A Mini-course on Multigrid Method: Nonlinear and constrained problems

Hardik Kothari

Euler Institute,
Università della Svizzera italiana,
Switzerland

May 16, 2023
Research school on Iterative Methods for Partial Differential Equations
Paris, France

Outline

Overview

- Nonlinear problems
    - Problem in residual form
    - Problem in minimization form
    - Full approximation scheme/MG-opt
    - Required modifications
- Inequality constrained problem
    - Variational inequality
    - Monotone Multigrid
    - Required modifications

Table of Contents

From linear to nonlinear problem

- Find $\boldsymbol{x}^* \in \mathcal{V}$ such that
$$\boldsymbol{A}(\boldsymbol{x}^*) = \boldsymbol{b}$$
where $\boldsymbol{A} : \mathcal{V} \to \mathcal{V}'$ is nonlinear residual function
- How to utilize multigrid method to solve nonlinear problem
- Key relation employed in the multigrid method
$$\boldsymbol{A}\boldsymbol{x}^* - \boldsymbol{A}\boldsymbol{x}^{(k)} = \boldsymbol{f} - \boldsymbol{A}\boldsymbol{x}^{(k)} = \boldsymbol{r}^{(k)}$$
- This relation does not hold for nonlinear problems
$$\boldsymbol{A}(\boldsymbol{x}^*) - \boldsymbol{A}(\boldsymbol{x}^{(k)}) \neq \boldsymbol{A}(\boldsymbol{e}^{(k)})$$
- Nonlinear residual equation
$$\boldsymbol{A}(\boldsymbol{x}^*) - \boldsymbol{A}(\boldsymbol{x}^{(k)}) = \boldsymbol{r}^{(k)}$$

Nonlinear problem

- Find $\boldsymbol{x}^* \in \mathcal{V}$ such that

$$F(\boldsymbol{x}^*) = 0$$

  where $F : \mathcal{V} \to \mathcal{V}'$ is nonlinear residual function.

- To solve above problem, we construct sequence of iterates $\boldsymbol{x}^{(k)} \in \mathcal{V}$ such that as $\boldsymbol{x}^* = \lim_{k\to\infty} x^{(k)}$

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \boldsymbol{s}^{(k)} \qquad \text{for } k = 0, 1, 2, \ldots$$

- Newton's method:
  - Linearize the nonlinear residual

$$F(\boldsymbol{x}^{(k)} + \boldsymbol{s}^{(k)}) \approx F(\boldsymbol{x}^{(k)}) + \nabla F(\boldsymbol{x}^{(k)})\boldsymbol{s}^{(k)} = 0$$

  - A Jacobian matrix $\boldsymbol{J}(\boldsymbol{x}) = \nabla F(\boldsymbol{x})$
  - Newton correction

$$\boldsymbol{s}^{(k)} = -(\boldsymbol{J}(\boldsymbol{x}^{(k)}))^{-1} F(\boldsymbol{x}^{(k)})$$

  - Gobalized update

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha^k \boldsymbol{s}^{(k)}$$

Minimization problem

- Model problem:

$$\text{Find } \boldsymbol{x}^* = \underset{\boldsymbol{x}^* \in \mathcal{V}}{\arg\min} f(\boldsymbol{x}^*)$$

  where $f : \mathcal{V} \to \mathbb{R}$ is twice continuously differentiable objective function
- First order optimality condition

$$\nabla f(\boldsymbol{x}^*) = 0$$

- To solve above problem, we construct sequence of iterates $\boldsymbol{x}^{(k)} \in \mathcal{V}$ such that as $\boldsymbol{x}^* = \lim_{k\to\infty} x^{(k)}$

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \boldsymbol{s}^{(k)}$$

- Newton's method:
  - Linearize the nonlinear residual

$$\nabla f(\boldsymbol{x}^{(k)} + \boldsymbol{s}^{(k)}) \approx \nabla f(\boldsymbol{x}^{(k)}) + \nabla^2 f(\boldsymbol{x}^{(k)})\boldsymbol{s}^{(k)} = 0$$

  - A Hessian matrix $\boldsymbol{H}(\boldsymbol{x}) = \nabla^2 f(\boldsymbol{x})$
  - Newton correction

$$\boldsymbol{s}^{(k)} = -(\boldsymbol{H}(\boldsymbol{x}^{(k)}))^{-1} \nabla f(\boldsymbol{x}^{(k)})$$

  - Gobalized update

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha^k \boldsymbol{s}^{(k)}$$

Multilevel solution approaches

- Newton-multigrid method:
  - Global linearization: requires global Jacobian
  - Solve linear system with multigrid
  - Jacobians are not necessarily SPD
  - Can solve the linear system inexactly → Inexact Newton strategy
- Full approximation scheme (Brandt 1977)
  - Perform nonlinear relaxation on each level → local linearization
  - Solve nonlinear defect equation
  - Initial guess in the basin of attraction
- MG-Opt method (Nash 2001)
  - Similar to FAS but in the optimization framework
  - Optimization method employs line-search framework
  - Convex problem
- Recursive multilevel trust region method (Gratton et al. 2008)
  - Solve nonlinear problem on each level
  - Optimization method employs trust-region framework
  - Non-convex problems

MG-Opt method

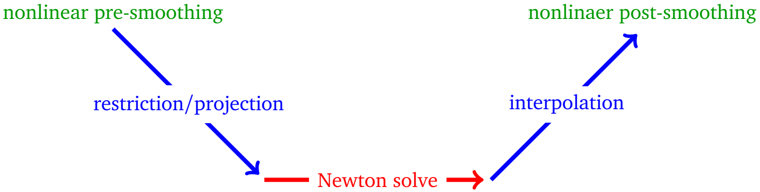
- Construct optimization problems on multiple levels

$$\arg\min_{u_\ell \in \mathcal{V}_\ell} f_\ell(u_\ell), \qquad \text{where } \ell = 1, \dots, L,$$

  where L denotes finest resolution, 1 denotes coarsest resolution
- Discretization of the objective function on different levels $\ell = 1, 2, \dots, L$ is necessary
- Nonlinear smoothing on each level
- Exact solution of the nonlinear system on coarsest level, e.g. Newton method
- In addition to standard restriction and projection operators a projection operator is required

Two-level FAS/MG-opt algorithm

Solve: $F(\boldsymbol{x}_h) = 0$ or $\nabla f(\boldsymbol{x}_h) = 0$



Coarse grid correction scheme:

1. Perform ν_{pre} pre-smoothing steps: $\boldsymbol{x}_h^{\nu_{pre}} \leftarrow \text{Nonlinear Smoother}(\boldsymbol{x}_h, \nu_{pre})$
2. Restrict residual to the coarse grid: $\boldsymbol{r}_H \leftarrow \boldsymbol{R}(\nabla f_h(\boldsymbol{x}_h^{\nu_{pre}}))$
3. Solve nonlinear problem on coarse space: $\nabla f_H(\boldsymbol{x}_H) + \boldsymbol{r}_H - \nabla f_H(\boldsymbol{P}(\boldsymbol{x}_h^{\nu_{pre}})) = 0$
4. Interpolate correction to the fine grid: $\boldsymbol{x}_h \leftarrow \boldsymbol{x}_h + \alpha \boldsymbol{I}(\boldsymbol{x}_H - \boldsymbol{P}(\boldsymbol{x}_h^{\nu_{pre}}))$
5. Perform ν_{post} post-smoothing steps: $\boldsymbol{x}_h^{\nu_{post}} \leftarrow \text{Nonlinear Smoother}(\boldsymbol{x}_h, \nu_{post})$

Transfer operators

- Prolongation operator (Interpolation operator):

$$\boldsymbol{I}_{\ell}^{\ell+1} : \mathcal{V}_{\ell} \to \mathcal{V}_{\ell+1}$$

- Restriction operator: $\boldsymbol{R}_{\ell+1}^{\ell} = (\boldsymbol{I}_{\ell+1}^{\ell})^{\top}$

$$\boldsymbol{R}_{\ell+1}^{\ell} : \mathcal{V}_{\ell+1}' \to \mathcal{V}_{\ell}'$$

- Projection operator:

$$\boldsymbol{P}_{\ell+1}^{\ell} : \mathcal{V}_{\ell+1} \to \mathcal{V}_{\ell}$$

- Property of interpolation and projection operator

$$x_{\ell+1} = \boldsymbol{I}_{\ell}^{\ell+1} \boldsymbol{P}_{\ell+1}^{\ell} \boldsymbol{x}_{\ell+1}$$

Nonlinear smoothers

Full approximation Scheme:

- Nonlinear Richardson
- Nonlinear Jacobi
- Nonlinear Gauss-Seidel method
- Pointwise linearization

MG-opt method:

- First order methods
- Gradient descent (nonlinear Richardson)
- Nonlinear Conjugate gradient
- Block coordinate descent
- Line-search at each iteration

First order consistency term

- Also known as: τ-correction, defect equation
- Original problem on coarse level

$$\nabla f_H(\boldsymbol{x}_H) = 0$$

- We need to construct error equation on coarse level $A(x) - A(x^k) = r$

$$\nabla f_H(\boldsymbol{x}_H) - \nabla f_H(\boldsymbol{P}(\boldsymbol{x}_h^{\nu_{pre}})) = -\boldsymbol{R}\nabla f_h(\boldsymbol{x}_h^{\nu_{pre}})$$

- Reformulating

$$\nabla f_H(\boldsymbol{x}_H) + \boldsymbol{R}\nabla f_h(\boldsymbol{x}_h^{\nu_{pre}}) - \nabla f_H(\boldsymbol{P}(\boldsymbol{x}_h^{\nu_{pre}})) = 0$$

- Correction

$$\boldsymbol{x}_h \leftarrow \boldsymbol{x}_h + \alpha \boldsymbol{I}(\boldsymbol{x}_H - \boldsymbol{P}(\boldsymbol{x}_h^{\nu_{pre}}))$$

Numerical results

Bratu equation:

$$-\nabla^2 u + \gamma e^u = 0 \quad \text{in } \Omega$$
$$u = 0 \quad \text{on } \partial\Omega$$

- convergence of MG-Opt method with V(3,3) cycle
- Meshes from 20×20 to 160×160
- nonlinear CG method as a smother

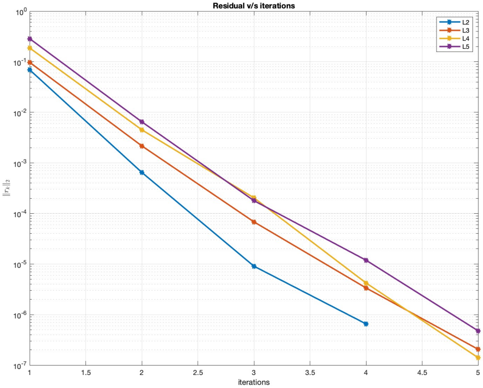


Table of Contents

Variational inequalities[1]

- Arises while handling inequality constraints
- such problems occur in many scientific fields, e.g., mechanics, physics, finance, etc.
- Weak formulation of the inequality constrained PDEs give rise to variational inequalities
- Examples of such problems can be given as:
  - obstacle problems
  - contact problems
  - porous media flow with minimal required-entry pressure
  - Stefan problem

[1] D. Kinderlehrer & G. Stampacchia, An Introduction to Variational Inequalities and Their Applications (1980)

Obstacle problem

- Find the equilibrium configuration of an elastic membrane with a fixed boundary, which is constrained to lie below a given obstacle
- Strong formulation:

$$\begin{aligned} -\nabla^2 u &\leqslant f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \\ u &\leqslant \overline{\psi} && \text{a.e. in } \Omega \end{aligned}$$

- Weak formulation:

$$\text{Find } u \in \mathcal{K} \text{ such that } \int_\Omega \nabla u \cdot \nabla(v-u)\, dx \geqslant \int_\Omega f(v-u)\, dx \quad \forall v \in \mathcal{K}$$

where $\mathcal{K} = \{v \in H_0^1(\Omega) \mid v \leqslant \overline{\psi}\}$

Reformulating obstacle problem

- Variational inequality: Find $\boldsymbol{u}_h \in \boldsymbol{\mathcal{K}}_h$ such that

$$a_j(\boldsymbol{u}_h, \boldsymbol{v}_h - \boldsymbol{u}_h) \geqslant F(\boldsymbol{v}_h - \boldsymbol{u}_h), \qquad \forall \boldsymbol{v} \in \boldsymbol{\mathcal{K}}_h$$

  where, $\mathcal{K}_h = \{v_h \in \mathcal{V}_h \mid v_h \leqslant \overline{\psi}_h\}$

- Constrained minimization problem: Find $\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h$ such that

$$\min_{\boldsymbol{u}_h \in \boldsymbol{\mathcal{V}}_h} \mathcal{J}(\boldsymbol{u}_h) = \frac{1}{2} a_j(\boldsymbol{u}_h, \boldsymbol{u}_h) - F(\boldsymbol{u}_h)$$
$$\text{subject to} \quad \boldsymbol{u}_h \leqslant \overline{\psi}_h$$

- Algebraic formulation:

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} J(\boldsymbol{x}) = \frac{1}{2} \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} - \boldsymbol{x}^\top \boldsymbol{b}$$
$$\text{subject to} \quad \boldsymbol{x} \leqslant \boldsymbol{g}$$

Methods for solving inequality constrained problems

$$\min_{\boldsymbol{x}} J(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}^\top \boldsymbol{b}$$
$$\text{subject to} \quad \boldsymbol{x} \leqslant \boldsymbol{g}$$

- Semismooth Newton method
- Interior point methods
- Primal-dual Active-set methods
- Monotone multigrid method

Monotone multigrid method[2]

- Standard solution methods requires linearization of constraints
- Monotone multigrid method circumvents the linearization of constraints
- It can solve the variational inequality problems with multigrid complexity
- The method preserves the monotonicity of the solution during the iterative process
- Employs the projected Gauss-Seidel method as a smoother on the finest level
  - Monotonically decreasing energy
  - Identifies the current active-set
  - The current iterate remains in the feasible set $\mathcal{K}$
- Coarse spaces are constructed using the truncated basis on the current active set
  - Achieved by treating the current active set as Dirichlet nodes
  - Modify the interpolation and restriction operators accordingly
- Ensures that the corrections from the coarse level do not modify the active set on the finest level

[2] R. Kornhuber, A Monotone multigrid methods for elliptic variational inequalities I & II, *Numerische Mathematik* (1994, 1996)

Projected Gauss-Seidel method

- Extension of the standard Guass-Seidel method
- Can solve constrained minimization problem
- Monotonically minimizes the energy function $\mathcal{J}$
- Retains the smoothing property of the Gauss-Seidel method
- Can be interpreted as a coordinate descent method

Projected Gauss-Seidel: Algorithm

Input : $\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{x}^{(0)}, \boldsymbol{lb}, \boldsymbol{ub}, \nu_*$
Output: $\boldsymbol{x}^{(\nu_*)}, \mathscr{A}$

Function: Projected GS$(\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{x}^{(0)}, \boldsymbol{lb}, \boldsymbol{ub}, \nu_*)$
for $k = 1, 2, \ldots, \nu_*$ **do**
  $\mathscr{A} \leftarrow \emptyset$; ▷ initialize empty active set
  for $i = 1, 2, \ldots, n$ **do**
    $\boldsymbol{x}_i^{(k)} = \frac{1}{\boldsymbol{A}_{ii}}(\boldsymbol{b}_i - \sum_{j<i} \boldsymbol{A}_{ij}\boldsymbol{x}_j^{(k)} - \sum_{j>i} \boldsymbol{A}_{ij}\boldsymbol{x}_j^{(k-1)})$; ▷ update the iterate
    if $\boldsymbol{x}_i^{(k)} < lb_i$ ***or*** $ub_i < \boldsymbol{x}_i^{(k)}$ **then**
      $\boldsymbol{x}_i^{(k)} = \max(lb_i, \min(\boldsymbol{x}_i^{(k)}, ub_i))$; ▷ project onto feasible set
      $\mathscr{A} \leftarrow \mathscr{A} \cup \{i\}$; ▷ add current index to the active set
    end
  end
end

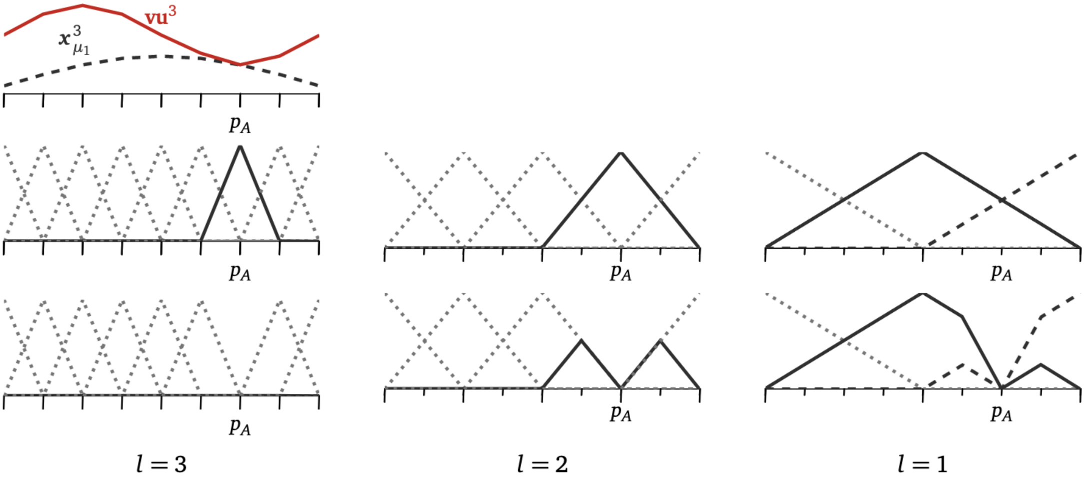
Algorithm 1: Projected Gauss-Seidel method

Truncated Transfer operators

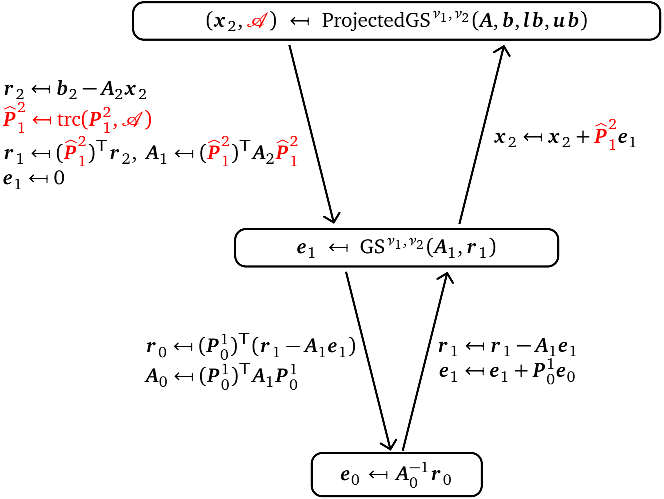
- The transfer operators are modified
- The activeset can be interpreted as Dirichlet nodes
- Eliminate the contribution at active-nodes from residuals and operators
- Can be achieved by truncating prolongation operator

$$\boldsymbol{P} = \begin{bmatrix} 1 & & & & & \\ 1/2 & 1/2 & & & & \\ & 1 & & & & \\ & 1/2 & 1/2 & & & \\ & & 1 & & & \\ & & 1/2 & 1/2 & & \\ & & & \ddots & & \\ & & & \ddots & 1/2 \\ & & & & 1 \end{bmatrix} \qquad \widehat{\boldsymbol{P}} = \begin{bmatrix} 1 & & & & & \\ 1/2 & 1/2 & & & & \\ & 1 & & & & \\ & {\color{red}0} & {\color{red}0} & & & \\ & & {\color{red}0} & & & \\ & & 1/2 & 1/2 & & \\ & & & \ddots & & \\ & & & \ddots & 1/2 \\ & & & & 1 \end{bmatrix}$$

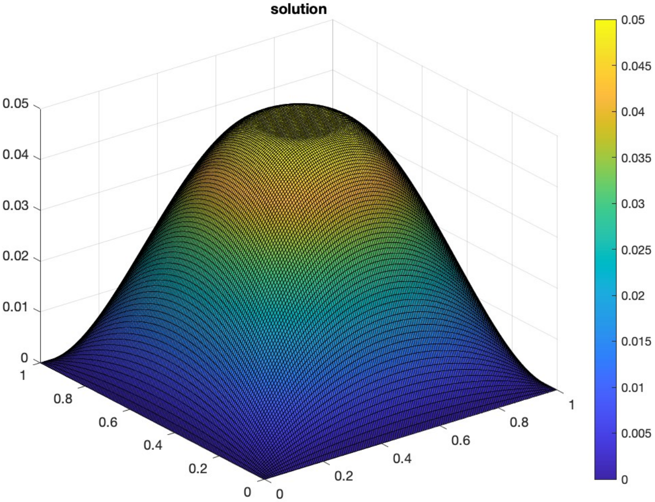
Truncated Basis



Monotone Multigrid algorithm



Numerical experiments



Comparing various MG-cycles

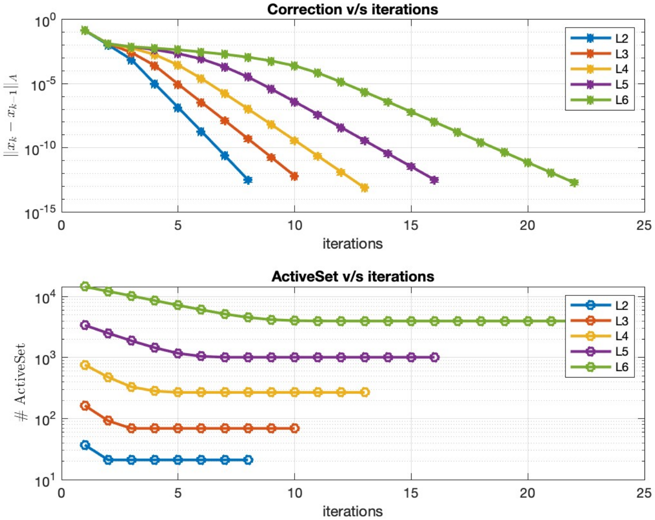
mesh size	# levels	V(3,3)	V(5,5)	W(3,3)	W(5,5)
20×20	2	8	8	8	8
40×40	3	10	9	9	9
80×80	4	13	11	10	9
160×160	5	16	14	12	11
320×320	6	22	18	16	14

Iterations

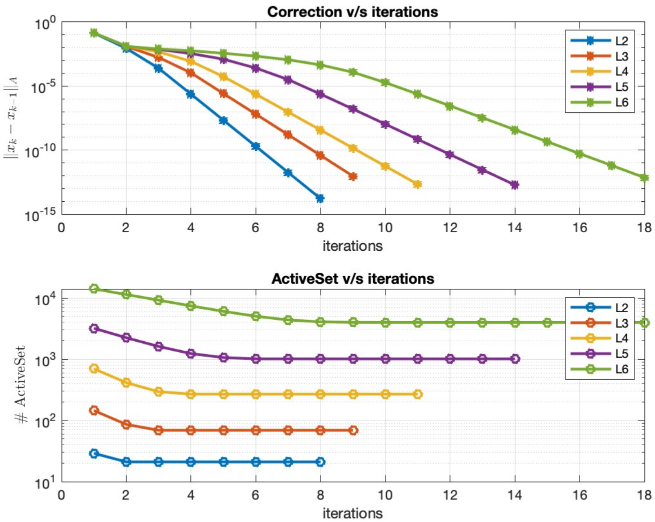
mesh size	# levels	V(3,3)	V(5,5)	W(3,3)	W(5,5)
20×20	2	0.0124	0.0091	0.0124	0.0091
40×40	3	0.0350	0.0241	0.0180	0.0108
80×80	4	0.0573	0.0383	0.0181	0.0109
160×160	5	0.0956	0.0648	0.0153	0.0101
320×320	6	0.1619	0.1165	0.0145	0.0093

Asymptotic convergence rate ($\rho^* = \dfrac{\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\|_A}{\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\|_A}$)

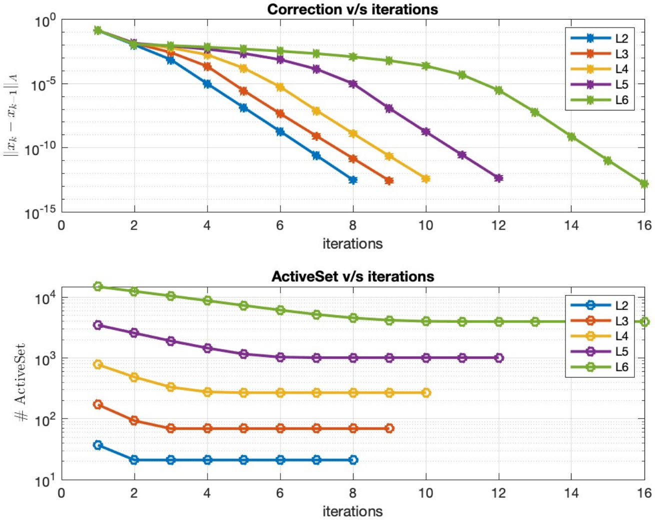
Convergence of V(3,3) cycle



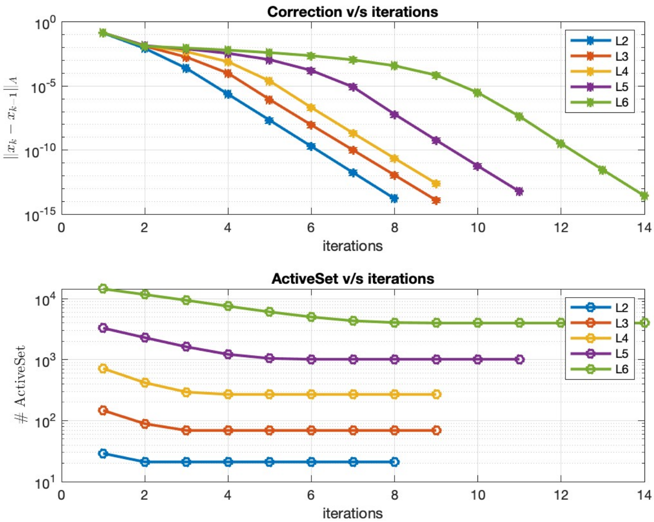
Convergence of V(5,5) cycle



Convergence of W(3,3) cycle



Convergence of W(5,5) cycle



Questions?

Contact: hardik.kothari@usi.ch