Active Set Expansion Improvements for MPGP Algorithm

Jakub Kruzik
jakub.kruzik@vsb.cz

15 January 2024
UGN, Ostrava

Problem definition

$$\operatorname*{argmin}_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{s.t.} \quad x \in \Omega$$

$$\operatorname*{argmin}_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{b} \quad \text{s.t.} \quad \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u},$$

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

Active/Free Set and Gradient Splitting

Active/Free set:

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

Active/Free Set and Gradient Splitting

Active/Free set:

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

Gradient splitting:

$$\boldsymbol{g} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}$$

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

Active/Free Set and Gradient Splitting

Active/Free set:

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

Gradient splitting:

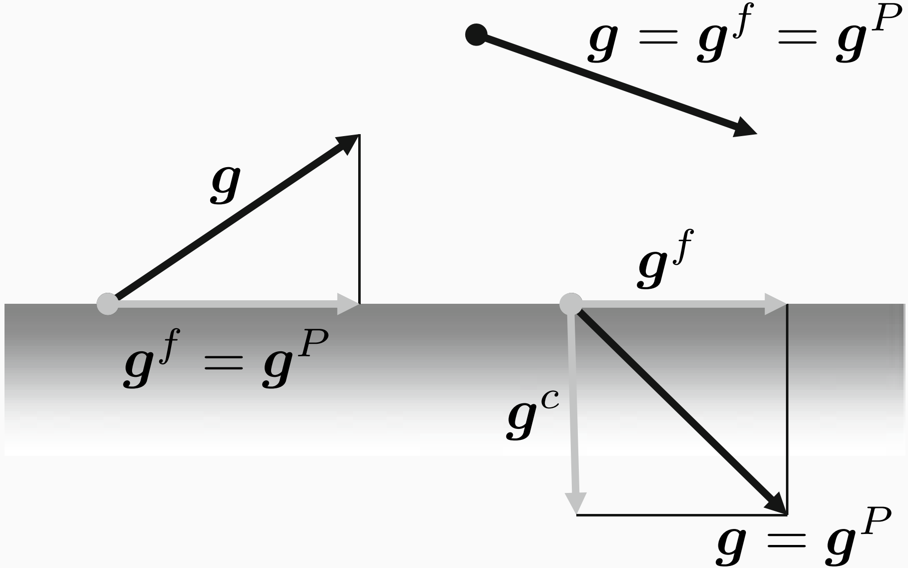
$$\boldsymbol{g} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}$$

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

Projected gradient:

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

Gradient Splitting - Example



Keeping x Feasible

Projection onto the feasible set Ω:

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

MPRGP - Modified Proportioning with Reduced Gradient Pojections

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

      CG step

    else:

      Expansion step

  else:

    Proportioning step

  $k = k + 1$

Output: $\boldsymbol{x}_k$

MPRGP - Modified Proportioning with Reduced Gradient Pojections

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

      CG step

    else:

      Expansion step

  else:

    Proportioning step

  $k = k + 1$

Output: $\boldsymbol{x}_k$

CG step:

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

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

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

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

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

   CG step

  else:

   Expansion step

 else:

  Proportioning step

 $k = k + 1$

Output: $\boldsymbol{x}_k$

CG step:

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

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

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

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

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 Pojections

```
Input: A, x_0 ∈ Ω, b, Γ > 0, ᾱ ∈ (0, 2||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}:
            CG step
        else:
            Expansion step
    else:
        Proportioning step
    k = k + 1
Output: x_k
```

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

            CG step

        else:

            Expansion step

    else:

        Proportioning step

    $k = k + 1$

Output: $\boldsymbol{x}_k$

Proportioning step:

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

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

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

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

MPRGP Operation Count

Step	Hess. mult.	Dot prod.	Vec. update	Grad. split.
CG	1	2	3	1
Expansion: Fixed	2	1	5	2
Proportioning	1	1	3	1

MPRGP R-linear Convergence

[Z. Dostál, J. Schöberl, 2005]

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

$$f(\boldsymbol{x}_{k+1}) - f(\widehat{\boldsymbol{x}}) \leq \eta \left(f(\boldsymbol{x}_k) - f(\widehat{\boldsymbol{x}}) \right),$$

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

MPRGP with Projected CG Step Expansion

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

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

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

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

    Projected CG

  else:

    Proportioning step

  $k = k + 1$

Output: $\boldsymbol{x}_k$

Projected CG:

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

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

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

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

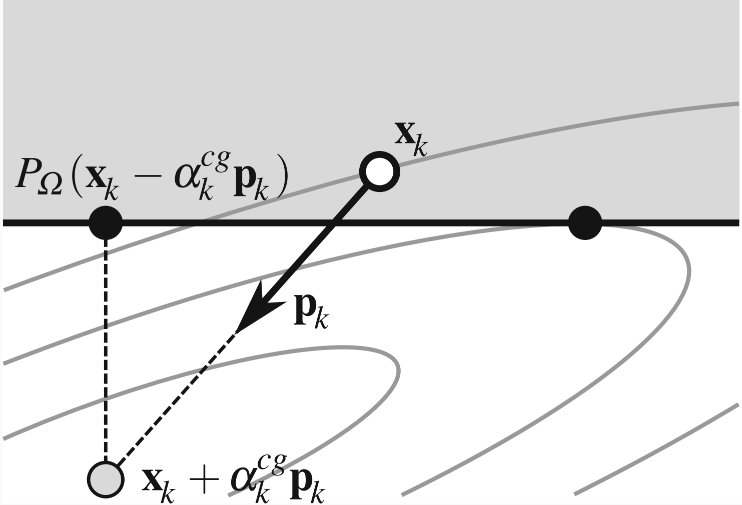
else:

  $\boldsymbol{x}_{k+1} = P_\Omega(\boldsymbol{x}_{k+1})$

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

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

Projected CG and Convergence



$$f(P_\Omega(\boldsymbol{x}_k - \alpha_k^{cg}\boldsymbol{p}_k)) > f(\boldsymbol{x}_k)$$

[Z. Dostál, 2009]

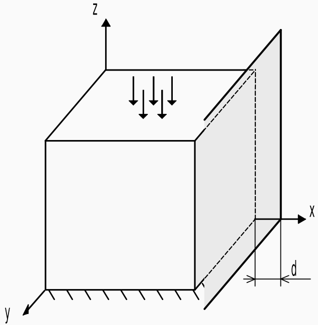
Benchmarks and Parameters

- $\Gamma = 1$
- $\overline{\alpha} = \alpha_u ||MA||^{-1}$, $\alpha_u = 1.9$
- $||\boldsymbol{A}||$ estimated by power method
  - 5 – 50 iterations (10^{-4} relative error)
  - not included in results (unless stated otherwise)

Benchmarks and Parameters

- $\Gamma = 1$
- $\overline{\alpha} = \alpha_u ||MA||^{-1}$, $\alpha_u = 1.9$
- $||\boldsymbol{A}||$ estimated by power method
  - 5 – 50 iterations (10^{-4} relative error)
  - not included in results (unless stated otherwise)

3D Linear Elasticity Contact Problem



- FETI dual formulation:

$$\operatorname*{argmin}_{\boldsymbol{\lambda}} \frac{1}{2}\boldsymbol{\lambda}^T \boldsymbol{F}\boldsymbol{\lambda} - \boldsymbol{\lambda}^T \boldsymbol{d} \text{ s.t. } \boldsymbol{\lambda}_I \geq \boldsymbol{o} \text{ and } \boldsymbol{G}\boldsymbol{\lambda} = \boldsymbol{e}$$

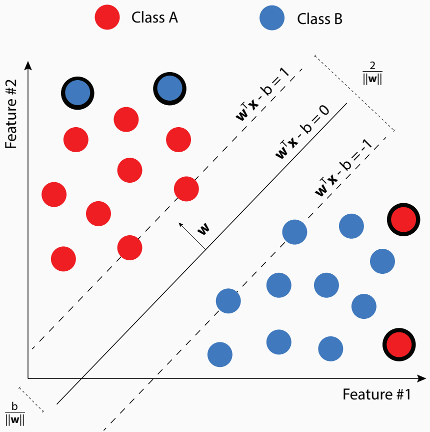
- SMALBE outer solver
- $81,812,703$ (undecomposed) degrees of freedom over $1,000$ subdomains
- relative tolerance 10^{-6}

Benchmarks and Parameters - Support Vector Machines (SVMs) Classification

$$\underset{\boldsymbol{\lambda}}{\text{argmin}} \ \frac{1}{2}\boldsymbol{\lambda}^T\boldsymbol{Y}\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{Y}\boldsymbol{\lambda} - \boldsymbol{\lambda}^T\boldsymbol{e} \quad \text{s.t.} \quad \boldsymbol{o} \leq \boldsymbol{\lambda} \leq C\boldsymbol{e}$$

- $l1$-loss function, dual formulation
- LIBSVM datasets
- relative tolerance 10^{-1}

Dataset	# samples	# features
Australian	690	14
Diabetes	678	8
Ionosphere	351	34



Benchmarks and Parameters - PERMON examples

Ex1: Centred finite difference discretization of

$$-u''(x) = -15, \qquad x \in [0,1]$$
$$s.t. \quad u(x) \geq \frac{sin(4\pi x - \frac{\pi}{6})}{2} - 2, \qquad x \in [0,1]$$

Ex2: Same as ex1 except constrained only on $x \in \left[0, \frac{1}{2}\right]$

Journal Bearing (lubricant pressure distribution): P1 discretization of

$$\underset{v \in K}{\operatorname{argmin}} \int_{\mathcal{D}} \left(\frac{1}{2} w_q(x) \|\nabla v(x)\|^2 - w_l(x) v(x) \right) dx,$$
$$K = \{v \in H_0^1(\mathcal{D})\colon\ v \geq 0\}, \qquad \mathcal{D} = (0, 2\pi) \times (0, 2d),$$

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 $d = 10$.

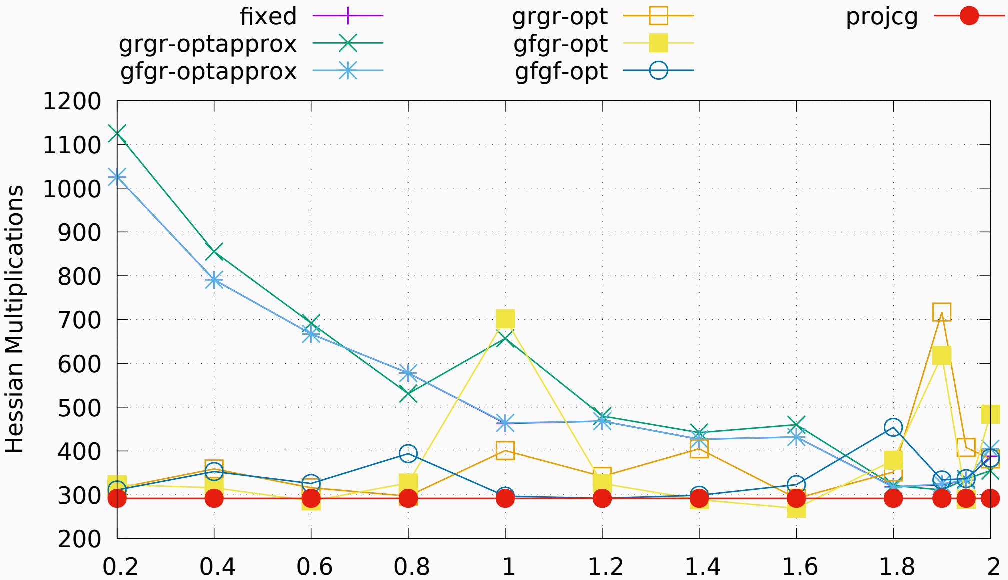
Benchmarks and Parameters - Generated BQP [Moré, Toraldo 1989]

- Size $n = 15000$
- $\kappa(\boldsymbol{A}) = 10^4$
- non-degenerate solution

The number of active variables at the solution:

- **BQP1**: 10%
- **BQP2**: 50%
- **BQP3**: 90%

3D contact - Choice of α_u

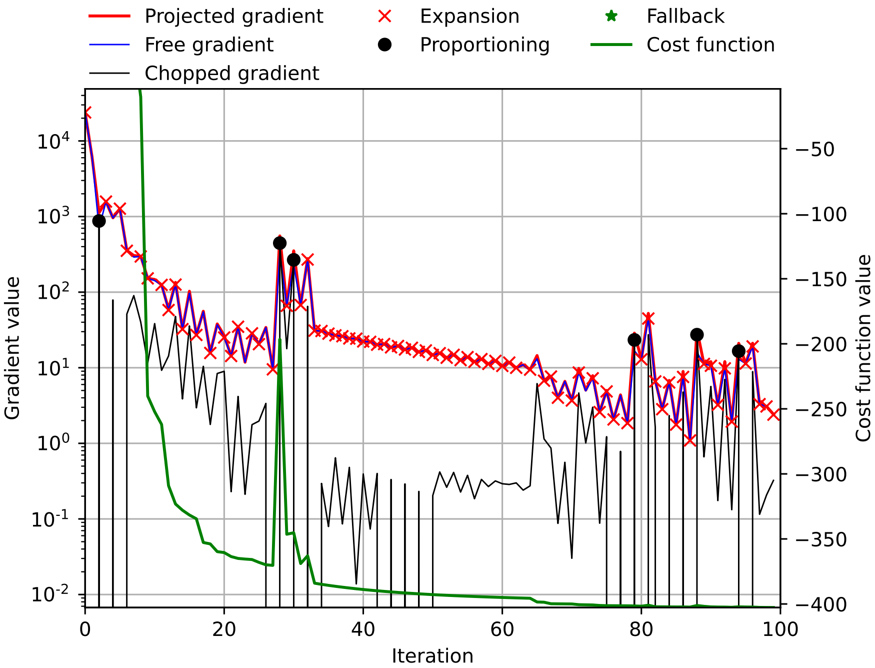


Results

Problem	Exp. type	Hess. mult.	CG	Exp.	Prop.
3D Contact	Fixed	323	144	83	3
	Projected CG	292	171	53	5
SVM: Australian	Fixed	195	8	92	2
	Projected CG	83	16	32	2
SVM: Diabetes	Fixed	630	2	313	1
	Projected CG	133	13	58	3
SVM: Ionosphere	Fixed	381	21	179	1
	Projected CG	125	14	54	2

SVM - tolerance $1e-4$

Dataset	Exp. Type	Hess. mult.	CG	Exp,.	Prop.	Cost inc.	Fall.
australian	fixed	4567	1134	1704	24	423	0
australian	projCG	3571	565	1455	95	127	0
diabetes	fixed	1108	124	491	1	18	0
diabetes	projCG	**1439**	113	627	71	109	0
ionosphere	fixed	628	149	237	4	8	0
ionosphere	projCG	265	104	78	4	6	0



Projected gradient
Free gradient
Chopped gradient
Expansion
Proportioning
Fallback
Cost function
Gradient value
Cost function value
Iteration

MPPCG with Fallback

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

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

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

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

    Projected CG with fallback

  else:

    Proportioning step

  $k = k + 1$

Output: $\boldsymbol{x}_k$

---

if $\alpha_k^{cg} \le \alpha^{feas}$:

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

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

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

else:

  $\boldsymbol{x}_{k+1} = P_\Omega(\boldsymbol{x}_{k+1})$

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

  if *fallback*:

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

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

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

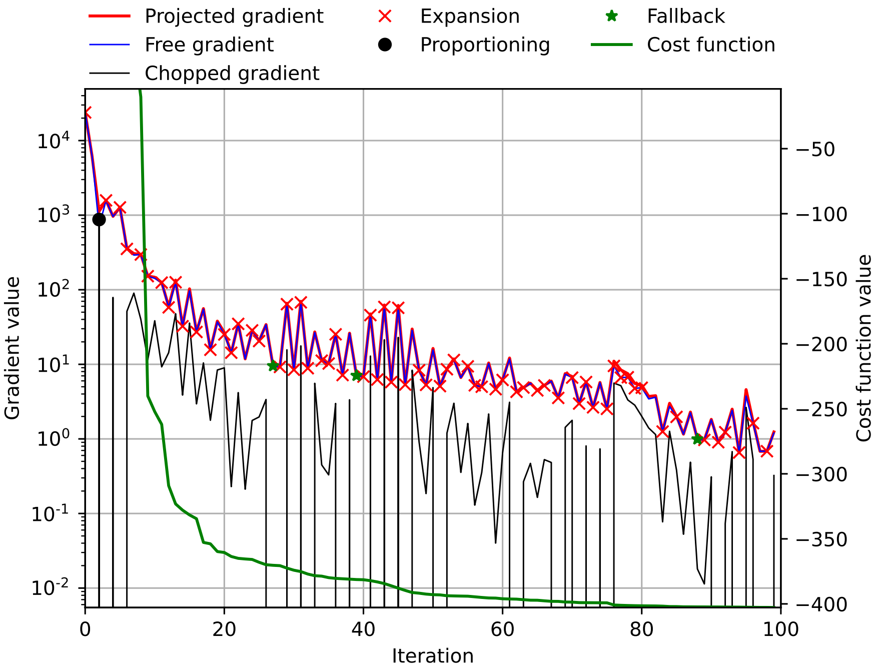
    $\boldsymbol{g}_{k+1} = \boldsymbol{A}\boldsymbol{x}_{k+1} - \boldsymbol{b}$

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

Fallback Variants

Fallback 1:

$$f(\boldsymbol{x}_{k+1}) > f(\boldsymbol{x}_k)$$



Projected gradient
Expansion
Fallback
Free gradient
Proportioning
Cost function
Chopped gradient
Gradient value
Cost function value
Iteration

SVM - tolerance $1e-4$

Dataset	Exp. Type	Hess. mult.	CG	Exp,.	Prop.	Cost inc.	Fall.
australian	fixed	4567	1134	1704	24	423	0
australian	projCG	3571	565	1455	95	127	0
australian	fallback 1	3298	1015	1014	36	218	218
diabetes	fixed	1108	124	491	1	18	0
diabetes	projCG	1439	113	627	71	109	0
diabetes	fallback 1	292	84	96	1	14	14
ionosphere	fixed	628	149	237	4	8	0
ionosphere	projCG	265	104	78	4	6	0
ionosphere	fallback 1	**320**	109	89	5	27	27

Fallback Variants

Fallback 1:

$$f(\boldsymbol{x}_{k+1}) > f(\boldsymbol{x}_k)$$

Fallback 2:

$$f(\boldsymbol{x}_{k+1}) > f(\boldsymbol{x}_k) \quad \wedge \quad ||\boldsymbol{g}^c_{k+1}|| > \Gamma||\boldsymbol{g}^f_{k+1}||$$

SVM - tolerance $1e-4$

Dataset	Exp. Type	Hess. Mult.	CG	Exp.	Prop.	Cost inc.	Fall.
australian	fixed	4567	1134	1704	24	423	0
australian	projCG	3571	565	1455	95	127	0
australian	fallback 1	3298	1015	1014	36	218	218
australian	fallback 2	2160	747	662	32	74	56
diabetes	fixed	1108	124	491	1	18	0
diabetes	projCG	1439	113	627	71	109	0
diabetes	fallback 1	292	84	96	1	14	14
diabetes	fallback 2	292	84	96	1	14	14
ionosphere	fixed	628	149	237	4	8	0
ionosphere	projCG	265	104	78	4	6	0
ionosphere	fallback 1	320	109	89	5	27	27
ionosphere	fallback 2	277	104	78	5	13	11

Spectral Projected Gradient (SPG) Method [Birgin et al. 2000]

Input: $\boldsymbol{x}^{(0)} \in \Omega,\ \delta, \sigma \in (0,1),\ M \in \mathbf{N},\ 0 < \alpha_{min} \leq \alpha_{max},\ \alpha_0 \in [\alpha_{min}, \alpha_{max}]$

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

  $\boldsymbol{d}^k = \Pi_\Omega\left(\boldsymbol{x}^k - \alpha_k g(\boldsymbol{x}^k)\right) - \boldsymbol{x}^k$

  $\nu_k = 1; \quad f_{ref} = \max\{f(\boldsymbol{x}^{k-i}),\ 0 \leq i \leq \min(k, M-1)\}$

  while $f(\boldsymbol{x}^k + \nu_k \boldsymbol{d}^k) > f_{ref} + \sigma \nu_k g(\boldsymbol{x}^k)^T \boldsymbol{d}^k$:

    $\nu_k = \delta \nu_k$;

  $\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \nu_k \boldsymbol{d}^k$

  define the step length $\alpha_k \in [\alpha_{min}, \alpha_{max}]$ by BoxVABB$_{\text{min}}$

  $k = k + 1$

Output: $\boldsymbol{x}^{k+1}$

BoxVABB$_{\text{min}}$ [Crisci et al. 2019]

$$\alpha_k^{\text{BB1}} = \frac{\| \boldsymbol{s}^{(k-1)} \|^2}{\boldsymbol{s}^{(k-1)\,T} \boldsymbol{y}^{(k-1)}}, \quad \alpha_k^{\text{BoxBB2}} = \frac{\boldsymbol{s}_{I_{k-1}}^{k-1\ T} \boldsymbol{y}_{I_{k-1}}^{k-1}}{\| \boldsymbol{y}_{I_{k-1}}^{k-1} \|^2},$$

where $\boldsymbol{s}^{k-1} = \boldsymbol{x}^k - \boldsymbol{x}^{k-1}$, $\boldsymbol{y}^k = \boldsymbol{g}^k - \boldsymbol{g}^{k-1}$ and

$$I_{k-1} = \{1, 2, \dots, n\} \setminus \{i \colon (x_i^{k-1} = \ell_i \wedge x_i^k = \ell_i) \vee (x_i^{k-1} = u_i \wedge x_i^k = u_i)\}$$

$\text{BoxVABB}_{\text{min}}$ [Crisci et al. 2019]

$$\alpha_k^{\text{BB1}} = \frac{\| \boldsymbol{s}^{(k-1)} \|^2}{\boldsymbol{s}^{(k-1)\,T} \boldsymbol{y}^{(k-1)}}, \qquad \alpha_k^{\text{BoxBB2}} = \frac{\boldsymbol{s}_{I_{k-1}}^{k-1\;T} \boldsymbol{y}_{I_{k-1}}^{k-1}}{\| \boldsymbol{y}_{I_{k-1}}^{k-1} \|^2},$$

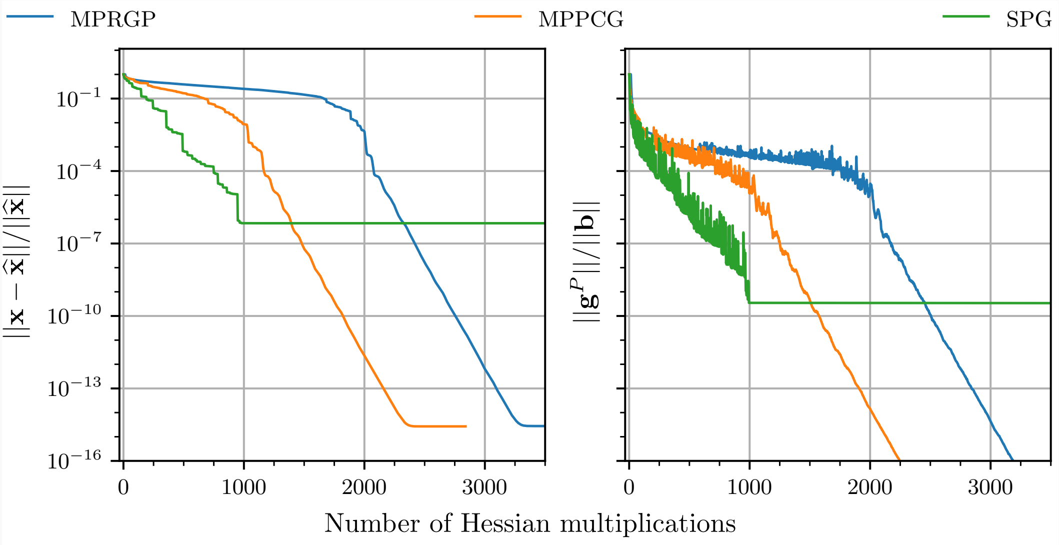
where $\boldsymbol{s}^{k-1} = \boldsymbol{x}^k - \boldsymbol{x}^{k-1}$, $\boldsymbol{y}^k = \boldsymbol{g}^k - \boldsymbol{g}^{k-1}$ and

$$I_{k-1} = \{1, 2, \dots, n\} \setminus \{i \colon (x_i^{k-1} = \ell_i \wedge x_i^k = \ell_i) \vee (x_i^{k-1} = u_i \wedge x_i^k = u_i)\}$$

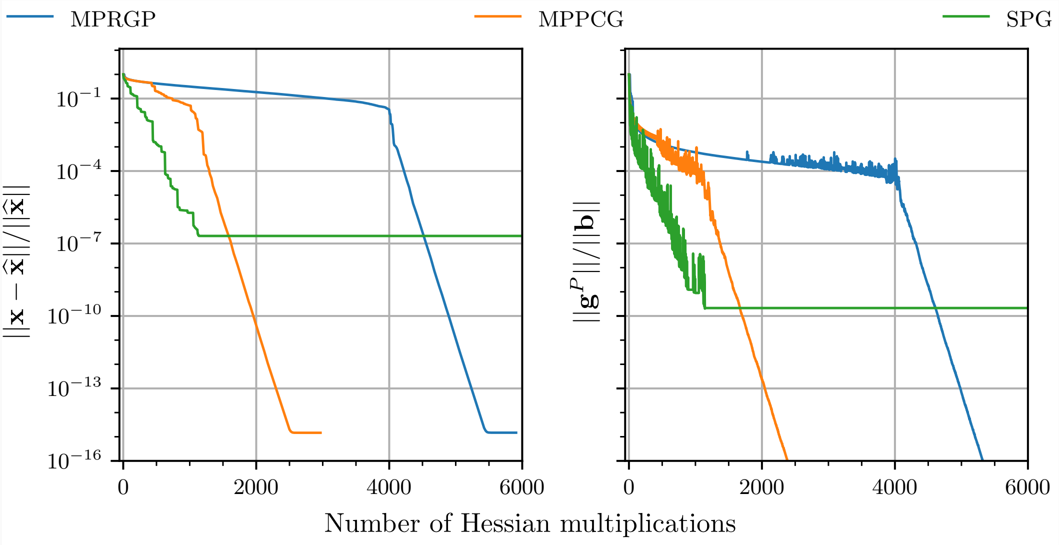
$$\tau_{k+1} = \begin{cases} \tau_k / \vartheta & \text{if} \quad \dfrac{\alpha_k^{\text{BoxBB2}}}{\alpha_k^{\text{BB1}}} < \tau_k \\ \tau_k \vartheta & \text{otherwise} \end{cases}$$

$$\alpha_k^{\text{BoxVABB}_{\text{min}}} = \begin{cases} \min \left\{ \alpha_j^{\text{BoxBB2}} \,:\, j = \max\{1, k - m_\alpha\}, \dots, k \right\} & \text{if} \quad \dfrac{\alpha_k^{\text{BoxBB2}}}{\alpha_k^{\text{BB1}}} < \tau_k \\ \alpha_k^{\text{BB1}} & \text{otherwise} \end{cases}$$

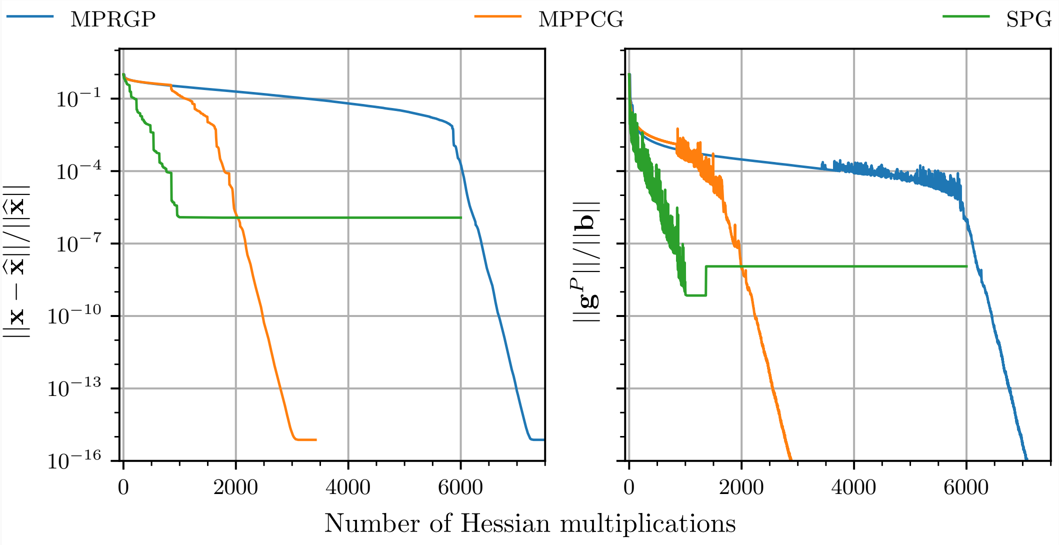
BQP1: Convergence curves for MPRGP, MPPCG and SPG.



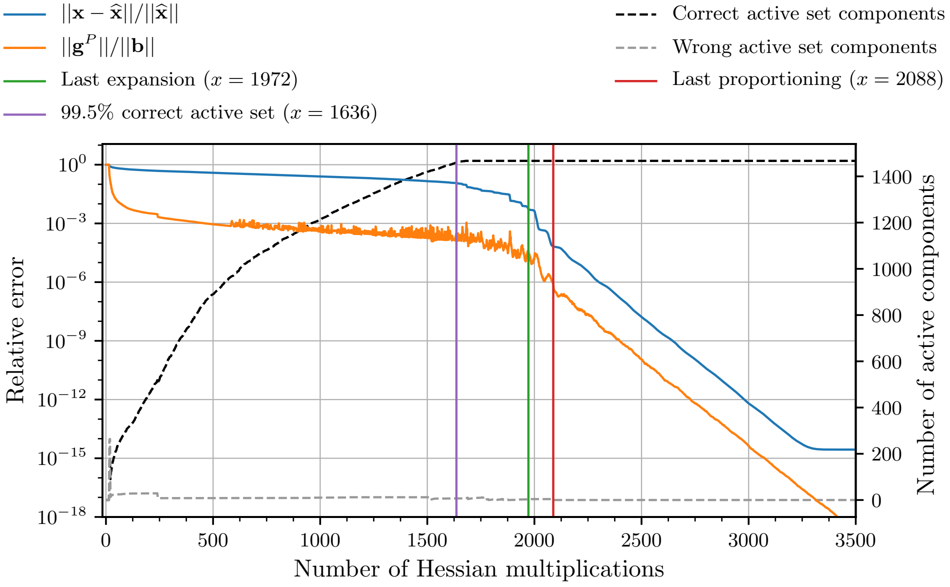
BQP2: Convergence curves for MPRGP, MPPCG and SPG.



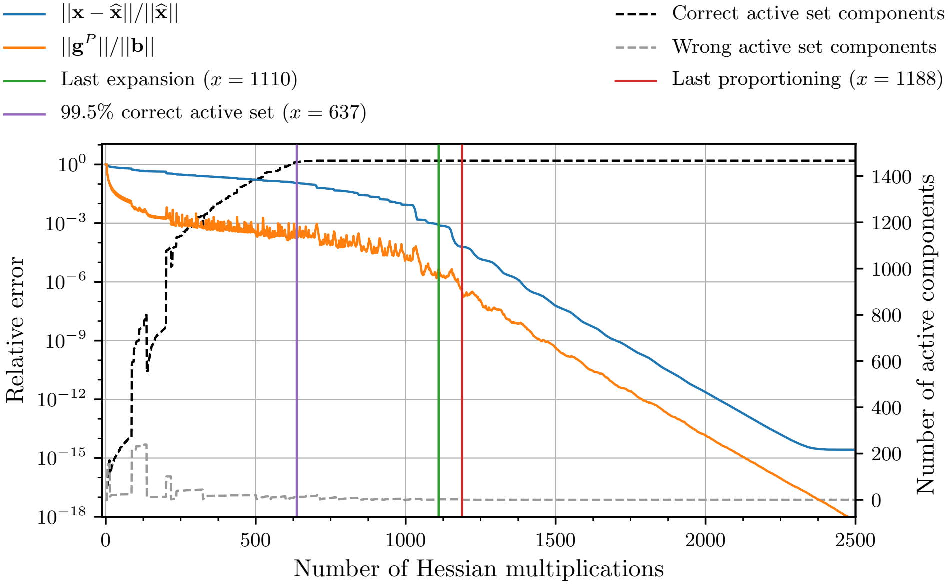
BQP3: Convergence curves for MPRGP, MPPCG and SPG.



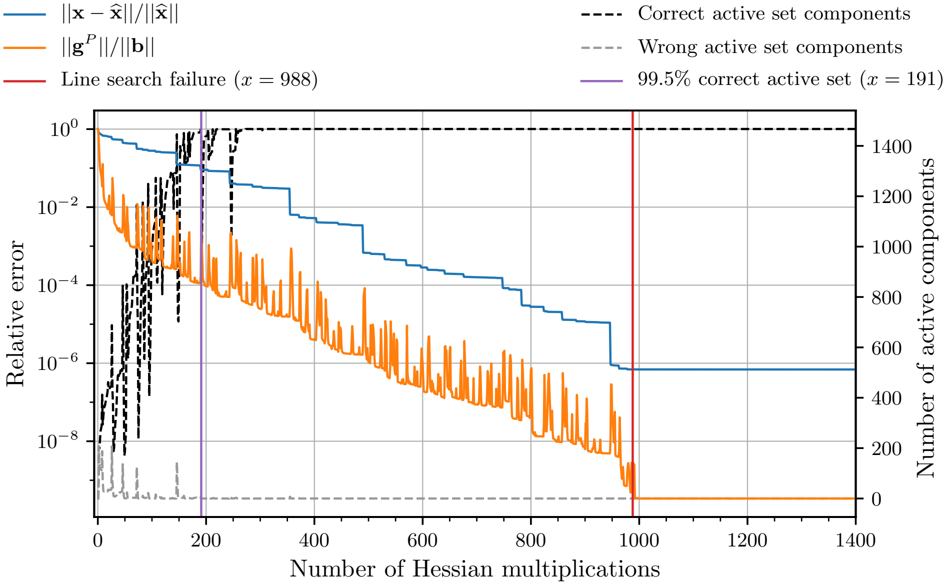
BQP1: Progress of MPRGP.



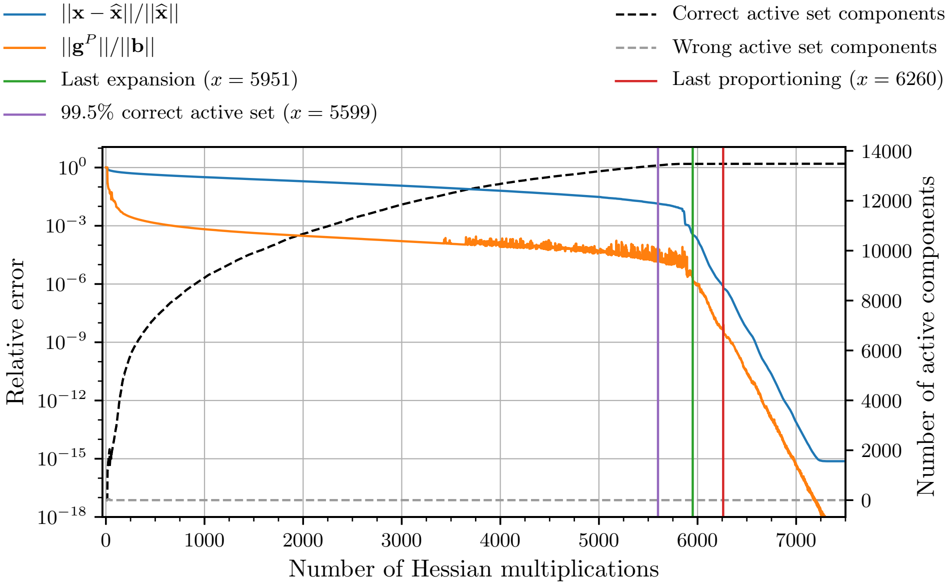
BQP1: Progress of MPPCG



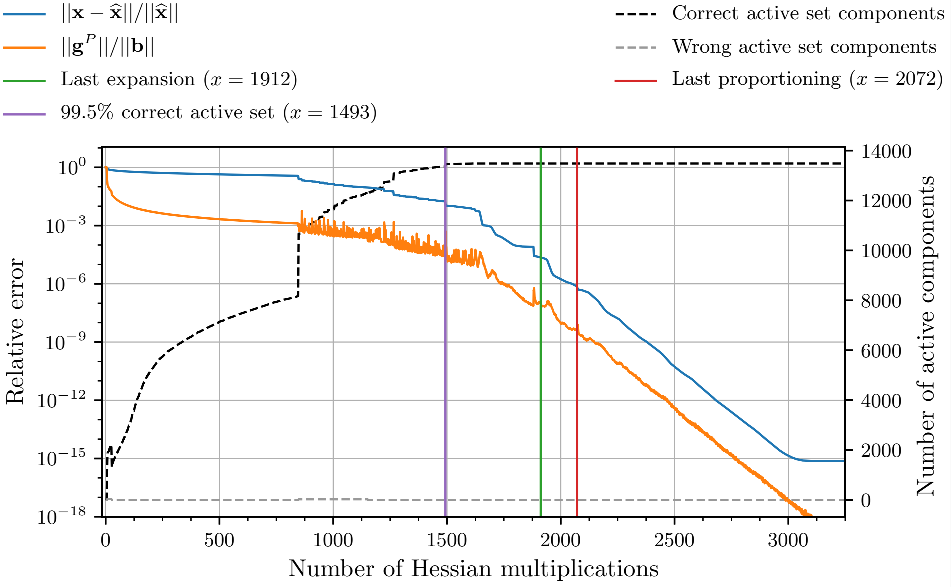
BQP1: Progress of SPG.



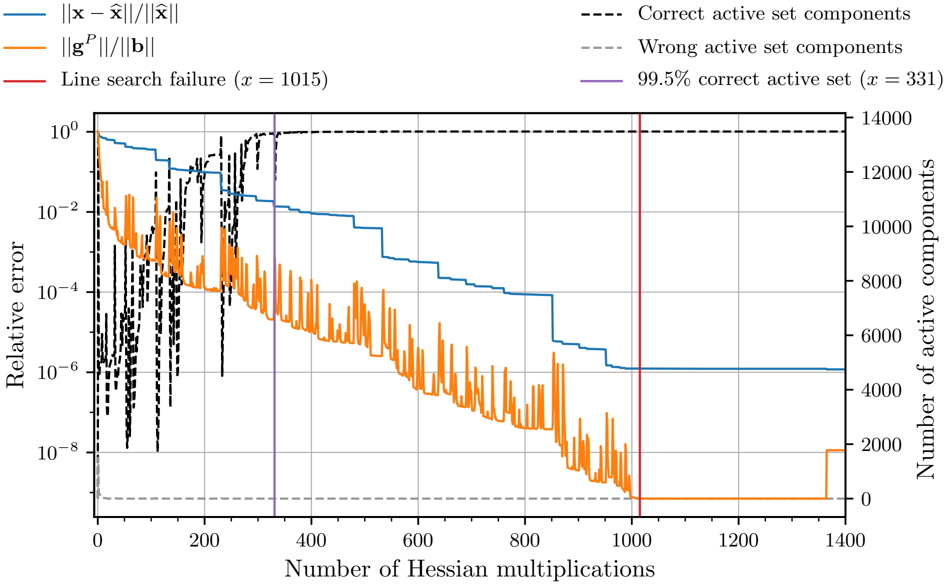
BQP3: Progress of MPRGP.



BQP3: Progress of MPPCG.



BQP3: Progress of SPG.



Spectral Projected Gradient (SPG) Expansion Step - MPSPG

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

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

define the step length $\alpha_k \in [\alpha_{min}, \alpha_{max}]$ by $\text{BoxVABB}_{\text{min}}$

$\boldsymbol{d}_k = P_\Omega \left(\boldsymbol{x}_{k+\frac{1}{2}} - \alpha_k \boldsymbol{g}_{k+\frac{1}{2}} \right) - \boldsymbol{x}_{k+\frac{1}{2}}$

$\nu_k = 1$

$f_{ref} = \max\{f(\boldsymbol{x}_{k+\frac{1}{2}}), f(\boldsymbol{x}_{k-i}),\ 0 \leq i \leq \min(k, M-1)\}$

while $f(\boldsymbol{x}_{k+\frac{1}{2}} + \nu_k \boldsymbol{d}_k) > f_{ref} + \sigma \nu_k (\boldsymbol{g}_{k+\frac{1}{2}})^T \boldsymbol{d}_k$:

    $\nu_k = \delta \nu_k$

$\boldsymbol{x}_{k+1} = \boldsymbol{x}_{k+\frac{1}{2}} + \nu_k \boldsymbol{d}_k$

$\boldsymbol{g}_{k+1} = \boldsymbol{g}_{k+\frac{1}{2}} + \nu_k \boldsymbol{A} \boldsymbol{d}_k$

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

Results - BQP

Problem	notes	MPRGP	MPPCG	SPG	MPSPG	MPSPGf
BQP1	rtol: 10^{-4}	1726	714	208	262	219 (217)
BQP2	rtol: 10^{-4}	3174	799	205	273	266 (265)
BQP3	rtol: 10^{-4}	3910	1268	248	401	234 (231)
BQP1	rtol: 10^{-6}	2075	1175	579	745	592 (590)
BQP2	rtol: 10^{-6}	4299	1230	522	744	726 (465)
BQP3	rtol: 10^{-6}	6081	1735	604	1224	692 (688)
GM of BQP speedups		1	2.94	9.19	6.25	8.12 (5.54)

Results - PERMON Examples

Problem	notes	MPRGP	MPPCG	SPG	MPSPG	MPSPGf
ex1	grid: 100	177	164	249	157	167 (165)
ex1	grid: 1000	3245	3037	2709	3946	2918 (2882)
ex1	grid: 5000	31657	25673	19806	74756	22541 (22232)
ex2	grid: 100	208	200	369	158	181 (178)
ex2	grid: 1000	2825	3366	3077	2993	2361 (2348)
ex2	grid: 5000	21525	16103	24505	20227	20767 (20543)
jbearing	grid: 50x50	154	142	154	135	139 (129)
jbearing	grid: 100x100	318	335	358	346	322 (307)
jbearing	grid: 200x50	663	664	740	803	690 (659)
jbearing	grid: 400x25	1463	1559	1597	1700	1443 (1383)
GM PERMON ex. speedups		1	1.05	0.92	0.91	1.10 (0.74)
Overall GM of speedups		1	1.54	2.19	1.87	2.32 (1.57)

Results - BQP, rtol 10^{-12}

Problem	MPRGP	MPPCG	MPSPG	MPSPGf
BQP1	2750	1787	1422	1274 (1272)
BQP2	4951	1932	1402	1249 (1248)
BQP3	6797	2463	1923	1360 (1355)
GM of BQP speedups	1	2.22	2.89	3.50 (2.33)

Conclusion and Outlook

- Unconstrained QP minimization, PCDEFLATION (a general multilevel deflation preconditioner)
- Active set methods, including MPRGP and SPG, active set expansion strategies and preconditioning for MPRGP
- Equality-constrained QP, such as augmented Lagrangian-type methods like SMALE and ALAPC
- Primal-dual interior point methods
- FETI-type domain decomposition methods
- Implementations and PERMON library
- Benchmarks and applications

Thank you for your attention!

Any questions?

Jakub Kruzik

jakub.kruzik@vsb.cz

PERMON

http://permon.vsb.cz