Accelerating MPGP-type Methods Through Preconditioning

Jakub Kruzik, David Horak
jakub.kruzik@ugn.cas.cz

May 19, 2026
HPCSE 2026, Solan

Institute of Geonics, Czech Academy of Sciences
Dept. Appl. Math., Fac. Electr. Eng. Comput. Sci., VSB-Technical University of Ostrava

Co-funded by the European Union
Ministry of Education, Youth and Sports
INODIN
Co-funded by the European Union under the project INODIN (CZ.02.01.01/00/23_020/0008487)

Quadratic Programming (QP) Problem

$$\arg\min_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{b} \quad \text{s.t.} \quad \boldsymbol{x} \in \Omega$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is symmetric and Ω is closed and convex.

Here $\boldsymbol{A}$ is also positive definite and $\Omega = \{\boldsymbol{x} \mid \boldsymbol{x} \geq \boldsymbol{l}\}$.

Applications:
contact problems, data denoising, machine learning, ice-sheet melting,...

Active / Free set:

$$\mathcal{A}(\boldsymbol{x}) = \{j : x_j = l_j\} \qquad \mathcal{F}(\boldsymbol{x}) = \{j : l_j < x_j\}$$

MPGP-type Algorithms: Ingredients

Active / Free set:

$$\mathcal{A}(\boldsymbol{x}) = \{j : x_j = l_j\} \qquad \mathcal{F}(\boldsymbol{x}) = \{j : l_j < x_j\}$$

Gradient splitting ($\boldsymbol{g} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}$):

$$g_j^f = \begin{cases} 0 & \text{if} \quad j \in \mathcal{A}, \\ g_j & \text{if} \quad j \in \mathcal{F}. \end{cases} \qquad g_j^c = \begin{cases} 0 & \text{if} \quad j \in \mathcal{F}, \\ \min(g_j, 0) & \text{if} \quad j \in \mathcal{A}, \end{cases}$$

MPGP-type Algorithms: Ingredients

Active / Free set:

$$\mathcal{A}(\boldsymbol{x}) = \{j : x_j = l_j\} \qquad \mathcal{F}(\boldsymbol{x}) = \{j : l_j < x_j\}$$

Gradient splitting ($\boldsymbol{g} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}$):

$$g_j^f = \begin{cases} 0 & \text{if} \quad j \in \mathcal{A}, \\ g_j & \text{if} \quad j \in \mathcal{F}. \end{cases} \qquad g_j^c = \begin{cases} 0 & \text{if} \quad j \in \mathcal{F}, \\ \min(g_j, 0) & \text{if} \quad j \in \mathcal{A}, \end{cases}$$

Projected gradient:

$$\boldsymbol{g}^P = \boldsymbol{g}^f + \boldsymbol{g}^c$$

MPGP-type Algorithms: Ingredients

Active / Free set:

$$\mathcal{A}(\boldsymbol{x}) = \{j : x_j = l_j\} \qquad \mathcal{F}(\boldsymbol{x}) = \{j : l_j < x_j\}$$

Gradient splitting ($\boldsymbol{g} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}$):

$$g_j^f = \begin{cases} 0 & \text{if} \quad j \in \mathcal{A}, \\ g_j & \text{if} \quad j \in \mathcal{F}. \end{cases} \qquad g_j^c = \begin{cases} 0 & \text{if} \quad j \in \mathcal{F}, \\ \min(g_j, 0) & \text{if} \quad j \in \mathcal{A}, \end{cases}$$

Projected gradient:

$$\boldsymbol{g}^P = \boldsymbol{g}^f + \boldsymbol{g}^c$$

Projection onto the feasible set Ω:

$$[P_\Omega(\boldsymbol{x})]_j = \max(l_j, x_j).$$

MPRGP - Modified Prop. with Reduced Grad. Projections [Dostál, Schöberl ’05]

Input: $\boldsymbol{A}$, $\boldsymbol{x}_0 \in \Omega$, $\boldsymbol{b}$, $\Gamma > 0$, $\overline{\alpha} \in (0, 2||\boldsymbol{A}||^{-1}]$

$\boldsymbol{g}_0 = \boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}$, $\boldsymbol{p}_0 = \boldsymbol{g}_0^f$, $k = 0$

while $||\boldsymbol{g}_k^P||$ *is not small*:

  if $||\boldsymbol{g}_k^c|| \leq \Gamma||\boldsymbol{g}_k^f||$:

    $\alpha_k^{feas} = \max\{\alpha : \boldsymbol{x}_k - \alpha\boldsymbol{p}_k \in \Omega\}$

    $\alpha_k^{cg} = \boldsymbol{g}_k^T\boldsymbol{p}_k / \boldsymbol{p}_k^T\boldsymbol{A}\boldsymbol{p}_k$

    if $\alpha_k^{cg} \leq \alpha_k^{feas}$:

      Conjugate Gradient (CG) step

    else:

      Expansion step

  else:

    Proportioning step

  $k = k + 1$

Output: $\boldsymbol{x}_k$

MPRGP - Modified Prop. with Reduced Grad. Projections [Dostál, Schöberl '05]

Input: $\boldsymbol{A}$, $\boldsymbol{x}_0 \in \Omega$, $\boldsymbol{b}$, $\Gamma > 0$, $\overline{\alpha} \in (0, 2||\boldsymbol{A}||^{-1}]$

```
g_0 = A x_0 - b, p_0 = g_0^f, k = 0
while ||g_k^P|| is not small:
    if ||g_k^c|| <= Γ||g_k^f||:
        α_k^feas = max{α : x_k - α p_k ∈ Ω}
        α_k^cg = g_k^T p_k / p_k^T A p_k
        if α_k^cg <= α_k^feas:
            Conjugate Gradient (CG) step
        else:
            Expansion step
    else:
        Proportioning step
    k = k + 1
```

Output: $\boldsymbol{x}_k$

CG step:

$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha_k^{cg} \boldsymbol{p}_k$

$\boldsymbol{g}_{k+1} = \boldsymbol{g}_k - \alpha_k^{cg} \boldsymbol{A} \boldsymbol{p}_k$

$\beta_k = \boldsymbol{p}_k^T \boldsymbol{A} \boldsymbol{g}_{k+1}^f / \boldsymbol{p}_k^T \boldsymbol{A} \boldsymbol{p}_k$

$\boldsymbol{p}_{k+1} = \boldsymbol{g}_{k+1}^f - \beta_k \boldsymbol{p}_k$

MPRGP - Modified Proportioning with Reduced Gradient Projections

Input: $\boldsymbol{A}$, $\boldsymbol{x}_0 \in \Omega$, $\boldsymbol{b}$, $\Gamma > 0$, $\overline{\alpha} \in (0, 2||\boldsymbol{A}||^{-1}]$

$\boldsymbol{g}_0 = \boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}$, $\boldsymbol{p}_0 = \boldsymbol{g}_0^f$, $k = 0$
while $||\boldsymbol{g}_k^P||$ *is not small*:
    if $||\boldsymbol{g}_k^c|| \leq \Gamma||\boldsymbol{g}_k^f||$:
        $\alpha_k^{feas} = \max\{\alpha : \boldsymbol{x}_k - \alpha\boldsymbol{p}_k \in \Omega\}$
        $\alpha_k^{cg} = \boldsymbol{g}_k^T\boldsymbol{p}_k/\boldsymbol{p}_k^T\boldsymbol{A}\boldsymbol{p}_k$
        if $\alpha_k^{cg} \leq \alpha_k^{feas}$:
            Conjugate Gradient (CG) step
        else:
            Expansion step
    else:
        Proportioning step
    $k = k + 1$

Output: $\boldsymbol{x}_k$

Expansion step:

$\boldsymbol{x}_{k+\frac{1}{2}} = \boldsymbol{x}_k - \alpha_k^{feas}\boldsymbol{p}_k$
$\boldsymbol{g}_{k+\frac{1}{2}} = \boldsymbol{g}_k - \alpha_k^{feas}\boldsymbol{A}\boldsymbol{p}_k$
$\boldsymbol{x}_{k+1} = P_\Omega(\boldsymbol{x}_{k+\frac{1}{2}} - \overline{\alpha}\boldsymbol{g}_{k+\frac{1}{2}}^f)$
$\boldsymbol{g}_{k+1} = \boldsymbol{A}\boldsymbol{x}_{k+1} - \boldsymbol{b}$
$\boldsymbol{p}_{k+1} = \boldsymbol{g}_{k+1}^f$

MPRGP - Modified Proportioning with Reduced Gradient Projections

Input: $\boldsymbol{A}$, $\boldsymbol{x}_0 \in \Omega$, $\boldsymbol{b}$, $\Gamma > 0$, $\overline{\alpha} \in (0, 2||\boldsymbol{A}||^{-1}]$

$\boldsymbol{g}_0 = \boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}$, $\boldsymbol{p}_0 = \boldsymbol{g}_0^f$, $k = 0$
while $||\boldsymbol{g}_k^P||$ *is not small*:
    if $||\boldsymbol{g}_k^c|| \leq \Gamma ||\boldsymbol{g}_k^f||$:
        $\alpha_k^{feas} = \max\{\alpha : \boldsymbol{x}_k - \alpha \boldsymbol{p}_k \in \Omega\}$
        $\alpha_k^{cg} = \boldsymbol{g}_k^T \boldsymbol{p}_k / \boldsymbol{p}_k^T \boldsymbol{A} \boldsymbol{p}_k$
        if $\alpha_k^{cg} \leq \alpha_k^{feas}$:
            Conjugate Gradient (CG) step
        else:
            Expansion step
    else:
        Proportioning step
    $k = k + 1$
Output: $\boldsymbol{x}_k$

Expansion step:

$\boldsymbol{x}_{k+\frac{1}{2}} = \boldsymbol{x}_k - \alpha_k^{feas} \boldsymbol{p}_k$
$\boldsymbol{g}_{k+\frac{1}{2}} = \boldsymbol{g}_k - \alpha_k^{feas} \boldsymbol{A} \boldsymbol{p}_k$
$\boldsymbol{x}_{k+1} = P_\Omega(\boldsymbol{x}_{k+\frac{1}{2}} - \overline{\alpha} \boldsymbol{g}^f_{k+\frac{1}{2}})$
$\boldsymbol{g}_{k+1} = \boldsymbol{A}\boldsymbol{x}_{k+1} - \boldsymbol{b}$
$\boldsymbol{p}_{k+1} = \boldsymbol{g}^f_{k+1}$

Proportioning step:

$\alpha_k = \boldsymbol{g}_k^T \boldsymbol{g}_k^c / (\boldsymbol{g}_k^c)^T \boldsymbol{A} \boldsymbol{g}_k^c$
$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha_k \boldsymbol{g}_k^c$
$\boldsymbol{g}_{k+1} = \boldsymbol{g}_k - \alpha_k \boldsymbol{A} \boldsymbol{g}_k^c$
$\boldsymbol{p}_{k+1} = \boldsymbol{g}^f_{k+1}$

MPRGP Q-linear and R-linear Convergence

[Z. Dostál, J. Schöberl '05]

Let $\boldsymbol{x}_k$ be generate by MRPGP, $\boldsymbol{x}_0 \in \Omega$, $\Gamma > 0$ and $\overline{\alpha} \in \left(0, 2||\boldsymbol{A}||^{-1}\right]$. Then

$$f(\boldsymbol{x}_{k+1}) - f(\widehat{\boldsymbol{x}}) \leq \eta \left(f(\boldsymbol{x}_k) - f(\widehat{\boldsymbol{x}})\right) \quad \text{and} \quad ||\boldsymbol{x}_k - \widehat{\boldsymbol{x}}||_{\boldsymbol{A}} \leq 2\eta^k \left(f(\boldsymbol{x}_0) - f(\widehat{\boldsymbol{x}})\right),$$

where $\widehat{\boldsymbol{x}}$ denotes the unique solution,

$$\eta = 1 - \frac{\widehat{\alpha}\lambda_{min}}{\vartheta\left(1 + \widehat{\Gamma}^2\right)},$$

$$\widehat{\Gamma} = \max\{\Gamma, \Gamma^{-1}\}, \qquad \vartheta = 2\max\{\overline{\alpha}||\boldsymbol{A}||, 1\}, \qquad \widehat{\alpha} = \min\{\overline{\alpha}, 2||\boldsymbol{A}||^{-1} - \overline{\alpha}\}.$$

$$\eta^{opt} = 1 - \kappa\left(\boldsymbol{A}\right)^{-1}/4$$

for $\Gamma = 1$ and $\overline{\alpha} = ||\boldsymbol{A}||^{-1}$

Preconditioning

$$\arg\min_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{b} \quad \text{s.t.} \quad \boldsymbol{l} \leq \boldsymbol{x},$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is SPD.

Apply SPD preconditioner

$$\boldsymbol{M} = \boldsymbol{L}\boldsymbol{L}^T$$

$$\arg\min_{\boldsymbol{u}} \frac{1}{2}\boldsymbol{u}^T\boldsymbol{L}^{-1}\boldsymbol{A}\boldsymbol{L}^{-T}\boldsymbol{u} - \boldsymbol{u}^T\boldsymbol{L}^{-1}\boldsymbol{b} \quad \text{s.t.} \quad \boldsymbol{l} \leq \boldsymbol{L}^{-T}\boldsymbol{u},$$

with $\boldsymbol{x} = \boldsymbol{L}^{-T}\boldsymbol{u}$.

Preconditioned MPRGP

Input: $\boldsymbol{A}$, $\boldsymbol{M}^{-1}$, $\boldsymbol{x}_0 \in \Omega$, $\boldsymbol{b}$, $\Gamma > 0$, $\overline{\alpha} \in (0, 2||\boldsymbol{A}||^{-1})$

```
g_0 = A x_0 - b, z_0 = M^{-1} g_0^f, p_0 = z_0, k = 0
while ||g_k^P|| is not small:
    if ||g_k^c||^2 <= Γ^2 ||g_k^f||^2:
        α_k^feas = max{α | x_k - α p_k ∈ Ω}
        α_k^cg = g_k^T z_k / p_k^T A p_k
        if α_k^cg <= α_k^feas:
            Preconditioned CG step
        else:
            Preconditioned expansion step
    else:
        Preconditioned proportioning step
    k = k + 1
```

$\boldsymbol{g}_0 = \boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}$, $\boldsymbol{z}_0 = \boldsymbol{M}^{-1}\boldsymbol{g}_0^f$, $\boldsymbol{p}_0 = \boldsymbol{z}_0$, $k = 0$

while $||\boldsymbol{g}_k^P||$ *is not small*:

- if $||\boldsymbol{g}_k^c||^2 \leq \Gamma^2 ||\boldsymbol{g}_k^f||^2$:
  - $\alpha_k^{feas} = \max\{\alpha \mid \boldsymbol{x}_k - \alpha \boldsymbol{p}_k \in \Omega\}$
  - $\alpha_k^{cg} = \boldsymbol{g}_k^T \boldsymbol{z}_k / \boldsymbol{p}_k^T \boldsymbol{A} \boldsymbol{p}_k$
  - if $\alpha_k^{cg} \leq \alpha_k^{feas}$:
    - Preconditioned CG step
  - else:
    - Preconditioned expansion step
- else:
  - Preconditioned proportioning step
- $k = k + 1$

Output: $\boldsymbol{x}_k$

Precond. CG step:

$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha_k^{cg} \boldsymbol{p}_k$

$\boldsymbol{g}_{k+1} = \boldsymbol{g}_k - \alpha_k^{cg} \boldsymbol{A} \boldsymbol{p}_k$

$\boldsymbol{z}_{k+1} = \boldsymbol{M}^{-1} \boldsymbol{g}_{k+1}^f$

$\beta_k = \boldsymbol{p}_k^T \boldsymbol{A} \boldsymbol{z}_{k+1} / \boldsymbol{p}_k^T \boldsymbol{A} \boldsymbol{p}_k$

$\boldsymbol{p}_{k+1} = \boldsymbol{z}_{k+1} - \beta_k \boldsymbol{p}_k$

Preconditioned MPRGP

Input: $\boldsymbol{A}$, $\boldsymbol{M}^{-1}$, $\boldsymbol{x}_0 \in \Omega$, $\boldsymbol{b}$, $\Gamma > 0$, $\overline{\alpha} \in (0, 2||\boldsymbol{A}||^{-1})$

$\boldsymbol{g}_0 = \boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}$, $\boldsymbol{z}_0 = \boldsymbol{M}^{-1}\boldsymbol{g}_0^f$, $\boldsymbol{p}_0 = \boldsymbol{z}_0$, $k = 0$

while $||\boldsymbol{g}_k^P||$ *is not small*:

  if $||\boldsymbol{g}_k^c||^2 \leq \Gamma^2 ||\boldsymbol{g}_k^f||^2$:

    $\alpha_k^{feas} = \max\{\alpha \mid \boldsymbol{x}_k - \alpha \boldsymbol{p}_k \in \Omega\}$

    $\alpha_k^{cg} = \boldsymbol{g}_k^T \boldsymbol{z}_k / \boldsymbol{p}_k^T \boldsymbol{A} \boldsymbol{p}_k$

    if $\alpha_k^{cg} \leq \alpha_k^{feas}$:

      Preconditioned CG step

    else:

      Preconditioned expansion step

  else:

    Preconditioned proportioning step

  $k = k + 1$

Output: $\boldsymbol{x}_k$

Precond. Expansion step:

$\boldsymbol{x}_{k+\frac{1}{2}} = \boldsymbol{x}_k - \alpha_k^{feas} \boldsymbol{p}_k$

$\boldsymbol{g}_{k+\frac{1}{2}} = \boldsymbol{g}_k - \alpha_k^{feas} \boldsymbol{A} \boldsymbol{p}_k$

$\boldsymbol{x}_{k+1} = P_\Omega(\boldsymbol{x}_{k+\frac{1}{2}} - \overline{\alpha} \boldsymbol{g}_k^f)$

$\boldsymbol{g}_{k+1} = \boldsymbol{A}\boldsymbol{x}_{k+1} - \boldsymbol{b}$

$\boldsymbol{z}_{k+1} = \boldsymbol{M}^{-1} \boldsymbol{g}_{k+1}^f$

$\boldsymbol{p}_{k+1} = \boldsymbol{z}_{k+1}$

Precond. Proportioning step:

$\alpha_k^{sd} = \boldsymbol{g}_k^T \boldsymbol{g}_k^c / (\boldsymbol{g}_k^c)^T \boldsymbol{A} \boldsymbol{g}_k^c$

$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha_k^{sd} \boldsymbol{g}_k^c$

$\boldsymbol{g}_{k+1} = \boldsymbol{g}_k - \alpha_k^{sd} \boldsymbol{A} \boldsymbol{g}_k^c$

$\boldsymbol{z}_{k+1} = \boldsymbol{M}^{-1} \boldsymbol{g}_{k+1}^f$

$\boldsymbol{p}_{k+1} = \boldsymbol{z}_{k+1}$

$\mathcal{A}$ is active set, $\mathcal{F}$ is free set

$$\overline{\boldsymbol{M}} = \begin{pmatrix} \boldsymbol{M}_{\mathcal{F}\mathcal{F}} & \boldsymbol{M}_{\mathcal{F}\mathcal{A}} \\ \boldsymbol{M}_{\mathcal{A}\mathcal{F}} & \boldsymbol{M}_{\mathcal{A}\mathcal{A}} \end{pmatrix}$$

Precondition only on the free set

$$\boldsymbol{z} = \begin{pmatrix} \boldsymbol{z}_{\mathcal{F}}^{f} \\ \boldsymbol{o} \end{pmatrix} = \boldsymbol{M}^{-1} \begin{pmatrix} \boldsymbol{g}_{\mathcal{F}}^{f} \\ \boldsymbol{o} \end{pmatrix} := \begin{pmatrix} \boldsymbol{M}_{\mathcal{F}\mathcal{F}}^{-1} & \boldsymbol{o} \\ \boldsymbol{o} & \boldsymbol{o} \end{pmatrix} \begin{pmatrix} \boldsymbol{g}_{\mathcal{F}}^{f} \\ \boldsymbol{o} \end{pmatrix}.$$

[1] Originally for Polyak's algorithm [1969]

Approximate Preconditioning in Face

$$\boldsymbol{z} = \begin{pmatrix} \widetilde{\boldsymbol{z}}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix} = \boldsymbol{M}^{-1} \begin{pmatrix} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix} := \begin{pmatrix} \boldsymbol{I}_{\mathcal{F}\mathcal{F}} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} \overline{\boldsymbol{M}}^{-1} \begin{pmatrix} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix}$$

Approximate Preconditioning in Face

$$\boldsymbol{z} = \begin{pmatrix} \widetilde{\boldsymbol{z}}_{\mathcal{F}}^{f} \\ \boldsymbol{o} \end{pmatrix} = \boldsymbol{M}^{-1} \begin{pmatrix} \boldsymbol{g}_{\mathcal{F}}^{f} \\ \boldsymbol{o} \end{pmatrix} := \begin{pmatrix} \boldsymbol{I}_{\mathcal{F}\mathcal{F}} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} \overline{\boldsymbol{M}}^{-1} \begin{pmatrix} \boldsymbol{g}_{\mathcal{F}}^{f} \\ \boldsymbol{o} \end{pmatrix}$$

assuming $\overline{\boldsymbol{M}}^{-1}$ is inverse

$$= \begin{pmatrix} (\boldsymbol{M}_{\mathcal{F}\mathcal{F}} - \boldsymbol{M}_{\mathcal{F}\mathcal{A}} \boldsymbol{M}_{\mathcal{A}\mathcal{A}}^{-1} \boldsymbol{M}_{\mathcal{A}\mathcal{F}})^{-1} \boldsymbol{g}_{\mathcal{F}}^{f} \\ \boldsymbol{o} \end{pmatrix} = \begin{pmatrix} \boldsymbol{S}^{-1} \boldsymbol{g}_{\mathcal{F}}^{f} \\ \boldsymbol{o} \end{pmatrix}$$

Approximate Preconditioning in Face

$$\boldsymbol{z} = \begin{pmatrix} \widetilde{\boldsymbol{z}}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix} = \boldsymbol{M}^{-1} \begin{pmatrix} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix} := \begin{pmatrix} \boldsymbol{I}_{\mathcal{F}\mathcal{F}} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} \overline{\boldsymbol{M}}^{-1} \begin{pmatrix} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix}$$

assuming $\overline{\boldsymbol{M}}^{-1}$ is inverse

$$= \begin{pmatrix} (\boldsymbol{M}_{\mathcal{F}\mathcal{F}} - \boldsymbol{M}_{\mathcal{F}\mathcal{A}} \boldsymbol{M}^{-1}_{\mathcal{A}\mathcal{A}} \boldsymbol{M}_{\mathcal{A}\mathcal{F}})^{-1} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix} = \begin{pmatrix} \boldsymbol{S}^{-1} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix}$$

$$= \begin{pmatrix} (\boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} + \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} \boldsymbol{M}_{\mathcal{F}\mathcal{A}} (\boldsymbol{M}_{\mathcal{A}\mathcal{A}} - \boldsymbol{M}_{\mathcal{A}\mathcal{F}} \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} \boldsymbol{M}_{\mathcal{F}\mathcal{A}})^{-1} \boldsymbol{M}_{\mathcal{A}\mathcal{F}} \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}}) \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix}$$

$$= \begin{pmatrix} (\boldsymbol{I} + \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} \boldsymbol{M}_{\mathcal{F}\mathcal{A}} (\boldsymbol{M}_{\mathcal{A}\mathcal{A}} - \boldsymbol{M}_{\mathcal{A}\mathcal{F}} \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} \boldsymbol{M}_{\mathcal{F}\mathcal{A}})^{-1} \boldsymbol{M}_{\mathcal{A}\mathcal{F}}) \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix}$$

Approximate Preconditioning in Face

$$\boldsymbol{z} = \begin{pmatrix} \widetilde{\boldsymbol{z}}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix} = \boldsymbol{M}^{-1} \begin{pmatrix} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix} := \begin{pmatrix} \boldsymbol{I}_{\mathcal{F}\mathcal{F}} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} \overline{\boldsymbol{M}}^{-1} \begin{pmatrix} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix} \qquad \text{assuming } \overline{\boldsymbol{M}}^{-1} \text{is inverse}$$

$$= \begin{pmatrix} (\boldsymbol{M}_{\mathcal{F}\mathcal{F}} - \boldsymbol{M}_{\mathcal{F}\mathcal{A}} \boldsymbol{M}^{-1}_{\mathcal{A}\mathcal{A}} \boldsymbol{M}_{\mathcal{A}\mathcal{F}})^{-1} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix} = \begin{pmatrix} \boldsymbol{S}^{-1} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix}$$

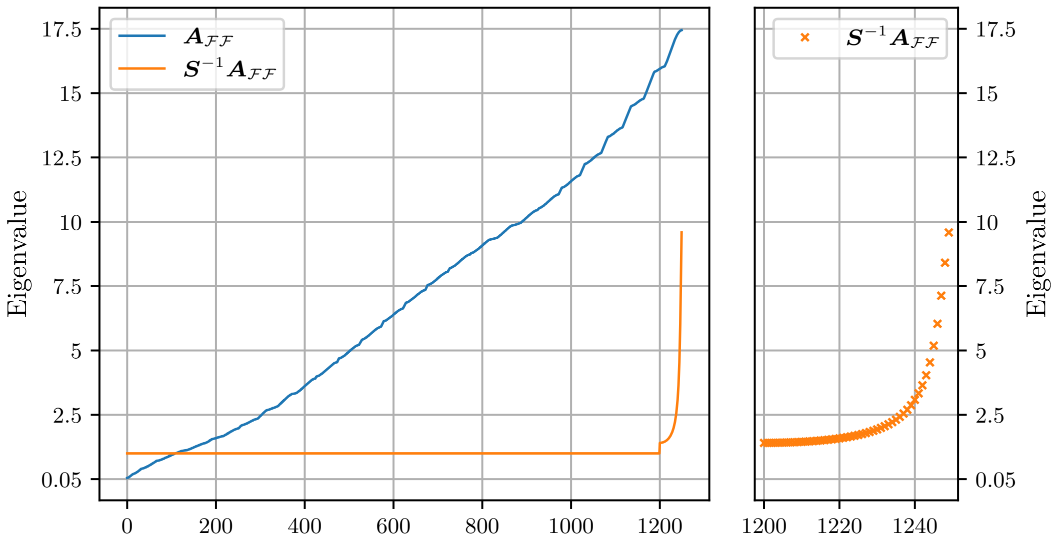
$$= \begin{pmatrix} (\boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} + \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} \boldsymbol{M}_{\mathcal{F}\mathcal{A}} (\boldsymbol{M}_{\mathcal{A}\mathcal{A}} - \boldsymbol{M}_{\mathcal{A}\mathcal{F}} \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} \boldsymbol{M}_{\mathcal{F}\mathcal{A}})^{-1} \boldsymbol{M}_{\mathcal{A}\mathcal{F}} \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}}) \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix}$$

$$= \begin{pmatrix} (\boldsymbol{I} + \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} \boldsymbol{M}_{\mathcal{F}\mathcal{A}} (\boldsymbol{M}_{\mathcal{A}\mathcal{A}} - \boldsymbol{M}_{\mathcal{A}\mathcal{F}} \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} \boldsymbol{M}_{\mathcal{F}\mathcal{A}})^{-1} \boldsymbol{M}_{\mathcal{A}\mathcal{F}}) \boldsymbol{M}^{-1}_{\mathcal{F}\mathcal{F}} \boldsymbol{g}^f_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix}$$

Let $\overline{\boldsymbol{M}} = \boldsymbol{A}$ and $r = \text{rank}(\boldsymbol{M}_{\mathcal{A}\mathcal{F}})$ then the preconditioned operator $\boldsymbol{S}^{-1} \boldsymbol{A}_{\mathcal{F}\mathcal{F}}$ has eigenvalues

$$1 = \lambda_1 = \cdots = \lambda_{n-r} \leq \cdots \leq \lambda_n$$

Eigenvalues – Inverse Precond. – Journal Bearing – 50x50 Grid Points (2500 DoFs)



Approximate Preconditioning in Face Condition Number

Let $\overline{\boldsymbol{M}} = \boldsymbol{A}$, the inner preconditioner be the inverse of $\overline{\boldsymbol{M}}$, and γ be the Cauchy–Bunyakowski–Schwarz (CBS) constant given by the active/free set splitting of the Hessian $\boldsymbol{A}$. Then

$$\kappa_{eff}\left(\boldsymbol{M}^{-1}\boldsymbol{A}\right) \leq \frac{1}{1-\gamma^2} \leq \frac{\left(\kappa\left(\boldsymbol{A}\right)+1\right)^2}{4\kappa(\boldsymbol{A})}.$$

The bound is sharp iff

$$\begin{pmatrix} \boldsymbol{v}_{\mathcal{F}} \\ \boldsymbol{o} \end{pmatrix} = \boldsymbol{v}_{min} + \boldsymbol{v}_{max} \quad \text{and} \quad \boldsymbol{S}\boldsymbol{v}_{\mathcal{F}} = \lambda \boldsymbol{v}_{\mathcal{F}}, \tag{1}$$

where $\boldsymbol{v}_{min}$ and $\boldsymbol{v}_{max}$ are eigenvectors belonging to minimal and maximal eigenvalues of $\boldsymbol{A}$, respectively, and

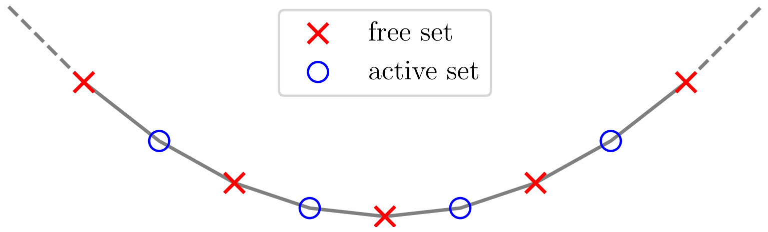
$$\text{nullity}\left(\boldsymbol{M}_{\mathcal{AF}}\right) \neq 0. \tag{2}$$

The Worst Case Example

Discretization of $-u''(x) = f$:

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \in \mathbb{R}^{n\times n},$$

where $n \geq 3$ is odd and $\mathcal{F} = \{1, 3, 5, \dots, n\}$.



The Worst Case Example

$$\kappa_{eff}\left(\boldsymbol{M}^{-1}\boldsymbol{A}\right) \leq \frac{1}{1-\gamma^2}$$

$\gamma \in [0,1)$ represents the ”strength of coupling” between $\boldsymbol{A}_{\mathcal{FF}}$ and $\boldsymbol{A}_{\mathcal{AA}}$

Numerically:

n	γ	$\kappa_{eff}\left(\boldsymbol{M}^{-1}\boldsymbol{A}\right)$	$\left(\kappa\left(\boldsymbol{A}\right)+1\right)^2 / \left(4\kappa(\boldsymbol{A})\right)$
3	0.7071	2.0000	2.0000
9	0.9511	10.4721	10.4721
99	0.9995	$1,013.5452$	$1,013.5452$
999	1.0000	$10,1321.5170$	$10,1321.5170$

The bounds are attained analytically as well.

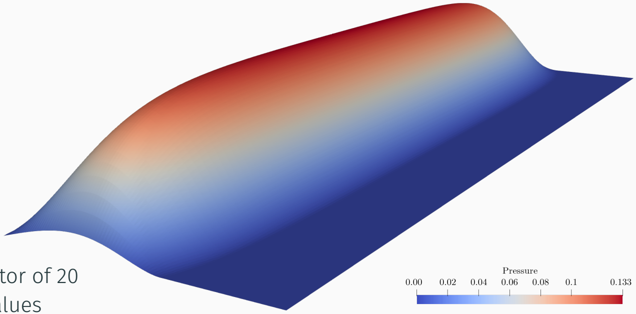
Benchmarks and Parameters

Journal Bearing (lubricant pressure distribution): P1 discretization of

$$\arg\min_{v \in K} \int_{\mathcal{D}} \left(\frac{1}{2} w_q(x) \|\nabla v(x)\|^2 - w_l(x) v(x) \right) dx,$$

$$K = \{ v \in H_0^1(\mathcal{D}) \colon\ v \geq 0 \}, \qquad \mathcal{D} = (0, 2\pi) \times (0, 2d),$$

where $w_q(x_1, x_2) = (1 + \epsilon \cos x_1)^3$, $w_l(x_1, x_2) = \epsilon \sin x_1$, $\epsilon = 0.1$ and $d = 10$.

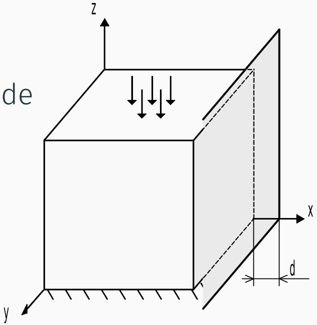


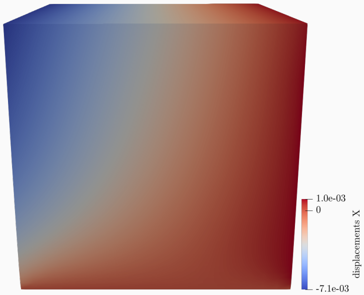
*pressure scaled by factor of 20
legend contains true values

Benchmarks and Parameters

3D Linear Elasticity Contact Problem (Primal):

- bottom fixed
- pushed from above
- obstacle close to the right side
- Q1 discretization





Solver:

- rtol $1\mathrm{e}{-10}$
- $\Gamma = 1$
- PERMON implementation
  - high-performance parallel library for solution of QPs
  - based on PETSc
- LUMI supercomputer, single core: AMD EPYC 7763 @ 2.45 GHz, Cray clang 16 with -O3

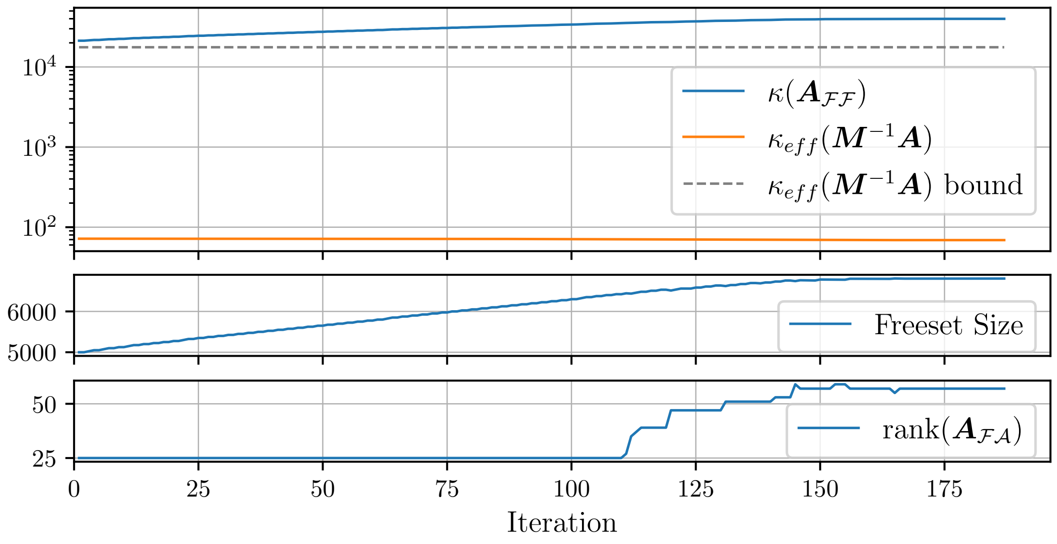
3D Cube Contact Problem with 20x40x80 Finite Elements (209,223 DoFs)

Method	Type	Precond.	**Hess.**	CG	Exp.	Prop.	Time [s]	S
MPPCG	None	None	**2766**	2269	244	8	37.93	**1.00**
MPPCG	Face	Cholesky	**14**	6	1	5	212.37	**0.18**
MPPCG	Approx	Cholesky	**57**	43	3	7	30.19	**1.26**
MPPCG	Face	ICC	**344**	212	60	11	72.38	**0.52**
MPPCG	Approx	ICC	**473**	297	84	7	10.38	**3.65**
MPPCG	Face	SSOR	**696**	439	125	6	82.46	**0.46**
MPPCG	Approx	SSOR	**715**	443	132	7	22.55	**1.68**

Journal Bearing Problem with 400x25 Discretization Points (10,000 DoFs)

Method	Type	Precond.	**Hess.**	CG	Exp.	Prop.	Time [s]	S
MPPCG	None	None	**2348**	2218	25	79	0.27	**1.00**
MPPCG	Face	Cholesky	**157**	78	0	78	1.95	**0.14**
MPPCG	Approx	Cholesky	**197**	97	10	79	0.91	**0.29**
MPPCG	Face	ICC	**179**	100	0	78	0.17	**1.59**
MPPCG	Approx	ICC	**208**	87	19	82	0.04	**7.28**
MPPCG	Face	SSOR	**748**	623	22	80	0.60	**0.44**
MPPCG	Approx	SSOR	**858**	699	38	82	0.19	**1.42**

MPPCG, Approx, Cholesky – Journal Bearing – 400x25 Grid Points (10,000 DoFs)



Journal Bearing Problem with 1600x100 Discretization Points (160,000 DoFs)

Method	Type	Precond.	**Hess.**	CG	Exp.	Prop.	Time [s]	S_b
MPPCG	None	None	**25166**	21632	1509	515	40.40	**1.00**
MPPCG	Face	Cholesky	**617**	308	0	308	317.43	**0.13**
MPPCG	Approx	Cholesky	**887**	379	93	321	59.38	**0.68**
MPPCG	Face	ICC	**776**	368	42	323	11.15	**3.62**
MPPCG	Approx	ICC	**1976**	238	564	609	4.47	**9.04**
MPPCG	Face	SSOR	**9609**	6194	1346	722	113.18	**0.36**
MPPCG	Approx	SSOR	**11661**	6902	1982	794	35.32	**1.14**

Conclusion and Outlook

- PERMON - open-source scalable library for convex QP
- Black box preconditioning of MPGP-type methods
- MPPCG with the approx. precond. in face gives large speedups
- Analysis of the error between preconditioning in face and its approx. variant
- Sharp bound on the condition number of the preconditioned operator

Outlook:

- Apply preconditioned MPGP-type methods to QP problems where good preconditioners for unconstrained problems are known (e.g., FETI)
- Optimal precond. for partially constrained problems

https://arxiv.org/abs/2507.00617

Thank you for your attention!

Any questions?

Jakub Kruzik, David Horak
jakub.kruzik@ugn.cas.cz

PERMON

https://permon.vsb.cz
https://github.com/permon

Czech Academy of Sciences
ÚGN
VSB TECHNICAL UNIVERSITY OF OSTRAVA | FACULTY OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE | DEPARTMENT OF APPLIED MATHEMATICS