An Efficient Inexact Accelerated Block Coordinate Descent Method for Least Squares Semidefinite Programming

Defeng Sun
Joint work with Kim-Chuan Toh and Liuqin Yang

Department of Mathematics, National University of Singapore

Outline

1. SDP and least squares SDP
2. Main ingredients
   - A Danskin-type theorem
   - Inexact APG
   - Inexact block symmetric Gauss-Seidel iteration with a non-smooth block
3. Inexact accelerated block coordinate gradient descent method for composite problem with 2 non-smooth terms and a multi-block coupled smooth term
4. Inexact accelerated block coordinate descent (ABCD) method for dual SDP
5. Numerical experiments for LSSDP

SDP with an additional polyhedral set and inequalities:

$$\begin{aligned} \min \quad & \langle C,\, X\rangle \\ \text{s.t.} \quad & \mathcal{A}_E(X) = b_E,\ \mathcal{A}_I X - s = 0,\ X \in \mathcal{S}^n_+,\ X \in \mathcal{P}, s \in \mathcal{K} \end{aligned}$$

$$\mathcal{P} = \{W \in \mathcal{S}^n :\ L \leq W \leq U\},\ \mathcal{K} = \{w \in \Re^{m_I} :\ l \leq w \leq u\}.$$

Applying a proximal point algorithm (PPA) to solve above SDP:

$$\begin{aligned} (X^{k+1}, s^{k+1}) = \arg\min \quad & \langle C,\, X\rangle + \tfrac{1}{2\sigma_k}(\|X - X^k\|^2 + \|s - s^k\|^2) \\ \text{s.t.} \quad & \mathcal{A}_E(X) = b_E,\ \mathcal{A}_I X - s = 0,\ X \in \mathcal{S}^n_+, \\ & X \in \mathcal{P},\ s \in \mathcal{K}. \end{aligned}$$

Least squares semidefinite programming (LSSDP)

LSSDP includes PPA subproblem as a particular case: Given G, g,

$$
\begin{array}{lll}
\textbf{(P)} & \min & \frac{1}{2}\|X-G\|^2 + \frac{1}{2}\|s-g\|^2 \\
 & \text{s.t.} & \mathcal{A}_E(X) = b_E,\ \mathcal{A}_I X - s = 0,\ X \in \mathcal{S}^n_+,\ X \in \mathcal{P},\ s \in \mathcal{K}.
\end{array}
$$

The dual of **(P)** is given by

$$
\begin{aligned}
\textbf{(D)}\ \min\quad & \boldsymbol{F(Z, v, S, y_E, y_I)} \\
:= {} & \delta^*_{\mathcal{P}}(-Z) + \delta^*_{\mathcal{K}}(-v) + \delta_{\mathcal{S}^n_+}(S) \\
& -\langle b_E,\, y_E\rangle + \frac{1}{2}\|\mathcal{A}^*_E y_E + \mathcal{A}^*_I y_I + S + Z + G\|^2 + \frac{1}{2}\|v - y_I + g\|^2 \\
& \qquad + \text{constant}
\end{aligned}
$$

$\delta_{\mathcal{C}}(\cdot)$ = indicator function over $\mathcal{C}$; $\delta_{\mathcal{C}}(u) = 0$ if $u \in \mathcal{C}$; ∞ otherwise

$\delta^*_{\mathcal{C}}(\cdot)$ is the conjugate function of $\delta_{\mathcal{C}}$ defined by

$$\delta^*_{\mathcal{C}}(\cdot) = \sup_{W \in \mathcal{C}} \langle \cdot,\, W\rangle.$$

Existing first-order methods for (**D**)

- Block coordinate descent (BCD) type method [Luo,Tseng,...] with iteration complexity of $O(1/k)$.
- Accelerated proximal gradient (APG) method [Nesterov, Beck-Teboulle] with iteration complexity of $O(1/k^2)$.
- Accelerated randomized BCD-type method [Beck, Nesterov, Richtarik,...] with iteration complexity of $O(1/k^2)$.

Elimination of a block via a Danskin-type theorem

Consider block vectors $x = (x_1, x_2, \ldots, x_s) \in \mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2 \cdots \times \mathcal{X}_s$, and

$$\min\{p(x_1) + \varphi(z) + \phi(z, x) \mid z \in \mathcal{Z},\ x \in \mathcal{X}\}$$
$$= \boxed{\min\{p(x_1) + f(x) \mid x \in \mathcal{X}\}}$$

where $p(\cdot)$, $\varphi(\cdot)$ are convex functions (possibly nonsmooth), and

$$f(x) = \min\{\varphi(z) + \phi(z, x) \mid z \in \mathcal{Z}\}$$
$$z(x) = \operatorname{argmin}\{\ldots\}$$

Assume that φ, ϕ satisfy the conditions in the next theorem, then f has Lipschitz continuous gradient $\nabla f(x) = \nabla_x \phi(z(x), x)$.

A Danskin-type theorem

$\varphi : \mathcal{Z} \to (-\infty, \infty]$ is a closed proper convex function;
$\phi(\cdot, \cdot) : \mathcal{Z} \times \mathcal{X} \to \Re$ is a convex function;
$\phi(z, \cdot) : \Omega \to \Re$ is continuously differentiable on Ω for each z;
$\nabla_x \phi(z, x)$ is continuous on $\mathrm{dom}(\varphi) \times \Omega$.
Consider $f : \Omega \to [-\infty, +\infty)$ defined by

$$f(x) = \inf_{z \in \mathcal{Z}} \{\varphi(z) + \phi(z, x)\}, \quad x \in \Omega. \tag{1}$$

Condition: The minimizer $z(x)$ is unique for each x and is bounded on a compact set.

A Danskin-type theorem

Theorem 1

(i) *If* $\exists$ *an open neighborhood* $\mathcal{N}_x$ *of* x *such that* $z(\cdot)$ *is bounded on any compact subset of* $\mathcal{N}_x$*, then the convex function* f *is differentiable on* $\mathcal{N}_x$ *and*

$$\nabla f(x') = \nabla_x \phi(z(x'), x') \quad \forall\, x' \in \mathcal{N}_x.$$

(ii) *Suppose that* $z(\cdot)$ *is bounded on any nonempty compact subset of* $\mathcal{Z}$*. Assume that for any* $z \in \mathrm{dom}(\varphi)$*,* $\nabla_x \phi(z, \cdot)$ *is Lipschitz continuous on* $\mathcal{Z}$ *and* $\exists\ \Sigma \succeq 0$ *such that for all* $x \in \mathcal{X}$ *and* $z \in \mathrm{dom}(\varphi)$*,*

$$\Sigma \succeq \mathcal{H} \quad \forall\, \mathcal{H} \in \partial^2_{xx} \phi(z, x).$$

Then, $\nabla f(\cdot)$ *is Lipschitz continuous on* $\mathcal{X}$ *with the Lipschitz constant* $||\Sigma\|_2$ *(the spectral norm of* Σ*) and for any* $x \in \mathcal{X}$*,*

$$\Sigma \succeq \mathcal{G} \quad \forall\, \mathcal{G} \in \partial^2_{xx} f(x),$$

where $\partial^2_{xx} f(x)$ *denotes the generalized Hessian of* f *at* x*.*

An inexact APG (accelerated proximal gradient)

Consider

$$\min\{F(x) := p(x) + f(x) \mid x \in \mathcal{X}\}$$

with $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \quad \forall\, x, y \in \mathcal{X}$.

Algorithm. Input $y^1 = x^0 \in \mathrm{dom}(p)$, $t_1 = 1$. Iterate

1. Find an approximate minimizer

$$x^k \approx \arg\min_{y \in \mathcal{X}} \Big\{ p(y) + f(y^k) + \langle \nabla f(y^k),\, y - y^k \rangle + \frac{1}{2} \langle y - y^k,\, \mathcal{H}_k (y - y^k) \rangle \Big\}$$

   where $\mathcal{H}_k \succ 0$ is an a priori given linear operator.

2. Compute $t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$, $y^{k+1} = x^k + \left(\frac{t_k - 1}{t_{k+1}}\right)(x^k - x^{k-1})$.

An inexact APG

Consider the following admissible conditions

$$F(x^k) \;\le\; p(x^k) + f(y^k) + \langle \nabla f(y^k),\, x^k - y^k\rangle + \tfrac{1}{2}\langle x^k - y^k,\, \mathcal{H}_k(x^k - y^k)\rangle$$

$$\nabla f(y^k) + \mathcal{H}_j(x^k - y^k) + \gamma^k =: \delta^k \quad \text{with } \|\mathcal{H}_k^{-1/2}\delta^k\| \le \frac{\epsilon_k}{\sqrt{2}t_k}$$

where $\gamma^k \in \partial p(x^k)$ = the set of subgradients of p at x^k, $\{\epsilon_k\}$ is a nonnegative summable sequence. Note $t_k \approx k/2$ for k large.

Theorem 2 (Jiang-Sun-Toh)

Suppose the above conditions hold and $\mathcal{H}_{k-1} \succeq \mathcal{H}_k \succ 0$ for all k. Then

$$0 \le F(x^k) - F(x^*) \le \frac{4}{(k+1)^2}\left(\sqrt{\tau} + \bar{\epsilon}_k\right)^2$$

where $\tau = \frac{1}{2}\|x^0 - x^*\|^2_{\mathcal{H}_1}$, $\bar{\epsilon}_k = \sum_{j=1}^{k} \epsilon_j$.

An inexact APG

Apply inexact APG to

$$\min\{F(x) := p(x_1) + f(x) \mid x \in \mathcal{X}\}.$$

Since $\nabla f(\cdot)$ is Lipschitz continuous, $\exists$ an symmetric and PSD linear operator $\mathcal{Q} : \mathcal{X} \to \mathcal{X}$ such that

$$\mathcal{Q} \succeq \mathcal{M}, \quad \forall\, \mathcal{M} \in \partial^2 f(x),\ \forall x \in \mathcal{X}$$

and $\mathcal{Q}_{ii} \succ 0$ for all i.
Given y^k, we have for all $x \in \mathcal{X}$

$$f(x) \leq q_k(x) := f(y^k) + \langle \nabla f(y^k),\, x - y^k \rangle + \frac{1}{2}\langle x - y^k,\, \mathcal{Q}(x - y^k)\rangle.$$

APG subproblem: need to solve a nonsmooth QP of the form

$$\min_{x \in \mathcal{X}}\{p(x_1) + q_k(x)\}, \quad x = (x_1, x_2, \ldots, x_s)$$

which is not easy to solve!
Idea: add an additional proximal term to make it easier!

An inexact block symmetric Gauss-Seidel (SGS) iteration

Given positive semidefinite linear operator $\mathcal{Q}$ such that

$$\mathcal{Q}x \equiv \begin{pmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12} & \cdots & \mathcal{Q}_{1s} \\ \mathcal{Q}_{12}^* & \mathcal{Q}_{22} & \cdots & \mathcal{Q}_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{Q}_{1s}^* & \mathcal{Q}_{2s}^* & \cdots & \mathcal{Q}_{ss} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_s \end{pmatrix}$$

where $\mathcal{Q}_{ii} \succ 0$. Consider the following block decomposition:

$$\mathcal{U}x \equiv \begin{pmatrix} 0 & \mathcal{Q}_{12} & \cdots & \mathcal{Q}_{1s} \\ & \ddots & & \vdots \\ & & \ddots & \mathcal{Q}_{s-1,s} \\ & & & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_s \end{pmatrix}$$

Then $\mathcal{Q} = \mathcal{U}^* + \mathcal{D} + \mathcal{U}$, where $\mathcal{D}x = (\mathcal{Q}_{11}x_1, \ldots, \mathcal{Q}_{ss}x_s)$.

An inexact block symmetric Gauss-Seidel (sGS) iteration

Consider the convex quadratic function:

$$q(x) := \frac{1}{2}\langle x,\, \mathcal{Q}x\rangle - \langle r,\, x\rangle, \quad x = (x_1, \ldots, x_s) \in \mathcal{X}.$$

Let $p : \mathcal{X}_1 \to (-\infty, +\infty]$ be a given closed proper convex function. Define

$$\mathcal{T} \;:=\; \mathcal{U}\mathcal{D}^{-1}\mathcal{U}^*$$

Let $y \in \mathcal{X}$ be given. Define

$$x^+ := \operatorname*{arg\,min}_{x\in\mathcal{X}} \left\{ p(x_1) + q(x) + \frac{1}{2}\|x - y\|_{\mathcal{T}}^2 \right\}. \tag{2}$$

The quadratic term has $\mathcal{H} := \mathcal{Q} + \mathcal{T} = (\mathcal{D} + \mathcal{U})\mathcal{D}^{-1}(\mathcal{D} + \mathcal{U}^*) \succ 0$.
(2) is easier to solve!

An inexact block symmetric Gauss-Seidel (sGS) iteration

Theorem 3 (Li-Sun-Toh)

Given y. For $i = s, \ldots, 2$, define

$$\begin{aligned} \widehat{x}_i &:= \arg\min_{x_i} \{ p(y_1) + q(y_{\le i-1}, x_i, \widehat{x}_{\ge i+1}) - \langle \widehat{\delta}_i, x_i \rangle \} \\ &= \mathcal{Q}_{ii}^{-1} \big(r_i + \widehat{\delta}_i - \textstyle\sum_{j=1}^{i-1} \mathcal{Q}_{ji}^* y_j - \sum_{j=i+1}^{s} \mathcal{Q}_{ij} \widehat{x}_j \big) \end{aligned}$$

computed in the backward GS cycle. The optimal solution x^+ in (2) *can be obtained exactly via*

$$\begin{aligned} x_1^+ &= \arg\min_{x_1} \{ p(x_1) + q(x_1, \widehat{x}_{\ge 2}) - \langle \delta_1^+, x_1 \rangle \} \\ x_i^+ &= \arg\min_{x_i} \{ p(x_1^+) + q(x_{\le i-1}^+, x_i, \widehat{x}_{\ge i+1}) - \langle \delta_i^+, x_i \rangle \} \\ &= \mathcal{Q}_{ii}^{-1} (r_i + \delta_i^+ - \textstyle\sum_{j=1}^{i-1} \mathcal{Q}_{ji}^* x_j^+ - \sum_{j=i+1}^{s} \mathcal{Q}_{ij} \widehat{x}_j) \end{aligned}$$

where x_i^+, $i = 1, 2, \ldots, s$, is computed in the forward GS cycle.

Very useful for multi-block ADMM! Reduces to classical block sGS if $p(\cdot) = 0$

An inexact accelerated block coordinate gradient descent

$$\min\{p(x_1) + \varphi(z) + \phi(z, x) \mid z \in \mathcal{Z},\ x \in \mathcal{X}\}$$

Algorithm 2. Input $y^1 = x^0 \in \mathrm{dom}(p) \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_s$, $t_1 = 1$. Let $\{\epsilon_k\}$ be a nonnegative summable sequence. Iterate

1. Suppose $\delta_i^k, \widehat{\delta}_i^k \in \mathcal{X}_i$, $i = 1, \ldots, s$, with $\widehat{\delta}_1^k = \delta_1^k$, are error vectors such that

$$\max\{\|\delta^k\|, \|\widehat{\delta}^k\|\} \le \epsilon_k/(\sqrt{2}t_k).$$

$$z^k = \arg\min_z \left\{\varphi(z) + \phi(z, y^k)\right\} \quad \text{(elimination via Danskin)}$$

$$x^k = \arg\min_x \left\{p(x_1) + q_k(x) + \frac{1}{2}\|x - y^k\|_{\mathcal{T}}^2 - \langle \Delta(\widehat{\delta}^k, \delta^k), x\rangle\right\}$$

(inexact sGS)

2. Compute $t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$, $y^{k+1} = x^k + \left(\frac{t_k - 1}{t_{k+1}}\right)(x^k - x^{k-1})$.

An inexact accelerated block coordinate gradient descent

Theorem 4

Let $\mathcal{H} = \mathcal{Q} + \mathcal{T}$ *and* $\beta = 2\|\mathcal{D}^{-1/2}\| + \|\mathcal{H}^{-1/2}\|$. *The sequence* $\{(z^k, x^k)\}$ *generated by Algorithm 2 satisfies*

$$0 \leq F(x^k) - F(x^*) \leq \frac{4}{(k+1)^2}\left(\sqrt{\tau} + \beta\bar{\epsilon}_k\right)^2$$

where $\tau = \frac{1}{2}\|x^0 - x^*\|_{\mathcal{H}}^2$, $\bar{\epsilon}_k = \sum_{j=1}^{k} \epsilon_j$.

Inexact ABCD for (D): version 1

Step 1. Suppose $\delta_E^k, \widehat{\delta}_E^k \in \mathcal{R}^{m_E}$, $\delta_I^k, \widehat{\delta}_I^k \in \mathcal{R}^{m_I}$ satisfy

$$\max\{\|\delta_E^k\|, \|\delta_I^k\|, \|\widehat{\delta}_E^k\|, \|\widehat{\delta}_I^k\|\} \leq \frac{\epsilon_k}{\sqrt{2}t_k}.$$

$$\begin{aligned}
(Z^k, v^k) &= \arg\min_{Z,v}\{F(Z, v, \widetilde{S}^k, \widetilde{y}_E^k, \widetilde{y}_I^k)\} && \text{(Projection onto } \mathcal{P}, \mathcal{K}) \\
\widehat{y}_E^k &= \arg\min_{y_E}\{F(Z^k, v^k, \widetilde{S}^k, y_E, \widetilde{y}_I^k) - \langle \widehat{\delta}_E^k, y_E\rangle\} && \text{(Chol or CG)} \\
\widehat{y}_I^k &= \arg\min_{y_I}\{F(Z^k, v^k, \widetilde{S}^k, \widehat{y}_E^k, y_I) - \langle \widehat{\delta}_I^k, y_I\rangle\} && \text{(Chol or CG)} \\
S^k &= \arg\min_{S}\{F(Z^k, v^k, S, \widehat{y}_E^k, \widehat{y}_I^k)\} && \text{(Projection onto } \mathcal{S}_+^n) \\
y_I^k &= \arg\min_{y_I}\{F(Z^k, v^k, S^k, \widehat{y}_E^k, y_I) - \langle \delta_I^k, y_I\rangle\} && \text{(Chol or CG)} \\
y_E^k &= \arg\min_{y_E}\{F(Z^k, v^k, S^k, y_E, y_I^k) - \langle \delta_E^k, y_E\rangle\} && \text{(Chol or CG)}
\end{aligned}$$

Step 2. Set $t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$ and $\beta_k = \frac{t_k-1}{t_{k+1}}$. Compute

$$(\widetilde{S}^{k+1}, \widetilde{y}_E^{k+1}, \widetilde{y}_I^{k+1}) = (1+\beta_k)(S^k, y_E^k, y_I^k) - \beta_k(S^{k-1}, y_E^{k-1}, y_I^{k-1}).$$

Inexact ABCD for (D): version 2

We can also treat (S, y_E, y_I) as a single block and use a semismooth Newton-CG (SNCG) algorithm introduced in [Zhao-Sun-Toh] to solve it inexactly. Choose $\tau = 10^{-6}$.

Step 1. Suppose $\delta_E^k \in \mathcal{R}^{m_E}$, $\delta_I^k \in \mathcal{R}^{m_I}$ are error vectors such that

$$\max\{\|\delta_E^k\|, \|\delta_I^k\|\} \leq \frac{\epsilon_k}{\sqrt{2}t_k}.$$

Compute

$$(Z^k, v^k) = \arg\min_{Z,v} \left\{ F(Z, v, \widetilde{S}^k, \widetilde{y}_E^k, \widetilde{y}_I^k) \right\} \quad \text{(Projection onto } \mathcal{P}, \mathcal{K})$$

$$(S^k, y_E^k, y_I^k) = \arg\min_{S, y_E, y_I} \left\{ \begin{array}{c} F(Z^k, v^k, S, y_E, y_I) + \frac{\tau}{2}\|y_E - \widetilde{y}_E^k\|^2 \\ -\langle \delta_E^k, y_E \rangle - \langle \delta_I^k, y_I \rangle \end{array} \right\} \quad \text{(SNCG)}$$

Step 2. Set $t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$, $\beta_k = \frac{t_k - 1}{t_{k+1}}$. Compute

$$(\widetilde{S}^{k+1}, \widetilde{y}_E^{k+1}, \widetilde{y}_I^{k+1}) = (1+\beta_k)(S^k, y_E^k, y_I^k) - \beta_k(S^{k-1}, y_E^{k-1}, y_I^{k-1}).$$

Numerical experiments

- We compare the performance of ABCD against BCD, APG and eARBCG (an enhanced accelerated randomized block coordinate gradient method) for solving LSSDP.
- We test the algorithms on LSSDP problem (**P**) by taking $G = -C$, $g = 0$ for the data arising from various classes of SDP of the form (**SDP**).

SDP problem sets

Let $\mathcal{P} = \{X \in \mathcal{S}^n \mid X \geq 0\}$.

- SDP relaxation of a binary integer nonconvex quadratic (BIQ) programming:

$$\begin{array}{ll} \min & \frac{1}{2}\langle Q,\, Y\rangle + \langle c,\, x\rangle \\ \text{s.t.} & \operatorname{diag}(Y) - x = 0, \quad \alpha = 1, \\ & X = \left[\begin{array}{cc} Y & x \\ x^T & \alpha \end{array}\right] \in \mathcal{S}^n_+, \quad X \in \mathcal{P} \end{array}$$

- SDP relaxation $\theta_+(G)$ of the maximum stable set problem of a graph G with edge set $\mathcal{E}$:

$$\max\{\langle ee^T,\, X\rangle \mid X_{ij} = 0, (i,j) \in \mathcal{E},\, \langle I,\, X\rangle = 1,\, X \in \mathcal{S}^n_+,\, X \in \mathcal{P}\}$$

- SDP relaxation of clustering problems (RCPs):

$$\min \left\{\langle W,\, I - X\rangle \mid Xe = e, \langle I,\, X\rangle = K, X \in \mathcal{S}^n_+, X \in \mathcal{P}\right\}$$

SDP problem sets

- SDP arising from computing lower bounds for quadratic assignment problems (QAPs):

$$\begin{aligned} v \;:=\; \min \quad & \langle B \otimes A,\, Y \rangle \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^{n} Y^{ii} = I, \quad \langle I,\, Y^{ij} \rangle = \delta_{ij} \quad \forall\, 1 \le i \le j \le n, \\ & \langle E,\, Y^{ij} \rangle = 1 \quad \forall\, 1 \le i \le j \le n, \\ & Y \in \mathcal{S}_{+}^{n^2},\; Y \in \mathcal{P} \end{aligned}$$

  where $\mathcal{P} = \{X \in \mathcal{S}^{n^2} \mid X \ge 0\}$.

- SDP relaxation of frequency assignment problems (FAPs):

SDP problem sets

- In order to get tighter bound for BIQ, we may add some valid inequalities to get the following problems:

$$
\begin{array}{ll}
\min & \frac{1}{2}\langle Q,\, Y\rangle + \langle c,\, x\rangle \\
\text{s.t.} & \operatorname{diag}(Y) - x = 0,\ \alpha = 1,\ X = \begin{bmatrix} Y & x \\ x^T & \alpha \end{bmatrix} \in \mathcal{S}^n_+,\ X \in \mathcal{P} \\
& 0 \le -Y_{ij} + x_i \le 1, \quad 0 \le -Y_{ij} + x_j \le 1 \\
& 0 \le x_i + x_j - Y_{ij} \le 1, \quad \forall\, 1 \le i < j,\ j \le n-1
\end{array}
$$

We call the above problem an extended BIQ (exBIQ).

Stop the algorithms after 25,000 iterations, or

$$\eta = \max\{\eta_1, \eta_2, \eta_3\} < 10^{-6},$$

where $\eta_1 = \frac{\|b_E - \mathcal{A}_E X\|}{1+\|b_E\|}$, $\eta_2 = \frac{\|X-Y\|}{1+\|X\|}$, $\eta_3 = \frac{\|s-\mathcal{A}_I X\|}{1+\|s\|}$

$X = \Pi_{\mathcal{S}^n_+}(\mathcal{A}^*_E y_E + \mathcal{A}^*_I y_I + Z + G)$, $Y = \Pi_{\mathcal{P}}(\mathcal{A}^*_E y_E + \mathcal{A}^*_I y_I + S + G)$, $s = \Pi_{\mathcal{K}}(g - y_I)$.

problem set (No.) \ solver	ABCD	APG	eARBCG	BCD
θ_+ (64)	64	64	64	11
FAP (7)	7	7	7	7
QAP (95)	95	95	24	0
BIQ (165)	165	165	165	65
RCP (120)	120	120	120	108
exBIQ (165)	165	141	165	10
Total (616)	616	592	545	201

Detailed numerical results

Problem	$m_E, m_I; n$ $\mathcal{P}, \mathcal{K}$	η ABCD | APG | eARBCG	time (hour:minute) ABCD | APG | eARBCG
1tc.2048	$18945, 0;$ 2048	9.8-7 | 9.8-7 | 9.4-7	7:35 | 22:18 | 31:38
fap25	$2118, 0;$ 2118	9.2-7 | 8.1-7 | 9.0-7	0:03 | 0:11 | 0:13
nug30	$1393, 0;$ 900	9.6-7 | 9.9-7 | 1.4-6	0:10 | 1:12 | 7:21
tho30	$1393, 0;$ 900	9.9-7 | 9.9-7 | 1.6-6	0:13 | 1:17 | 3:51
ex-gka5f	$501, 0.37M;$ 501	9.8-7 | 1.6-6 | 9.9-7	0:24 | 2:26 | 4:00

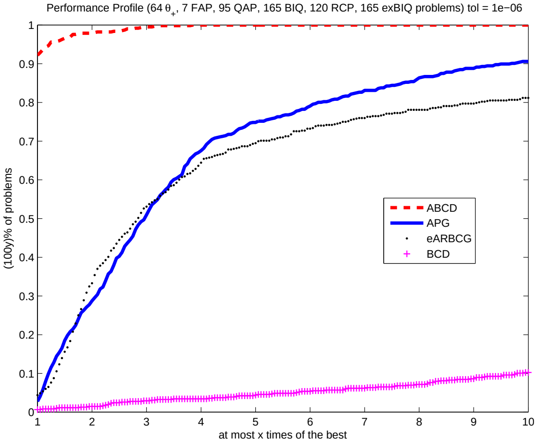

Figure: Performance profiles of ABCD, APG, eARBCG and BCD on $[1, 10]$

Higher accuracy results for ABCD

Number of problems which are solved to the accuracy of 10^{-6}, 10^{-7}, 10^{-8} by the ABCD method.

problem set (No.)	10^{-6}	10^{-7}	10^{-8}
θ_+ (64)	64	58	52
FAP (7)	7	7	7
QAP (95)	95	95	95
BIQ (165)	165	165	165
RCP (120)	120	120	118
exBIQ (165)	165	165	165
Total (616)	616	610	602

Tolerance profiles of the ABCD

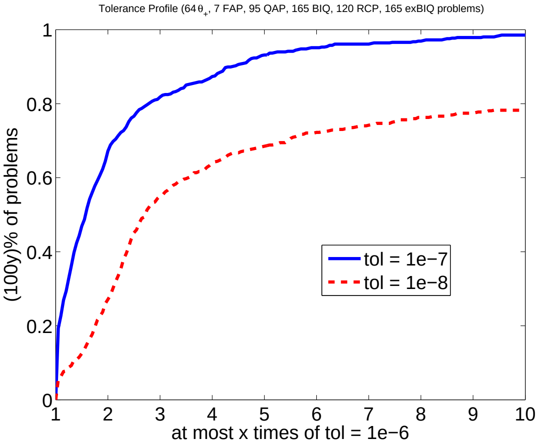

Figure: Tolerance profiles of ABCD on $[1, 10]$

Thank you for your attention!