Preconditioning for Bound Constrained Quadratic Programming Problems Arising from Discretization of Variational Inequalities

Marta Domorádová and Zdeněk Dostál

Department of Applied Mathematics
VŠB-Technical University of Ostrava

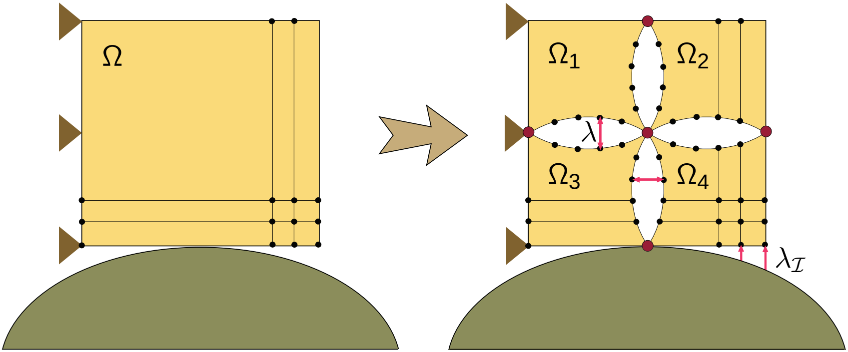
Supported by GA CR 201/07/0294, AS CR 1ET400300415
and MSM6198910027.

Computational Methods with Applications
August 19-25, 2007, HARRACHOV

Outline

Motivation

FETI-DP Domain Decomposition



To improve rate of convergence for variational inequalities (FETI-DP, BETI-DP, . . .).

Partially Bound Constrained QP Problem

Find

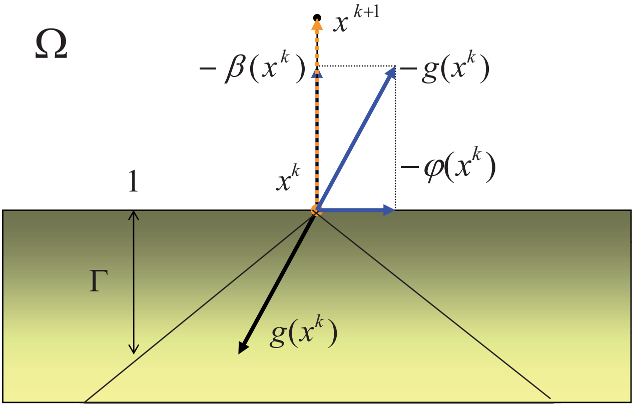
$$\min_{\mathbf{x}\in\Omega} f(\mathbf{x}), \quad \Omega = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{x}_{\mathcal{I}} \geq \ell_{\mathcal{I}}\}, \quad \mathcal{I} = \{1, \ldots, m\},$$

where

- $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{x}^T\mathbf{b}$,
- ℓ and **b** are given column *n*-vectors,
- $1 \leq m \ll n$,
- and **A** is an $n \times n$ spd matrix.

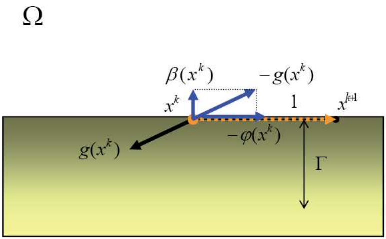
MPRGP - Proportioning

$\mathbf{x}^k$ strictly proportional $||\beta(\mathbf{x}^k)||^2 \leq \Gamma^2 \widetilde{\varphi}(\mathbf{x}^k)^T \varphi(\mathbf{x}^k)$

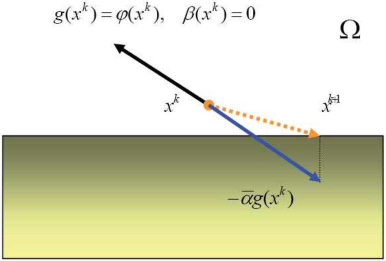


MPRGP - Proportional Iterations

Feasible CG step



Projection step:
expansion of the active set



Modified Proportioning with Reduced Gradient Projections - MPRGP

if $\mathbf{x}^k$ is *strictly proportional*

{trial conjugate gradient step}

if $\mathbf{x}^{k+1}$ is *feasible*

accept $\mathbf{x}^{k+1}$

else

{expansion step}

$x^{k+1} = P_\Omega(x^k - \overline{\alpha}\varphi(x^k))$

end

else

{proportioning step}

minimization in direction $-\beta(x^k)$

end

MPRGP Algorithm

Rate of Convergence

THEOREM

$\lambda_{\min}$ = min eigenvalue of **A**, $\Gamma > 0$, $\hat{\Gamma} = \max\{\Gamma, \Gamma^{-1}\}$

R-linear rate of convergence in energy norm

$$\|\mathbf{x}^k - \widehat{\mathbf{x}}\|_{\mathbf{A}}^2 \le 2\eta \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}}) \right),$$

where

$$\eta = 1 - \frac{\overline{\alpha}\lambda_{\min}}{2 + 2\widehat{\Gamma}^2}.$$

Dostál and Schöberl, Computational Optimization and Applications, 2005

Conjugate Projectors

P is an **A**-*conjugate projector*

if $\mathrm{Im}\mathbf{P} \perp_{\mathbf{A}} \mathrm{Ker}\mathbf{P} \iff \mathbf{P}^T\mathbf{A}(\mathbf{I}-\mathbf{P}) = \mathbf{P}^T\mathbf{A} - \mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{O}$

$\Downarrow$

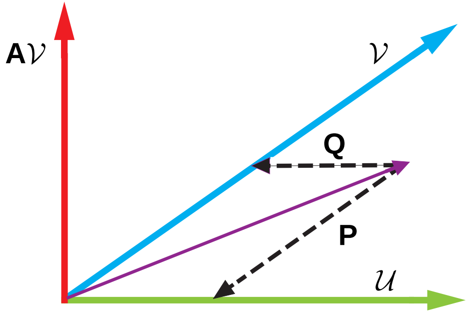
- $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ is also a conjugate projector
- $\mathbf{P}^T\mathbf{A} = \mathbf{A}\mathbf{P} = \mathbf{P}^T\mathbf{A}\mathbf{P}$ and $\mathbf{Q}^T\mathbf{A} = \mathbf{A}\mathbf{Q} = \mathbf{Q}^T\mathbf{A}\mathbf{Q}$.

Conjugate Projectors

$\mathcal{U}$ = subspace spanned by columns of full rank matrix $\mathbf{U} \in \mathbb{R}^{n \times p}$

$$\mathbf{U} = \begin{bmatrix} \mathbf{O} \\ \mathbf{U}_2 \end{bmatrix}, \quad \mathbf{U}_2 \in \mathbb{R}^{n-m \times p}$$

- $\mathbf{P} = \mathbf{U}(\mathbf{U}^T\mathbf{A}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{A}$ onto $\mathcal{U}$
- $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ onto $\mathcal{V}$



Invariant Subspace

$\mathcal{V} = \mathrm{Im}\mathbf{Q}$. For all $\mathbf{x} \in \mathbf{A}\mathcal{V}$

$$\mathbf{Q}^T\mathbf{A}\mathbf{Q}\mathbf{x} = \mathbf{A}\mathbf{Q}\mathbf{x}$$

$$\Downarrow$$

$$\mathbf{Q}^T\mathbf{A}\mathbf{Q}(\mathbf{A}\mathcal{V}) \subseteq \mathbf{A}\mathcal{V}.$$

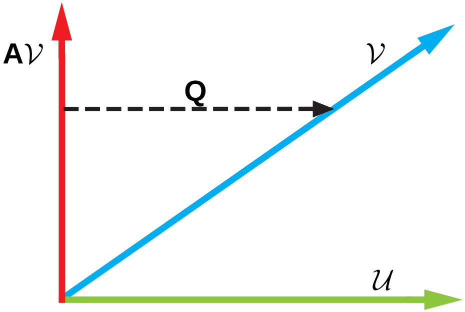
$\Rightarrow$ $\mathbf{A}\mathcal{V}$ = *invariant subspace* of $\mathbf{Q}^T\mathbf{A}\mathbf{Q}$.

Invariant Subspace

$\mathbf{Q}$ = conjugate projector on $\mathcal{V}$. Then for any $\mathbf{x} \in \mathbf{A}\mathcal{V}$

$$\|\mathbf{Qx}\| \geq \|\mathbf{x}\| \quad \text{and} \quad \mathcal{V} = \mathbf{Q}(\mathbf{A}\mathcal{V}).$$

Domorádová and Dostál, Numerical linear Algebra with Applications, 2007



Proof

- Let $\mathbf{x} \in \mathbf{A}\mathcal{V}$, so that $\exists\ \mathbf{y} \in \mathbb{R}^n : \mathbf{x} = \mathbf{AQy}$

  $$\Rightarrow \mathbf{Q}^T\mathbf{x} = \mathbf{Q}^T\mathbf{AQy} = \mathbf{AQy} = \mathbf{x}$$

- and $\mathbf{x}^T\mathbf{Qx} = \mathbf{x}^T\mathbf{Q}^T\mathbf{x} = \|\mathbf{x}\|^2$, so that

$$\begin{aligned} \|\mathbf{Qx}\|^2 &= \mathbf{x}^T\mathbf{Q}^T\mathbf{Qx} = \mathbf{x}^T\left((\mathbf{Q}^T - \mathbf{I}) + \mathbf{I}\right)\left((\mathbf{Q}^T - \mathbf{I}) + \mathbf{I}\right)\mathbf{x} \\ &= \|(\mathbf{Q} - \mathbf{I})\mathbf{x}\|^2 + \|\mathbf{x}\|^2 \geq \|\mathbf{x}\|^2. \end{aligned}$$

- $\mathcal{V} = \mathrm{Im}\mathbf{Q}$, so that $\mathcal{V} = \mathbf{Q}(\mathbb{R}^n) \supseteq \mathbf{Q}(\mathbf{A}\mathcal{V})$
- $\mathbf{A}$ is nonsingular and mapping $\mathbf{A}\mathcal{V} \ni \mathbf{x} \to \mathbf{Qx}$ is injective
- dimension argument $\Rightarrow \mathbf{Q}(\mathbb{R}^n) = \mathbf{Q}(\mathbf{A}\mathcal{V})$

□

Proof

- Let $\mathbf{x} \in \mathbf{A}\mathcal{V}$, so that $\exists\, \mathbf{y} \in \mathbb{R}^n : \mathbf{x} = \mathbf{AQy}$

$$\Rightarrow \mathbf{Q}^T\mathbf{x} = \mathbf{Q}^T\mathbf{AQy} = \mathbf{AQy} = \mathbf{x}$$

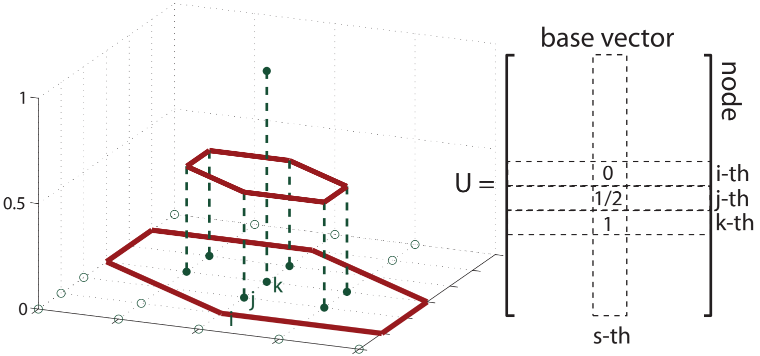
- and $\mathbf{x}^T\mathbf{Qx} = \mathbf{x}^T\mathbf{Q}^T\mathbf{x} = \|\mathbf{x}\|^2$, so that

$$\begin{aligned} \|\mathbf{Qx}\|^2 &= \mathbf{x}^T\mathbf{Q}^T\mathbf{Qx} = \mathbf{x}^T\left((\mathbf{Q}^T - \mathbf{I}) + \mathbf{I}\right)\left((\mathbf{Q}^T - \mathbf{I}) + \mathbf{I}\right)\mathbf{x} \\ &= \|(\mathbf{Q} - \mathbf{I})\mathbf{x}\|^2 + \|\mathbf{x}\|^2 \geq \|\mathbf{x}\|^2. \end{aligned}$$

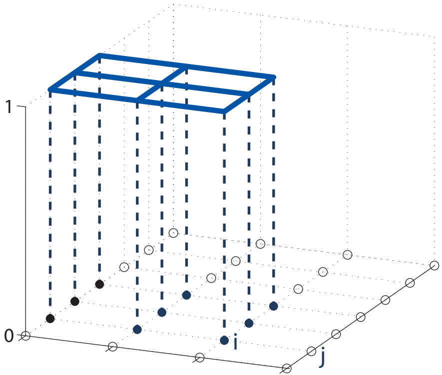
- $\mathcal{V} = \mathrm{Im}\mathbf{Q}$, so that $\mathcal{V} = \mathbf{Q}(\mathbb{R}^n) \supseteq \mathbf{Q}(\mathbf{A}\mathcal{V})$
- $\mathbf{A}$ is nonsingular and mapping $\mathbf{A}\mathcal{V} \ni \mathbf{x} \to \mathbf{Qx}$ is injective
- dimension argument $\Rightarrow \mathbf{Q}(\mathbb{R}^n) = \mathbf{Q}(\mathbf{A}\mathcal{V})$

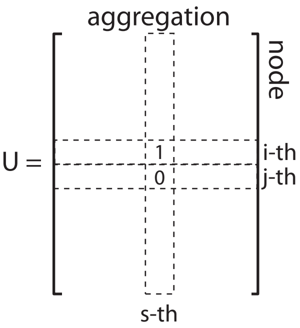
$\square$

Projector defined by traces of linear functions on coarse grid

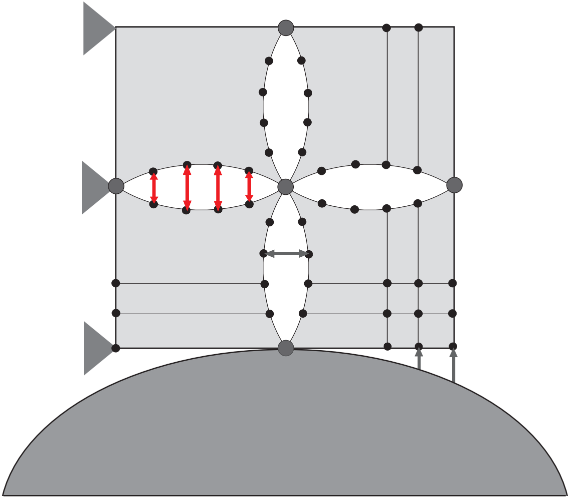


Projector defined by aggregations





Projector for FETI-DP



Preconditioning by Conjugate Projector

Decomposition by means of projectors

- $\mathbf{P} = \mathbf{U}(\mathbf{U}^T\mathbf{A}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{A}$ onto $\mathcal{U}$
- $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ onto $\mathcal{V} = \mathrm{Im}\mathbf{Q}$

$$\begin{aligned}
\min_{\mathbf{x}\in\Omega} f(\mathbf{x}) &= \min_{\substack{\mathbf{y}\in\mathcal{U},\mathbf{z}\in\mathcal{V}\\ \mathbf{y}+\mathbf{z}\in\Omega}} f(\mathbf{y}+\mathbf{z}) = \min_{\mathbf{y}\in\mathcal{U}} f(\mathbf{y}) + \min_{\mathbf{z}\in\mathcal{V}\cap\Omega} f(\mathbf{z})\\
&= f(\mathbf{x}^0) + \min_{\mathbf{z}\in\mathcal{V}\cap\Omega} f(\mathbf{z})\\
&= f(\mathbf{x}^0) + \min_{\substack{\mathbf{z}\in\mathbf{A}\mathcal{V}\\ \mathbf{z}_{\mathcal{I}}\geq\ell_{\mathcal{I}}}} \frac{1}{2}\mathbf{z}^T\mathbf{Q}^T\mathbf{A}\mathbf{Q}\mathbf{z} - \mathbf{b}^T\mathbf{Q}\mathbf{z}
\end{aligned}$$

Domorádová and Dostál, Numerical linear Algebra with Applications, 2007

Minimization over $\mathcal{U}$

$\mathbf{U} \in \mathbb{R}^{n\times p}$, $\mathbf{P} = \mathbf{U}(\mathbf{U}^T\mathbf{A}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{A}$

$$\min_{\mathbf{x}\in\mathcal{U}} f(\mathbf{x}) = \min_{\mathbf{y}\in\mathbb{R}^p} f(\mathbf{U}\mathbf{y}) = \min_{\mathbf{y}\in\mathbb{R}^p} \frac{1}{2}\mathbf{y}^T\mathbf{U}^T\mathbf{A}\mathbf{U}\mathbf{y} - \mathbf{b}^T\mathbf{U}\mathbf{y}$$

by the gradient argument

$$\mathbf{U}^T\mathbf{A}\mathbf{U}\mathbf{y} = \mathbf{U}^T\mathbf{b}.$$

Then minimizer $\mathbf{x}^0 = \mathbf{U}\mathbf{y}^0$ of f over $\mathcal{U}$ is defined by

$$\mathbf{x}^0 = \underbrace{\mathbf{U}(\mathbf{U}^T\mathbf{A}\mathbf{U})^{-1}\mathbf{U}^T}_{\mathbf{P}\mathbf{A}^{-1}}\mathbf{b} = \mathbf{P}\mathbf{A}^{-1}\mathbf{b}.$$

Observation

$$\mathbf{P}^{\mathsf{T}}\mathbf{A} = \mathbf{A}\mathbf{P} \Rightarrow \mathbf{P}^{\mathsf{T}} = \mathbf{A}\mathbf{P}\mathbf{A}^{-1}$$

- $\mathbf{g}^0 = \mathbf{A}\mathbf{x}^0 - \mathbf{b} = \mathbf{A}\underbrace{\mathbf{P}\mathbf{A}^{-1}\mathbf{b}}_{x^0} - \mathbf{b} = (\mathbf{A}\mathbf{P}\mathbf{A}^{-1} - \mathbf{I})\mathbf{b}$
  $= (\mathbf{P}^T - \mathbf{I})\mathbf{b} = -\mathbf{Q}^T\mathbf{b} \Rightarrow \mathbf{g}^0 \in \mathrm{Im}\mathbf{Q}^{\mathrm{T}}$
- $\mathrm{Im}\mathbf{Q}^{\mathrm{T}} = \mathrm{Im}(\mathbf{Q}^{\mathrm{T}}\mathbf{A}) = \mathrm{Im}(\mathbf{A}\mathbf{Q}) = \mathbf{A}\mathcal{V}$

$\Rightarrow \mathbf{g}^0 \in \mathbf{A}\mathcal{V}$.

Lemma

Let $\mathbf{z}^1, \mathbf{z}^2, \ldots$ be generated by the MPRGP algorithm for the problem

$$\min_{\mathbf{z}_{\mathcal{I}} \geq \ell_{\mathcal{I}}} \frac{1}{2} \mathbf{z}^T \mathbf{Q}^T \mathbf{A} \mathbf{Q} \mathbf{z} + \left(\mathbf{g}^0\right)^T \mathbf{z}$$

starting from $\mathbf{z}^0 = P_\Omega\left(\mathbf{g}^0\right)$.

Then $\mathbf{z}^k \in \mathbf{A}\mathcal{V},\ k = 0, 1, 2, \ldots$.

$\Rightarrow$ $\mathbf{A}\mathcal{V}$ is invariant subspace of P_Ω

Preconditioning Effect

- $\lambda_{max} \geq \cdots \geq \lambda_{min}$ the eigenvalues of $\mathbf{A}$
- $\mathcal{E} = p$-dimensional subspace spanned by the eigenvectors corresponding to the p smallest eigenvalues of $\mathbf{A}$
- $\overline{\gamma} = \|\mathbf{R}_{\mathbf{A}\mathcal{U}} - \mathbf{R}_{\mathcal{E}}\| \ldots$ gap between $\mathbf{A}\mathcal{U}$ and $\mathcal{E}$

$$\overline{\lambda}_{min} \geq \sqrt{(1-\overline{\gamma}^2)\lambda_{min-m}^2 + \overline{\gamma}^2\lambda_{min}^2}$$

$\overline{\lambda}_{min}$ = min eigenvalue of $(\mathbf{Q}^T\mathbf{A}\mathbf{Q}|\mathbf{A}\mathcal{V})$

Dostál, International Journal of Computer Mathematics, 1988

Rate of Convergence of Reduced Problem

$\hat{\mathbf{z}}$ = unique solution, $\mathbf{g}^0 = -\mathbf{Q}^T\mathbf{b}$

$$\min_{\substack{\mathbf{z}\in\mathbf{A}\mathcal{V}\\ \mathbf{z}_\mathcal{I}\geq\ell_\mathcal{I}}} \underbrace{\frac{1}{2}\mathbf{z}^T\mathbf{Q}^T\mathbf{A}\mathbf{Q}\mathbf{z} + \left(\mathbf{g}^0\right)^T\mathbf{z}}_{f_{0,\mathbf{Q}}}$$

$$f_{0,\mathbf{Q}}(\mathbf{z}^{k+1}) - f_{0,\mathbf{Q}}(\widehat{\mathbf{z}}) \leq \overline{\eta}\left(f_{0,\mathbf{Q}}(\mathbf{z}^k) - f_{0,\mathbf{Q}}(\widehat{\mathbf{z}})\right),$$

where

$$\overline{\eta} = 1 - \frac{\overline{\alpha}\overline{\lambda}_{\min}}{2 + 2\widehat{\Gamma}^2}$$

Improved Estimate

MPRGP

$$\|\mathbf{x}^k - \widehat{\mathbf{x}}\|_{\mathbf{A}}^2 \leq 2\eta \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}}) \right)$$

MPRGP-CP

$$f_{0,\mathbf{Q}}(\mathbf{z}^{k+1}) - f_{0,\mathbf{Q}}(\widehat{\mathbf{z}}) \leq \overline{\eta} \left(f_{0,\mathbf{Q}}(\mathbf{z}^k) - f_{0,\mathbf{Q}}(\widehat{\mathbf{z}}) \right)$$

$$\overline{\eta} < \eta$$

because for nonsingular $\mathbf{U}^T\mathbf{E}$ and $\lambda_{min} < \lambda_{min-m}$

$$\overline{\eta} = 1 - \frac{\overline{\alpha}\overline{\lambda}_{\min}}{2 + 2\widehat{\Gamma}^2} < 1 - \frac{\overline{\alpha}\lambda_{\min}}{2 + 2\widehat{\Gamma}^2} = \eta$$

Domorádová and Dostál, Numerical linear Algebra with Applications, 2007

Numerical Experiments

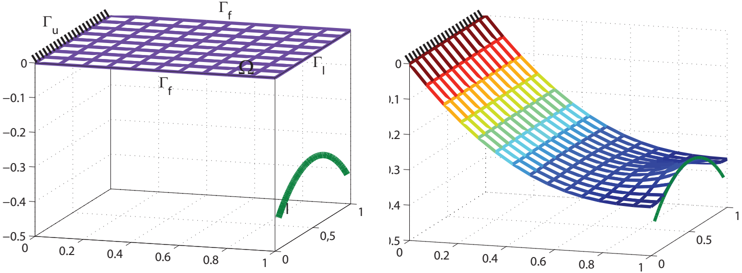
Deflection of the membrane

Find

$$\min f(u) = \frac{1}{2}\int_{\Omega} \|\nabla u(x)\|^2 \mathrm{d}\Omega + \int_{\Omega} u \,\mathrm{d}\Omega \quad \text{subject to} \quad u \in \mathcal{K},$$

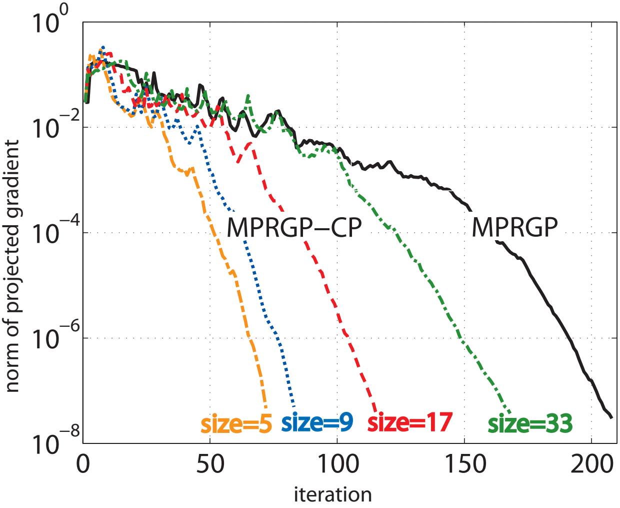
where

$$\mathcal{K} = \{u \in H^1(\Omega) : \; u = 0 \;\text{ on }\; \Gamma_u \quad \text{and} \quad c \le u \;\text{ on }\; \Gamma_c\}.$$

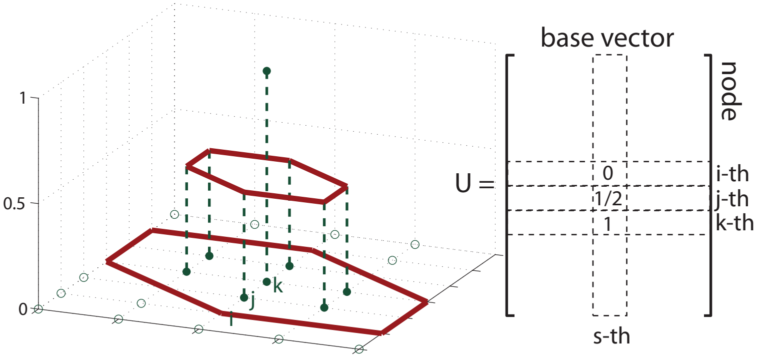


Numerical Experiments

Projector deffined by traces of linear functions on coarse grid

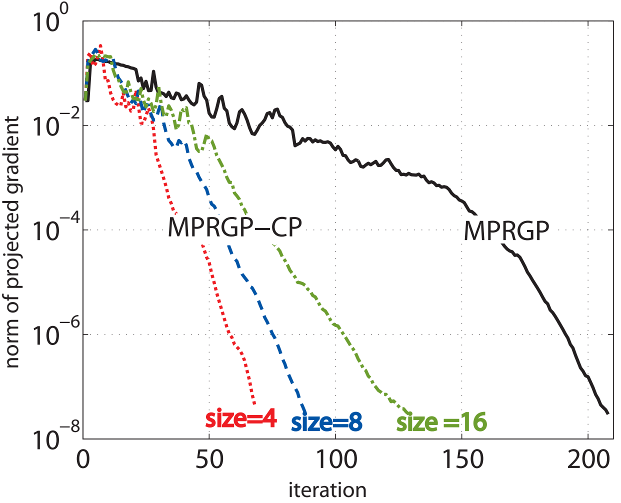


Projector defined by traces of linear functions on coarse grid

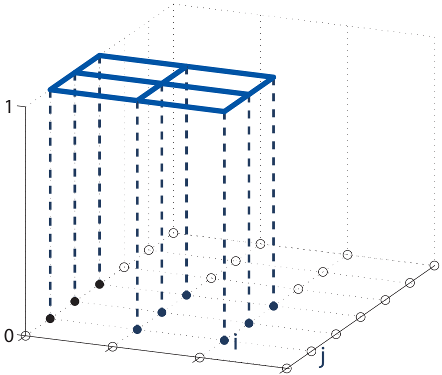


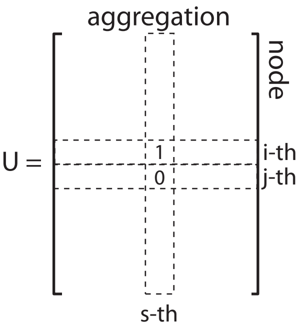
Numerical Experiments

Projector defined by aggregations



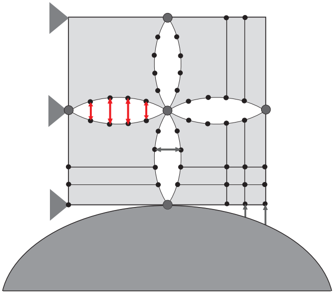
Projector defined by aggregations





Numerical Experiments

FETI-DP



number of subdomains	number of nodes 2		4		8		16	
2	6	8	9	13	16	18	21	29
4	9	11	13	17	19	30	25	53
8	12	18	19	29	25	56	36	80
16	17	24	26	55	39	85	-	-

MPRGP-CP MPRGP

Summary

- The preconditioning effects all steps of the algorithm.
- Better rate of convergence is proved.
- It is confirmed by numerical experiments.

Our Outlook

- to examine implementation by averaging.
- to implement the algorithm into our research FETI and BETI software OOSOL.
- to consider iterative implementation of inversion in the projector.