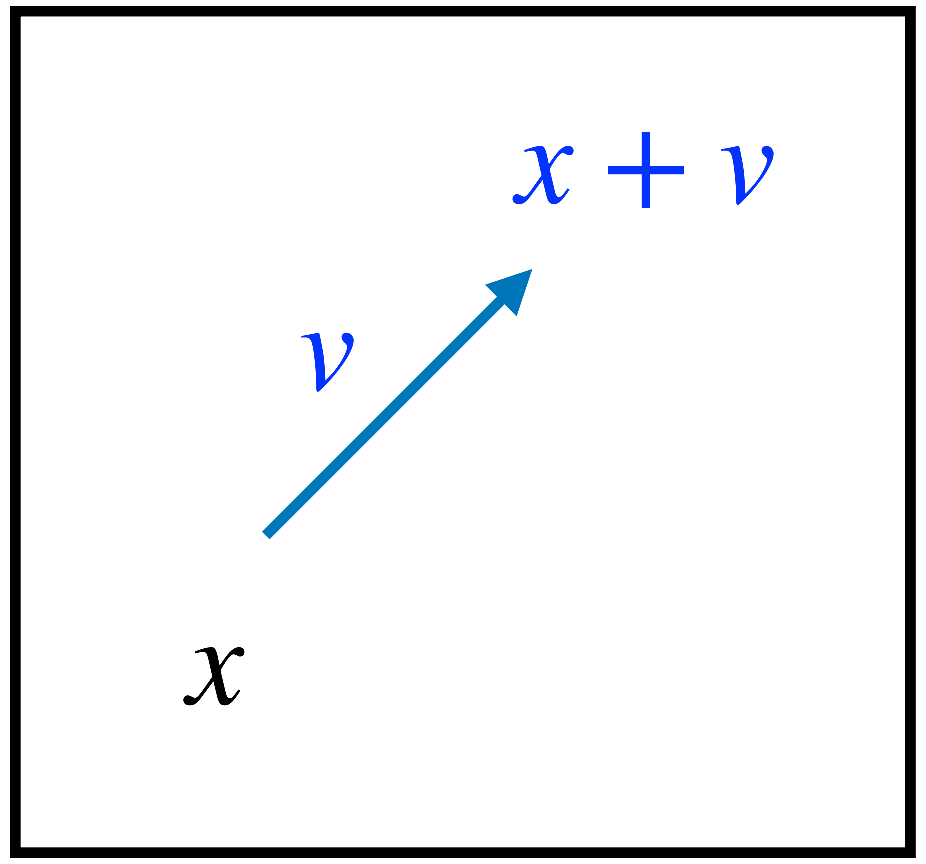
$$f(x) = \mathrm{KL}(Ax, b) + \lambda \, \|\nabla x\|_{1,\rho}, \quad x \in ((0,1)^n, h) = (\mathscr{M}, h), \quad \lambda > 0$$

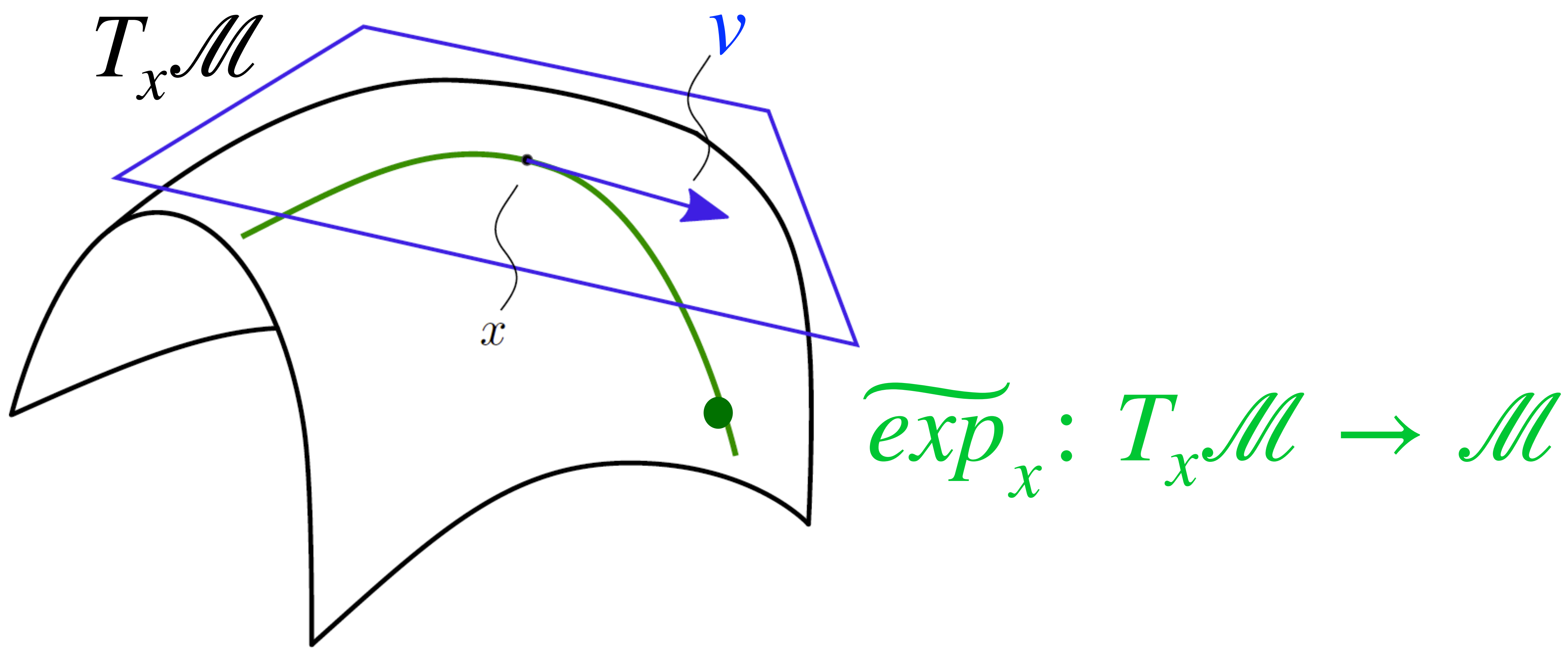
smooth non-quadratic data term with sparsity prior

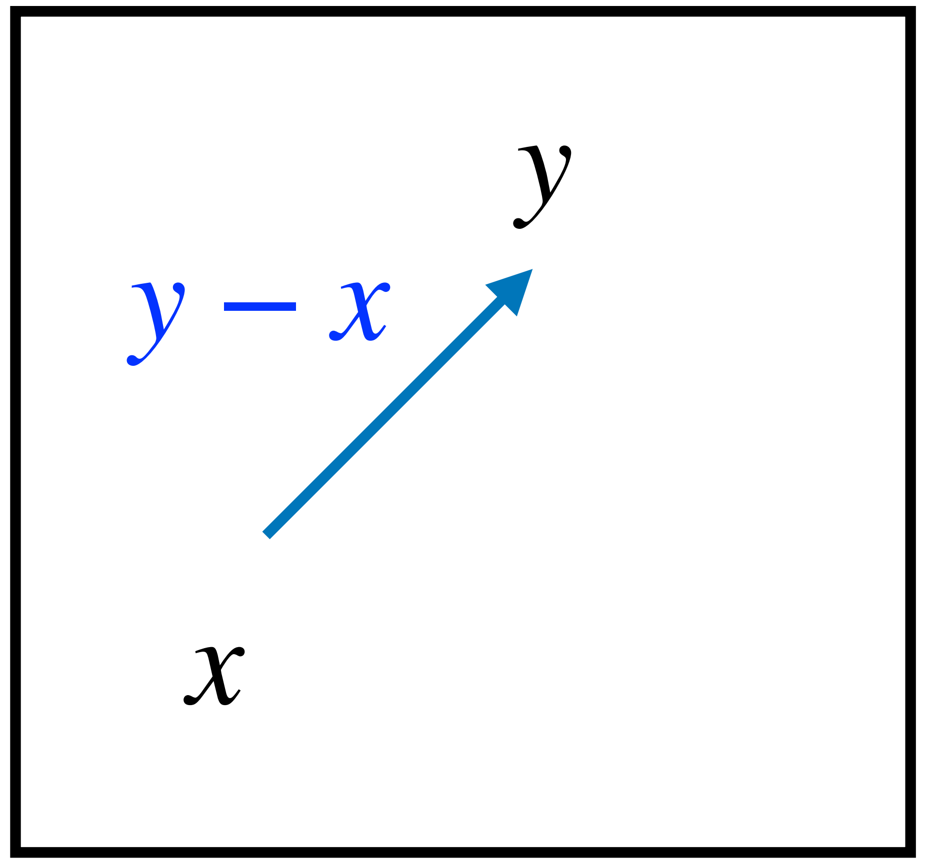
$$\mathrm{KL}(x, y) = \sum_{i \in [n]} \left(x_i \log \frac{x_i}{y_i} + y_i - x_i \right)$$

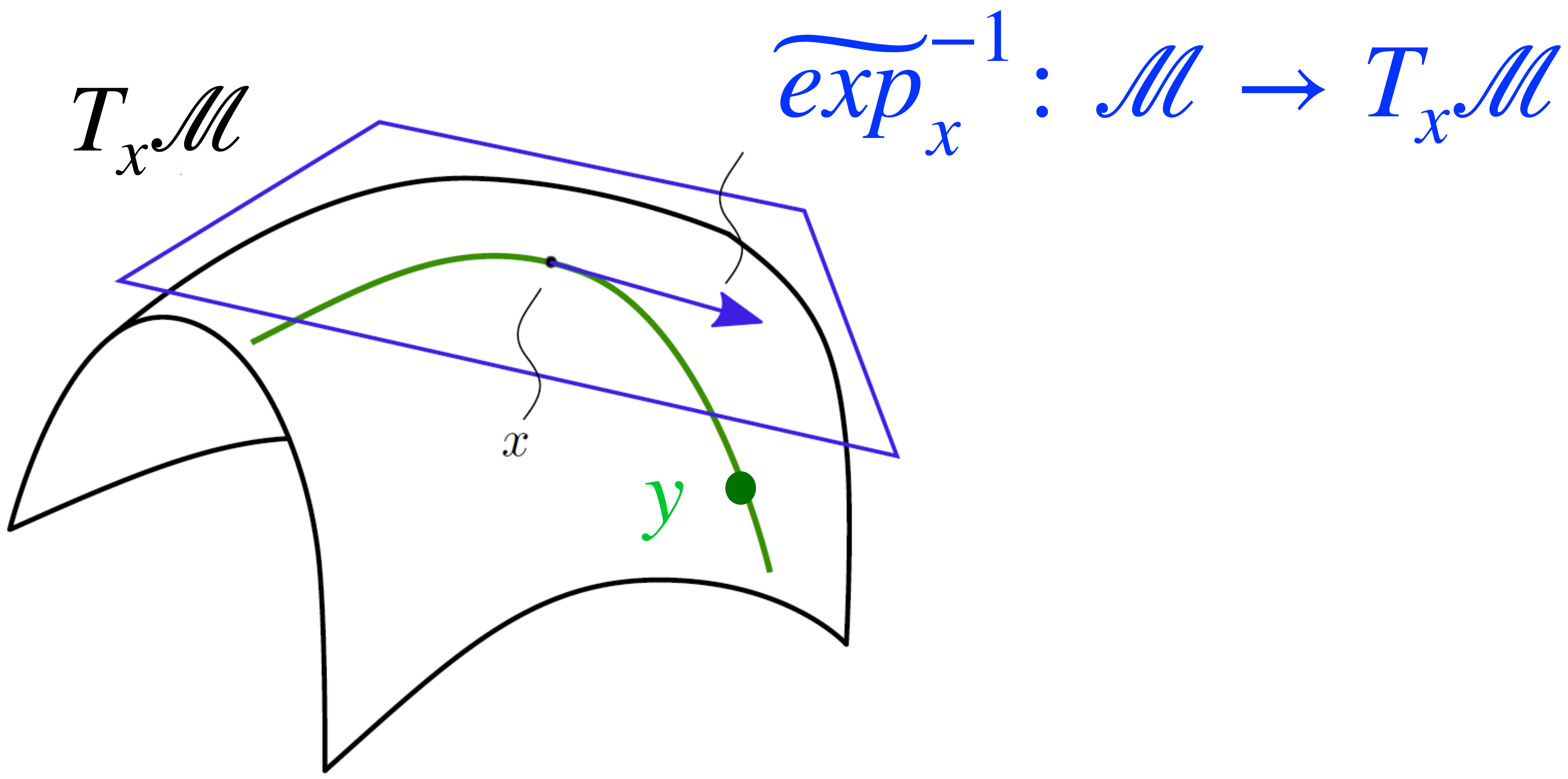
positive x : positive (non-normalised) discrete measure

Generalize Algorithmic Operations to Riemannian Manifolds





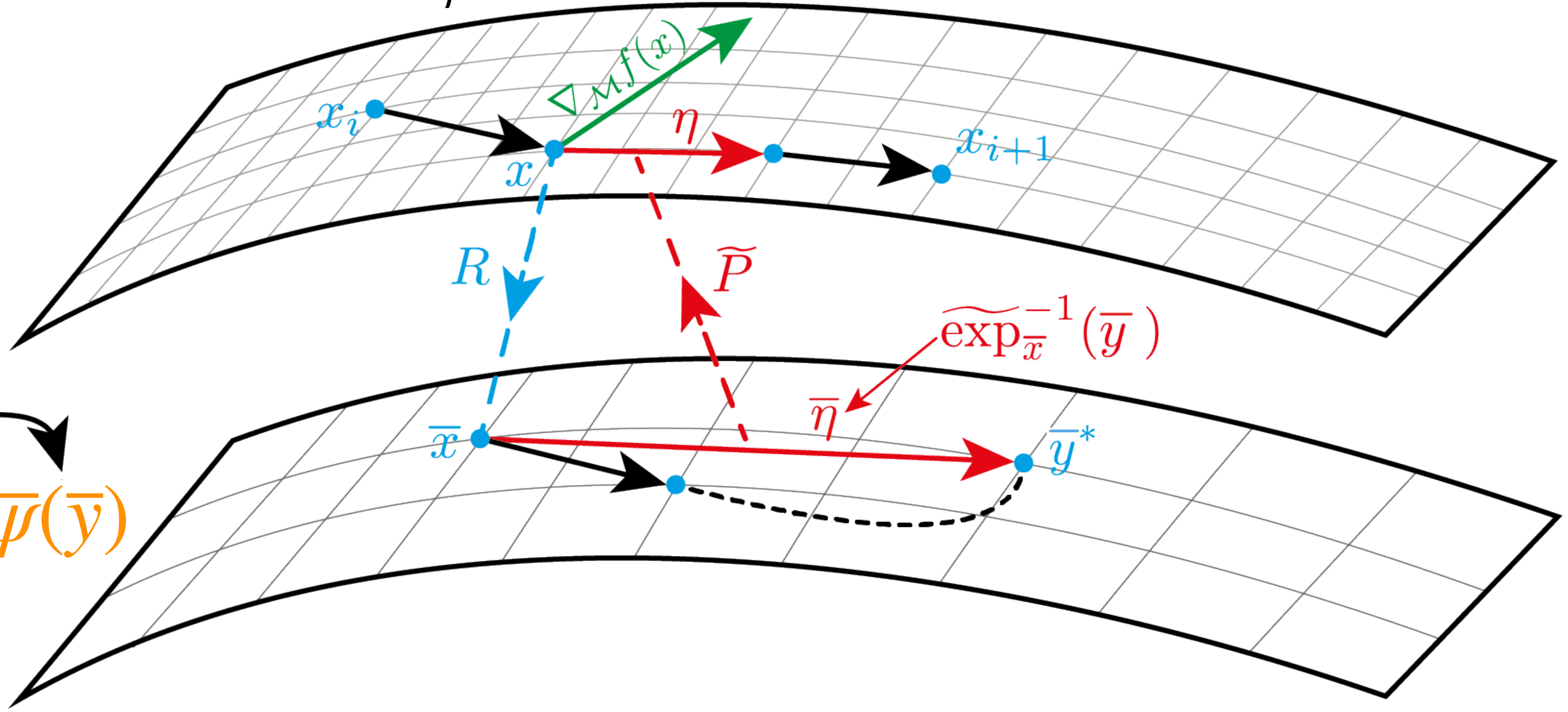




Geometric Multilevel Optimization

$$\min_{x \in \mathscr{M}} f(x) := \mathrm{KL}(Ax, b) + \lambda \|\nabla x\|_{1,\rho}$$

Plier, Savarino, Kocvara, P., *SSVM, 2021*



$$\min_{\overline{y} \in \overline{\mathscr{M}} \subset \mathbf{R}^{\overline{n}}} \overline{\psi}(\overline{y})$$

to into account **constraints** smoothly: $\mathscr{M} := (l, u) = (0,1)^n$

$\Longrightarrow$ change to a Riemannian **metric**

$\Longrightarrow$ devise a **retraction** for first-order optimization

Geometric Multilevel Optimization

potential: convex Legendre-type function

Alvarez, Bolte, Brahic, *SIAM J Control Optim, 2004*

$$\varphi(x) = \left\langle u - l, (x - l)\log(x - l) + (u - x)\log(u - x) \right\rangle$$

metric $\quad g_x(v, w) = \left\langle v, \nabla^2\varphi(x)\, w \right\rangle$

---

retraction

$$\mathcal{M} = ((l_i, u_i), h) \longleftrightarrow \underset{F_i^{-1}}{\overset{F_i}{\rightleftarrows}} \longleftrightarrow \mathcal{N} = (\dot{\Delta}_2, g_{FR})$$

Plier, Savarino, Kocvara, P., *SSVM, 2021*

pullback $\quad \widetilde{exp}_x^{\mathcal{M}} : T_x\mathcal{M} \to \mathcal{M}, \qquad \widetilde{exp}_x^{\mathcal{M}}(v)_i = F_i^{-1}\left(\widetilde{exp}_{F_i(x)}^{\mathcal{N}}\left(d_x F_i(v)\right)\right)$

information geometry / e-connection $\quad \mathrm{Exp}_x = \dfrac{x e^{\frac{v}{x}}}{\langle x, e^{\frac{v}{x}} \rangle}$

Plier, Savarino, Kocvara, P., *SSVM, 2021*

Geometry of the Box

metric

$$g_x(v, w) = \left\langle v, \nabla^2 \varphi(x)\, w \right\rangle =: \left\langle v, H_x w \right\rangle$$

$$\nabla_{\mathscr{M}} f(x) = H_x^{-1} \nabla f(x) = \frac{(x-l)(u-x)}{(u-l)^2} \nabla f(x)$$

Riemannian gradient

retraction

$$\widetilde{exp}_x^{\mathscr{M}} : T_x\mathscr{M} \to \mathscr{M}, \qquad \widetilde{\exp}_x^{\mathscr{M}}(v) = l + \frac{(u-l)(x-l)e^{\frac{u-l}{(x-l)(u-x)}v}}{(u-x) + (x-l)e^{\frac{u-l}{(x-l)(u-x)}v}}$$

Geometry of the Box

Plier, Savarino, Kocvara, P., *SSVM, 2021*

metric $\quad g_x(v,w) = \left\langle v, \nabla^2 \varphi(x)\, w \right\rangle =: \left\langle v, H_x w \right\rangle$

$$\nabla_{\mathscr{M}} f(x) = H_x^{-1} \nabla f(x) = \frac{(x-l)(u-x)}{(u-l)^2} \nabla f(x)$$

Riemannian gradient

retraction

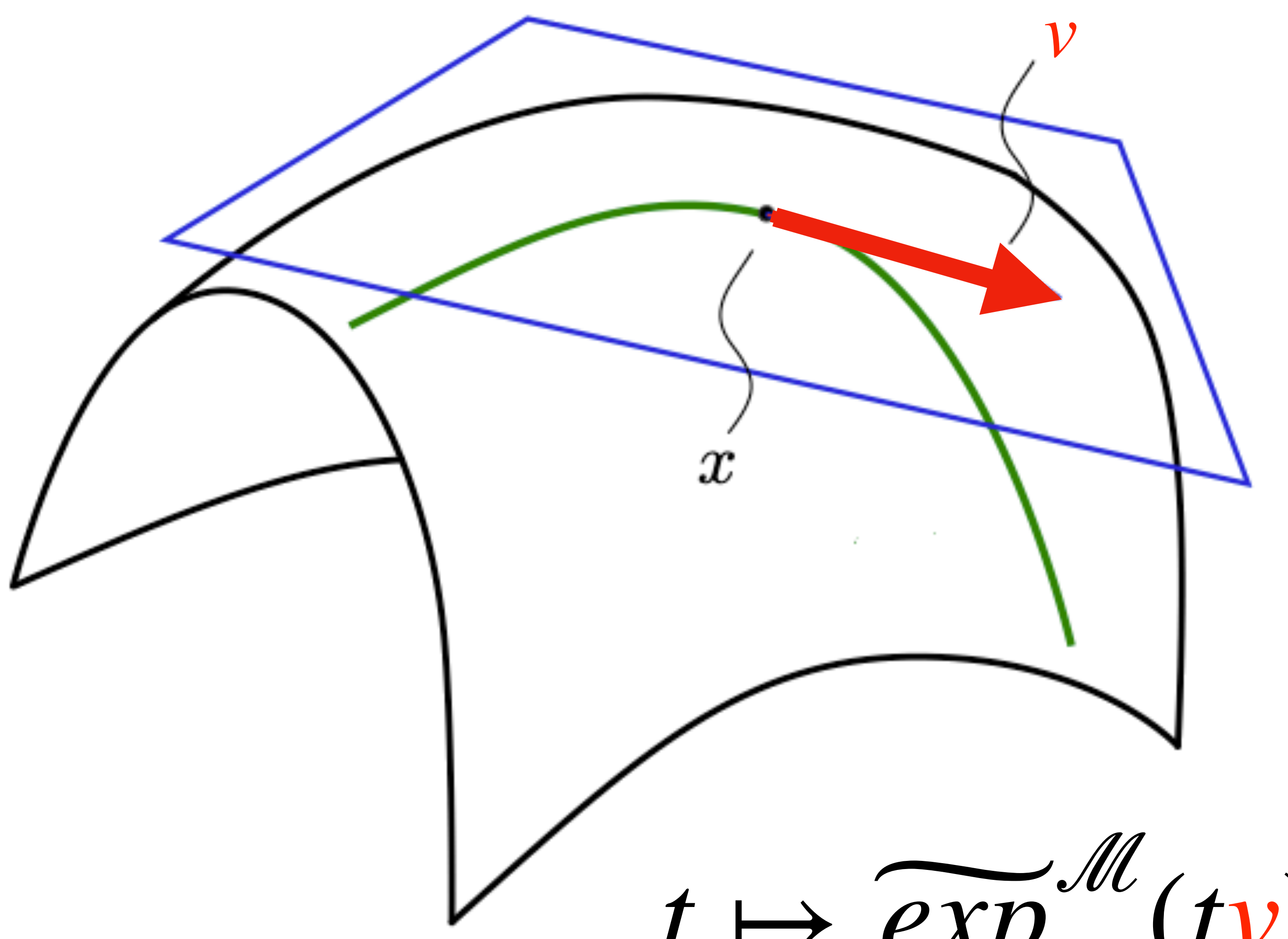
$$\widetilde{\exp}_x^{\mathscr{M}} : T_x\mathscr{M} \to \mathscr{M},$$

$$\Updownarrow$$

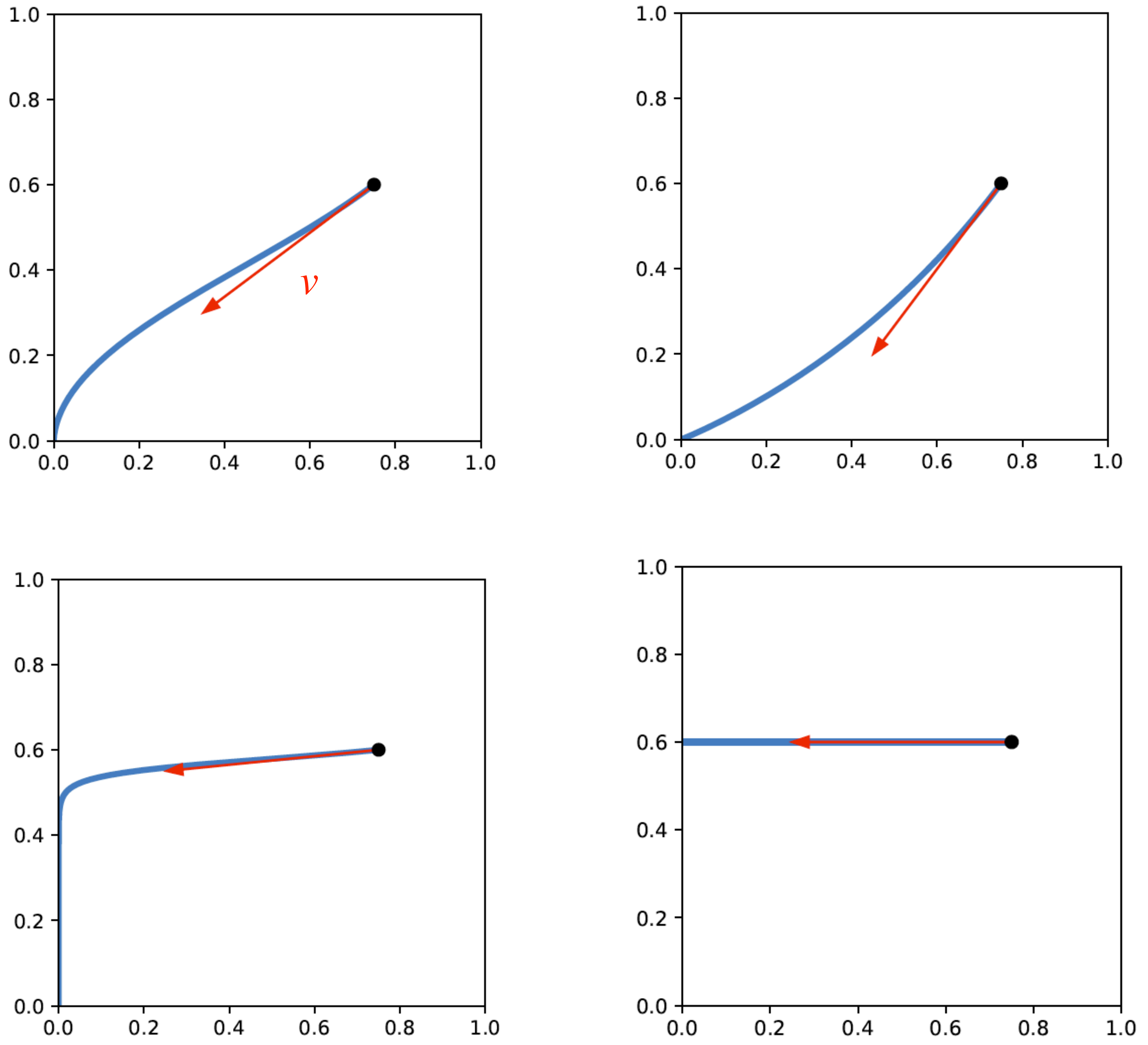
$$\widetilde{\exp}_x^{-1} : \mathscr{M} \to T_x\mathscr{M}$$

$$\widetilde{\exp}_x^{\mathscr{M}}(v) = l + \frac{(u-l)(x-l)e^{\frac{u-l}{(x-l)(u-x)}v}}{(u-x) + (x-l)e^{\frac{u-l}{(x-l)(u-x)}v}}$$

Retraction

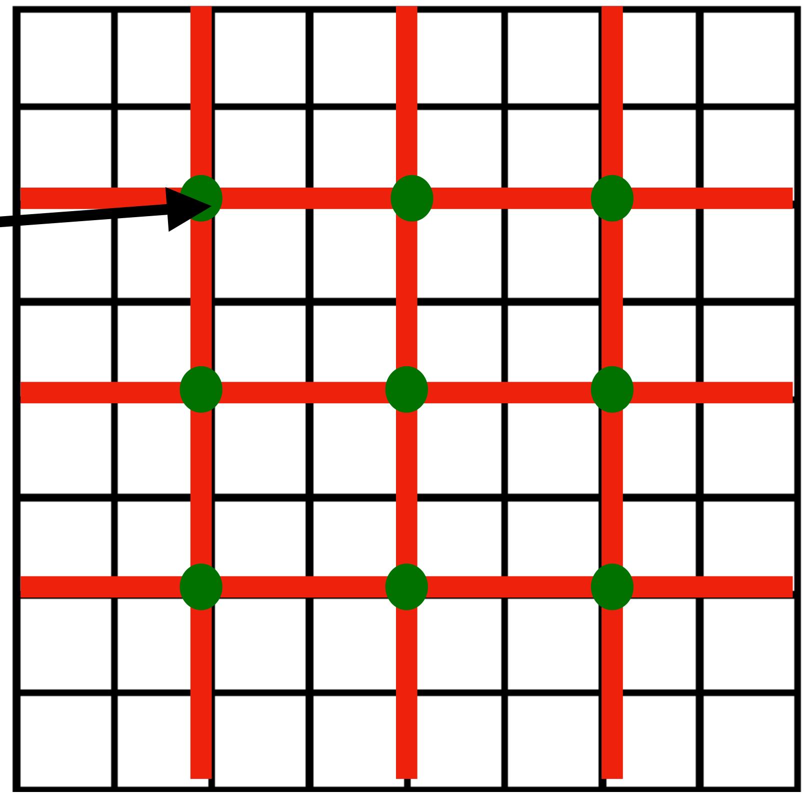


$$t \mapsto \widetilde{exp}_x^{\mathscr{M}}(tv)$$



Prolongation by Geometric Averaging

$$P(\bar{x})_i = x_i := \begin{cases} x_i, & \textbf{if } i \in I_{\overline{\mathscr{M}}}, \\ \mathrm{mean}_{\Omega_i}(\bar{x}_j)_{j \in N_i}, & \textbf{if } i \in I^c_{\overline{\mathscr{M}}} \end{cases}$$



With $\bar{x} = R(x)$ and $u \in T_{\bar{x}}\overline{\mathscr{M}}$ we have that $\mathrm{d}P_{\bar{x}}u \in T_x\mathscr{M}$

$$g_x(\mathrm{d}P_{\bar{x}}u, v) = g_{\bar{x}}(u, \mathrm{d}R_x v), \quad v \in T_x\mathscr{M}$$

Galerkin condition

$$R = P^\top$$

Two Grid Approach, Geometric Coarse Model

for *current* fine grid variable $x \in \mathscr{M}$ define coarse grid model

$$\overline{\psi}(\overline{y}) = \bar{f}(\overline{y}) - g_{\overline{x}}\left(\overline{v}_x, \widetilde{exp}_{\overline{x}}^{-1}(\overline{y})\right)$$

with $\overline{x} = R(x)$ and $\overline{v}_x = \nabla \bar{f}(\overline{x}) - \mathrm{d}R_x \nabla f(x)$

starting iterate at coarse level

Two Grid Approach, Coarse Model

coarse

no

do a fine iteration

if **condition** at x

yes

set $\bar{x} = Rx$, $\overline{\mathscr{M}} = (\bar{l}, \bar{u})$

find $\bar{y} \in \overline{\mathscr{M}}$ with $\bar{\phi}(\bar{y}) < \bar{\phi}(\bar{x})$

set $\bar{\eta} = \widetilde{exp}_{\bar{x}}^{-1}(\bar{y})$ and $\eta = \mathrm{d}P_{\bar{x}}\bar{\eta}$

find α with $f(\widetilde{exp}_x^{\mathscr{M}}(\alpha\eta)) < f(x)$

update $x^+ = \widetilde{exp}_x^{\mathscr{M}}(\alpha\eta)$

fine

Outline

Optimization: local approximation

Two-grid optimization

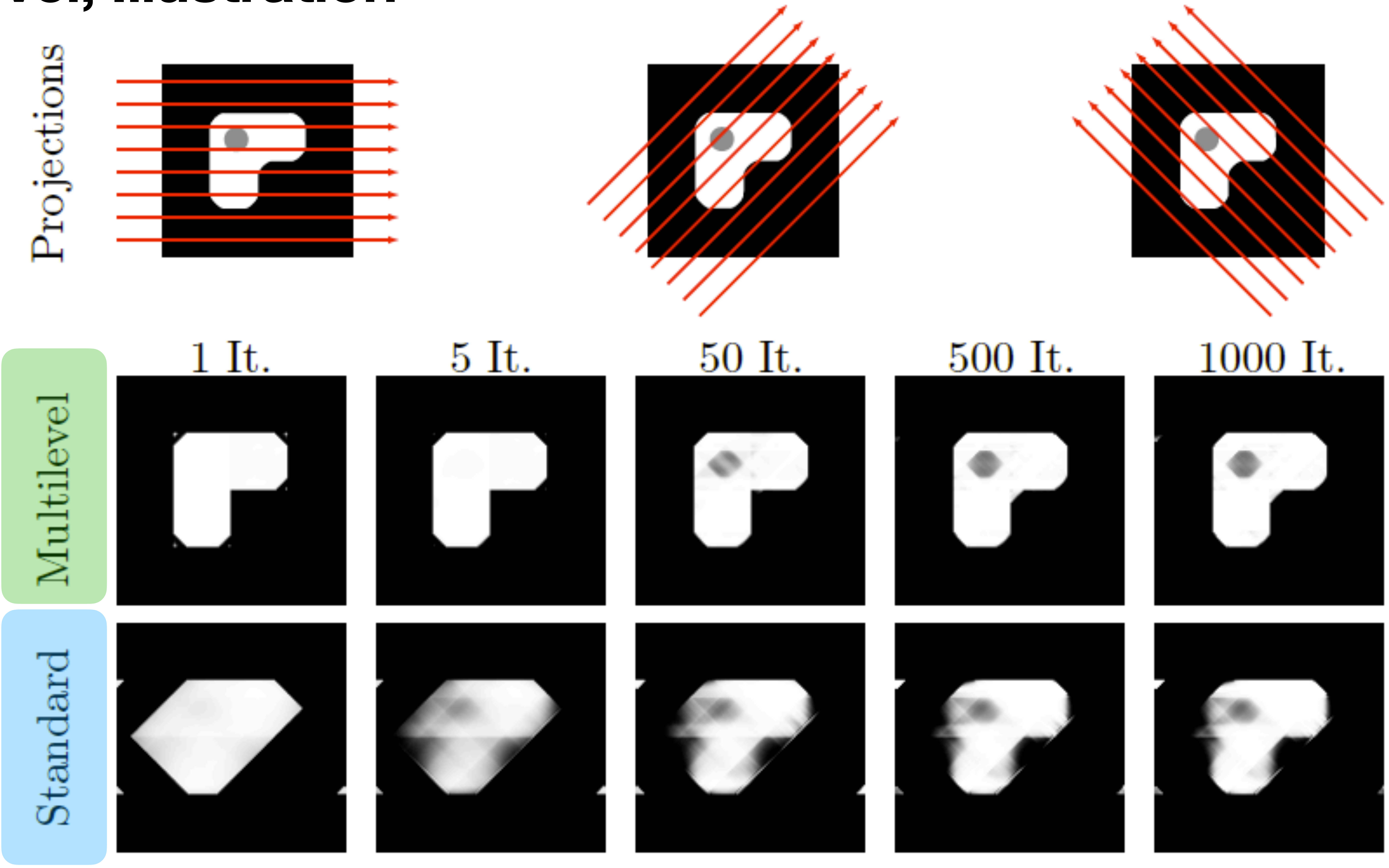
Connection to nonlinear multigrid (FAS)

Connection to standard local approximation

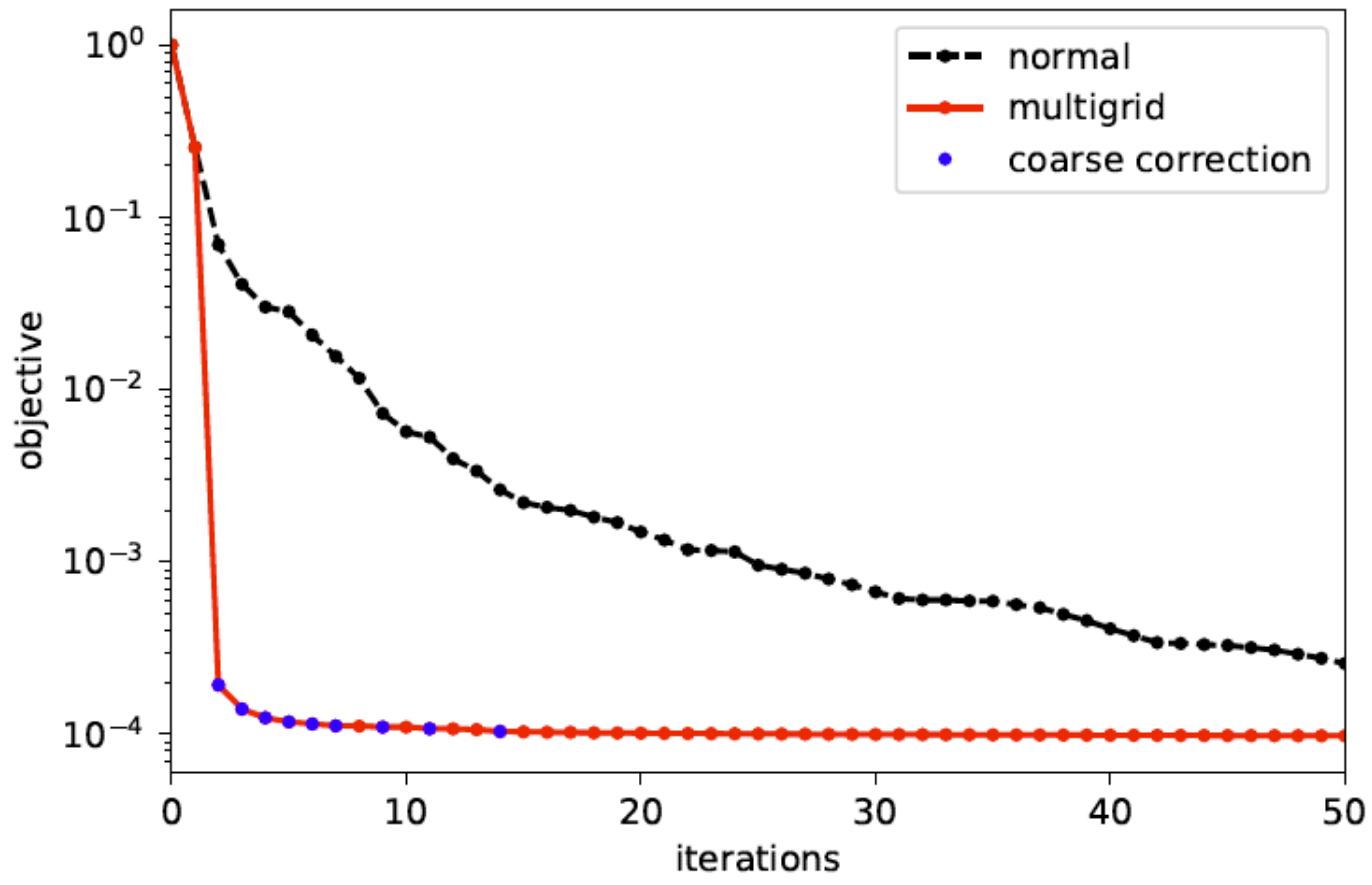
Geometric two-grid optimization

Application to discrete tomography

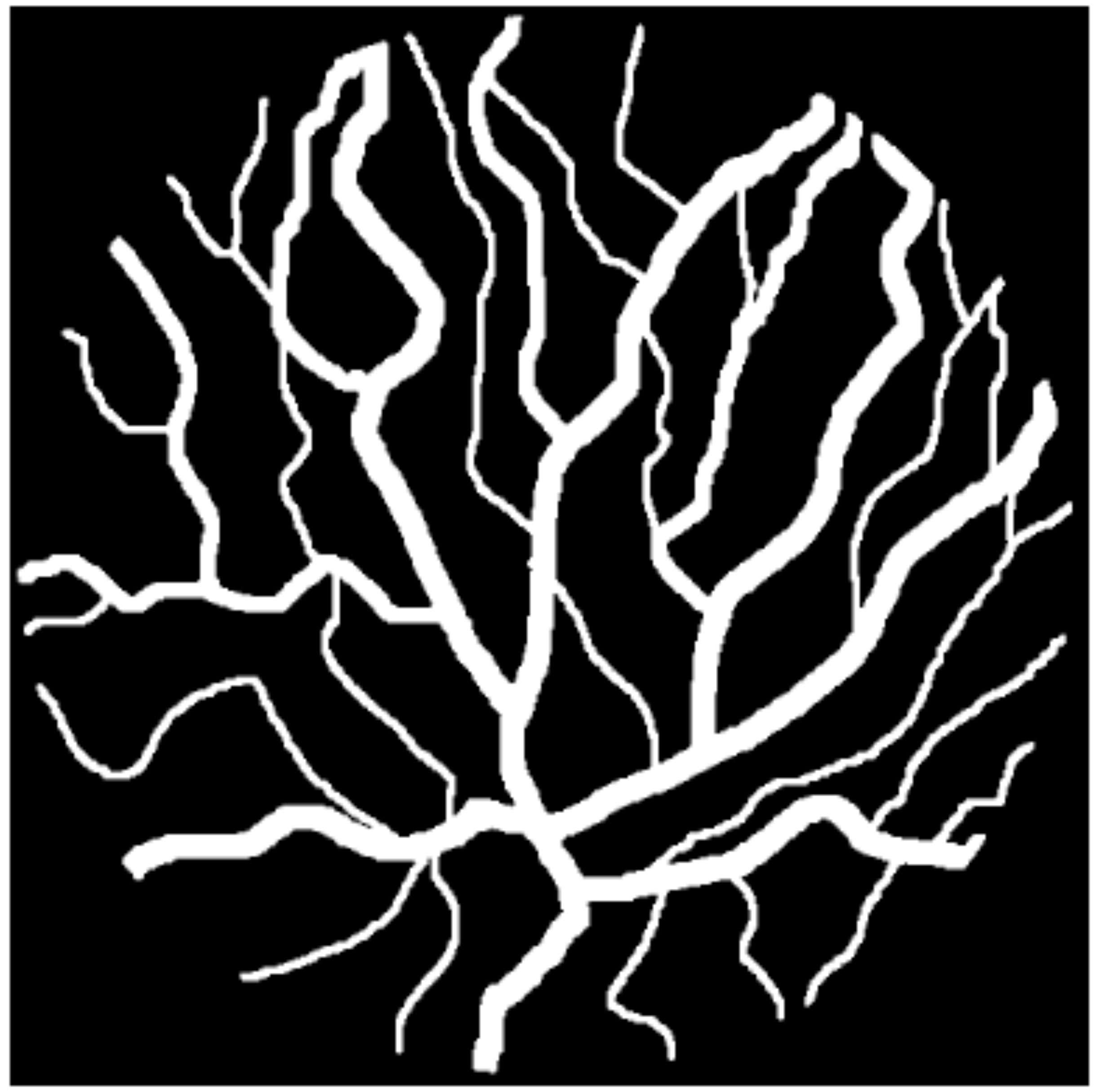
Multilevel, Illustration



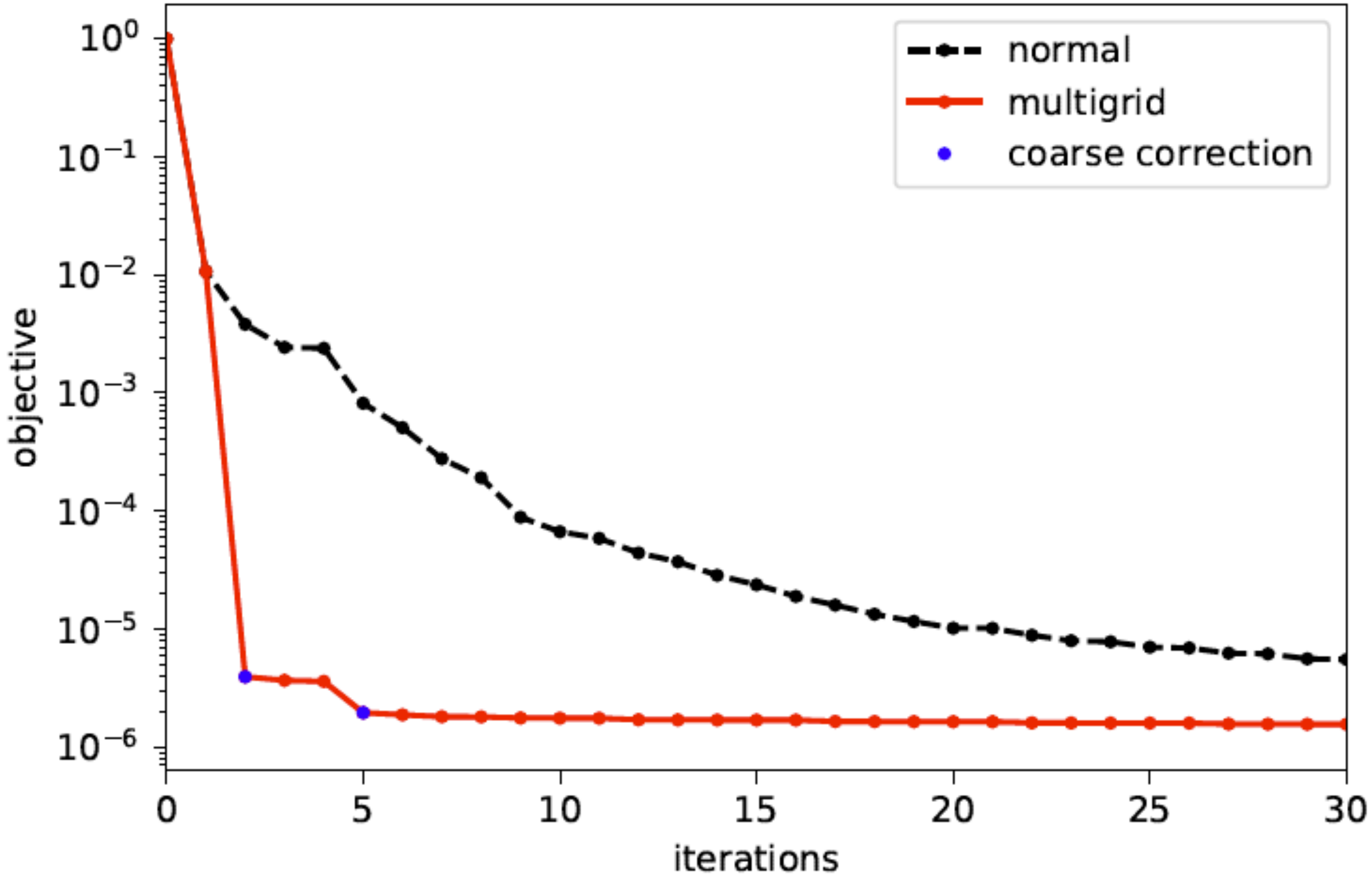
Multilevel, Numerical Results



1024 x 1024



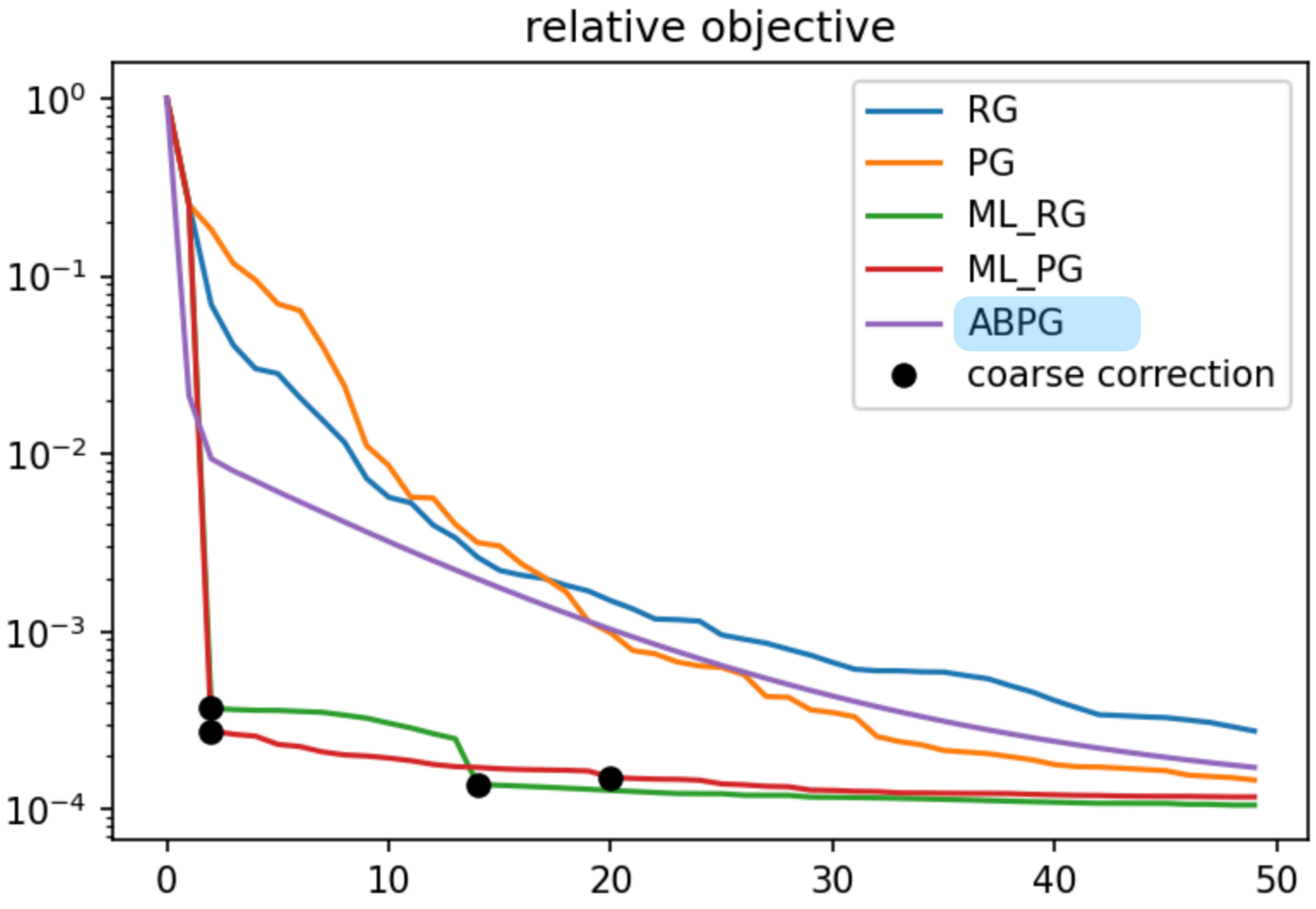
Multilevel, Numerical Results

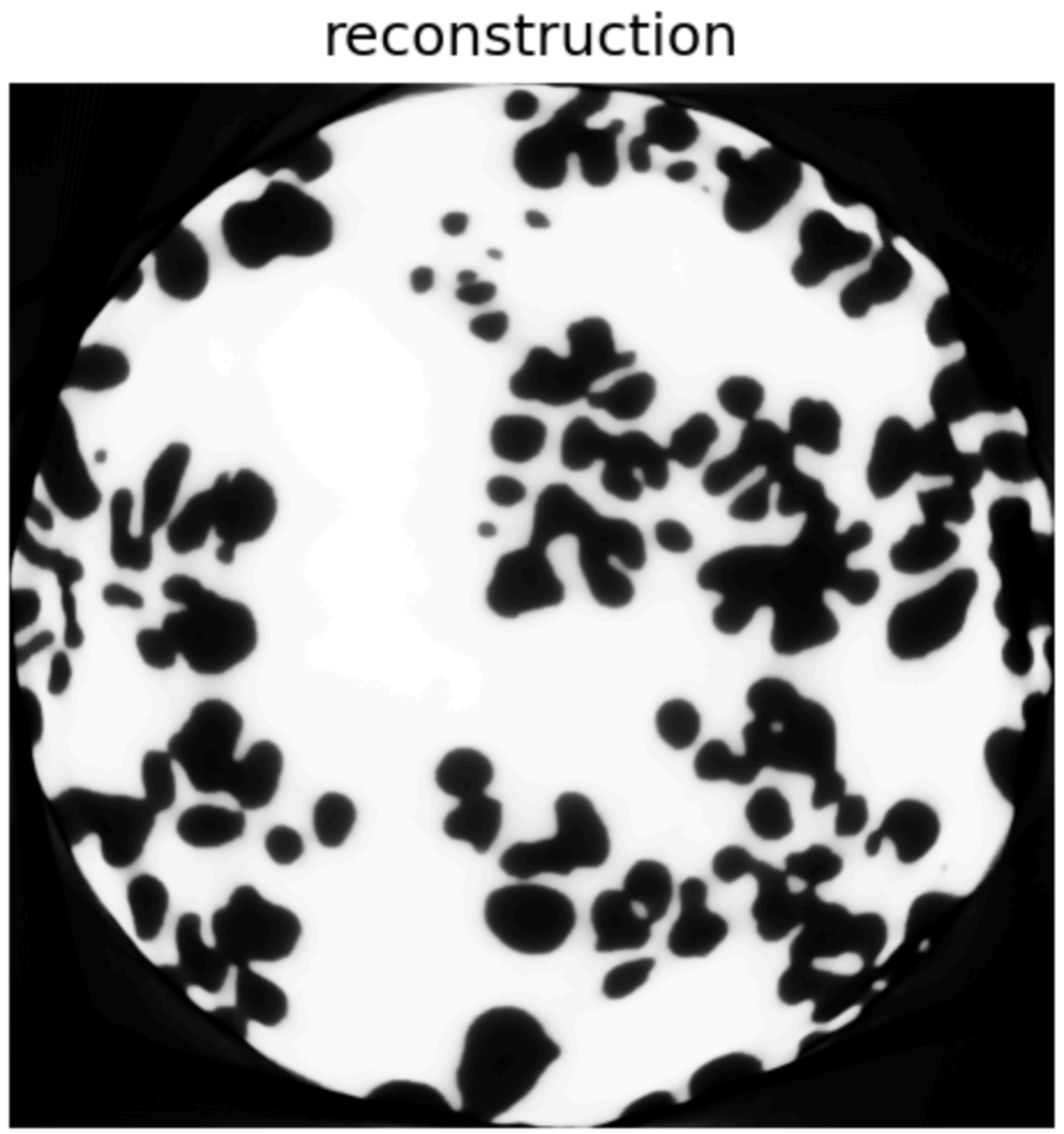


1024 x 1024

Multilevel vs FISTA-type Acceleration

Hanzely et al, *ArXiv, 2021*





1024 x 1024

Conclusion

- *Multilevel / multigrid optimization* approach
- *Coarse model:* efficient descent direction computation
- *Geometry* takes into account constraints
- *State-dependent* restriction and prolongation
- *Recursive procedure*: more levels can be used

- Convergence rates; no line search; geodesic convexity
- Coarse models for *non-convex* problems

References I

U. Trottenberg, C. Oosterlee, A. Schüller, **Multigrid,** Academic Press, 2001

S. G. Nash, **A multigrid approach to discretized optimization problems.** Optimization Methods and Software, 2000

S. Gratton, A. Sartenaer, P. L. Toint, **Recursive trust-region methods for multiscale nonlinear optimization.** SIAM Journal on Optimization, 2008

Z. Wen, D. Goldfarb, **A line search multigrid method for large-scale nonlinear optimization.** SIAM Journal on Optimization, 2009

S. Gratton, M. Mouffe, P. L. Toint, M. Weber-Mendonca, **A recursive formula-trust-region method for bound-constrained nonlinear optimization.** IMA Journal of Numerical Analysis, 2008

M. Kocvara and S. Mohammed, **A first-order multigrid method for bound-constrained convex optimization.** Optimization Methods and Software, 2016

P. Parpas, **A multilevel proximal gradient algorithm for a class of composite optimization problems.** SIAM J. Sci. Comp., 2017

References II

S. Roux, H. Leclerc, F. Hild, **Efficient Binary Tomographic Reconstruction,** J Math Imaging Vis, 2014

A. *Dabravolski, K. J. Batenburg, J. Sijbers,* **A Multiresolution Approach to Discrete Tomography Using DART,** PLoS One. 2014

J. Plier, **Theoretical and Numerical Approaches to Co-/Sparse Recovery in Discrete Tomography**, PhD Thesis, HeiDOK, 2020

J. Plier, F. Savarino, M. Kocvara, S. Petra, F**irst-Order Geometric Multilevel Optimization for Discrete Tomography**, SSVM, 2021

S. I. Amari, H. Nagaoka, **Methods of Information Geometry,** Amer. Math. Soc. and Oxford Univ. Press, 2000

P. A. Absil, R. Mahony, R. Sepulchre, **Optimization Algorithms on Matrix Manifolds,** Princeton University Press, 2008

Geometric Multilevel Optimization For Discrete Tomography

Stefania Petra
Mathematical Imaging Group
Heidelberg University

IN MEMORIAM OF CARLA PERI

TAIR 2022
Politecnico di Milano, May 2, 2022 – May 4, 2022