A block-decomposition framework for solving large-scale convex programming

Camilo Ortiz [1]

Joint work: Renato D.C Monteiro[1] and Benar F. Svaiter[2]

SIAM IS12,
May 20-22, 2012

[1]Georgia Tech
[2]IMPA

Talk Outline

Motivation

- Nonsmooth convex optimization applications:
  - ℓ_1 minimization: Imaging
  - Solve SDPs:
    - Transportation
    - Hard assignment problems (wireless communication networks, microchip production)
    - Binary quadratic problems
    - Various relaxations of (NP) hard combinatorial problems
- Large-scale optimization:
  - Interior point methods (second order): best for small/medium SDPs (at most a few thousands of variables and constraints).
  - First order methods: best for large SDPs (millions of variables and constraints).

Motivation

- Nonsmooth convex optimization applications:
  - ℓ_1 minimization: Imaging
  - Solve SDPs:
    - Transportation
    - Hard assignment problems (wireless communication networks, microchip production)
    - Binary quadratic problems
    - Various relaxations of (NP) hard combinatorial problems
- Large-scale optimization:
  - Interior point methods (second order): best for small/medium SDPs (at most a few thousands of variables and constraints).
  - First order methods: best for large SDPs (millions of variables and constraints).

Optimization and Inclusion Problems

Block Structured Monotone Inclusion Problem (MIP)

Find $(\mathbf{x},\mathbf{u}) \in \mathbb{R}^{\mathbf{n_1}} \times \mathbb{R}^{\mathbf{n_2}}$ such that

$$\mathbf{0} \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D})](\mathbf{x},\mathbf{u}) \subset \mathbb{R}^{\mathbf{n_1}} \times \mathbb{R}^{\mathbf{n_2}} \equiv \mathbf{0} \in \mathbf{F_x}(\mathbf{x},\mathbf{u}) + \mathbf{C}(\mathbf{x}),\ \mathbf{0} \in \mathbf{F_u}(\mathbf{x},\mathbf{u}) + \mathbf{D}(\mathbf{u}).$$

General Convex Optimization Problem (GCOP)

$$\min\ \mathbf{f}(\mathbf{x}) + \mathbf{h_1}(\mathbf{x}) + \mathbf{h_2}(\mathbf{x}).$$

- $\mathbf{f}, \mathbf{h_1}, \mathbf{h_2} : \mathbb{R}^{\mathbf{n}} \to \mathbb{R} \cup \{+\infty\}$ are convex lower semicontinuous proper functions.
- $\mathbf{f}$ is differentiable with Lipschitz constant L on the gradient, i.e.,

$$\|\nabla \mathbf{f}(\mathbf{x}) - \nabla \mathbf{f}(\mathbf{u})\| \leq \mathbf{L}\, \|\mathbf{x} - \mathbf{u}\| \qquad \forall \mathbf{x}, \mathbf{u}. \in \mathbb{R}^{\mathbf{n}}.$$

Setting $\mathbf{F}(\mathbf{x},\mathbf{u}) := \begin{pmatrix} u \\ -x \end{pmatrix}$, $\mathbf{C}(\mathbf{x}) = \nabla \mathbf{f}(\mathbf{x}) + \partial \mathbf{h_1}(\mathbf{x})$ and $\mathbf{D}(\mathbf{u}) = \partial \mathbf{h_2^*}(\mathbf{u})$, implies

Optimization and Inclusion Problems

Block Structured Monotone Inclusion Problem (MIP)

Find $(x,u) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that

$$0 \in [F + C \otimes D](x,u) \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \equiv 0 \in F_x(x,u) + C(x),\ 0 \in F_u(x,u) + D(u).$$

- $C : \mathbb{R}^{n_1} \rightrightarrows \mathbb{R}^{n_1}$ and $D : \mathbb{R}^{n_2} \rightrightarrows \mathbb{R}^{n_2}$ are maximal monotone operators;
- $F : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ is a continuous monotone map, $F(x,u) = (F_x(x,u), F_u(x,u)) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$;
- $F_u(\cdot, u')$ is L_{ux}-Lipschitz continuous.

General Convex Optimization Problem (GCOP)

$$\min\ f(x) + h_1(x) + h_2(x).$$

Setting $F(x,u) := \begin{pmatrix} u \\ -x \end{pmatrix}$, $C(x) = \nabla f(x) + \partial h_1(x)$ and $D(u) = \partial h_2^*(u)$, implies that solving MIP is equivalent to satisfying the optimality conditions of GCOP:

Optimization and Inclusion Problems

Block Structured Monotone Inclusion Problem (MIP)

Find $(\mathbf{x},\mathbf{u}) \in \mathbb{R}^{\mathbf{n_1}} \times \mathbb{R}^{\mathbf{n_2}}$ such that

$$\mathbf{0} \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D})](\mathbf{x},\mathbf{u}) \subset \mathbb{R}^{\mathbf{n_1}} \times \mathbb{R}^{\mathbf{n_2}} \equiv \mathbf{0} \in \mathbf{F_x}(\mathbf{x},\mathbf{u}) + \mathbf{C}(\mathbf{x}),\ \mathbf{0} \in \mathbf{F_u}(\mathbf{x},\mathbf{u}) + \mathbf{D}(\mathbf{u}).$$

General Convex Optimization Problem (GCOP)

$$\min\ \mathbf{f}(\mathbf{x}) + \mathbf{h_1}(\mathbf{x}) + \mathbf{h_2}(\mathbf{x}).$$

Setting $\mathbf{F}(\mathbf{x},\mathbf{u}) := \begin{pmatrix} u \\ -x \end{pmatrix}$, $\mathbf{C}(\mathbf{x}) = \nabla \mathbf{f}(\mathbf{x}) + \partial \mathbf{h_1}(\mathbf{x})$ and $\mathbf{D}(\mathbf{u}) = \partial \mathbf{h_2^*}(\mathbf{u})$, implies that solving MIP is equivalent to satisfying the optimality conditions of GCOP:

$$\begin{matrix} 0 \in u + \nabla f(x) + \partial h_1(x) \\ 0 \in -x + \partial h_2^*(u) \end{matrix} \quad \equiv \quad \boxed{\mathbf{0} \in \nabla \mathbf{f}(\mathbf{x}) + \partial \mathbf{h_1}(\mathbf{x}) + \partial \mathbf{h_2}(\mathbf{x}).}$$

Optimization and Inclusion Problems

General Convex Optimization Problem (GCOP)

$$\min \ f(x) + h_1(x) + h_2(x).$$

Conic Optimization Problem (COP)

PRIMAL		DUAL	
min	$\langle \mathbf{c}, \mathbf{x} \rangle$	max	$\langle \mathbf{b}, \mathbf{y} \rangle + \langle \mathbf{d}, \mathbf{s} \rangle$
s.t	$\mathbf{A}\mathbf{x} - \mathbf{b} \in \mathbf{K}$	s.t	$\mathbf{A}^*\mathbf{y} + \mathbf{B}^*\mathbf{s} = \mathbf{c}$
	$\mathbf{B}\mathbf{x} - \mathbf{d} \in \mathbf{L}$		$\mathbf{y} \in \mathbf{L}*, \mathbf{s} \in \mathbf{K}^*.$

- $\mathbf{K}, \mathbf{L}$ are closed convex cones (usually $\mathbf{K} = \{0\}$ and $\mathbf{L} = \mathbb{R}_+^{n_l} \times \mathbf{S}_+^{n_s}$);
- $\mathbf{A} : \mathbb{R}^{n_l} \times \mathbf{S}^{n_s} \to \mathbb{R}^{m}$, $\mathbf{B} : \mathbb{R}^{n_l} \times \mathbf{S}^{n_s} \to \mathbb{R}^{p}$ are onto linear maps;
- $\mathbf{c} \in \mathbb{R}^{n_l} \times \mathbf{S}^{n_s}$, $\mathbf{b} \in \mathbb{R}^{m}$ and $\mathbf{d} \in \mathbb{R}^{p}$.

Optimization and Inclusion Problems

General Convex Optimization Problem (GCOP)

$$\min \; f(x) + h_1(x) + h_2(x).$$

Conic Optimization Problem (COP)

PRIMAL	DUAL
$\min \; \langle c, x\rangle$ s.t $Ax - b \in K$, $Bx - d \in L$	$\max \; \langle b, y\rangle + \langle d, s\rangle$ s.t $A^*y + B^*s = c$, $y \in L*, s \in K^*$.

COP is a special case of GCOP [Option 1]:

$$\begin{aligned} f(\cdot) &= \langle c, \cdot\rangle \\ h_1(\cdot) &= \delta_L(B(\cdot) - d) \\ h_2(\cdot) &= \delta_K(A(\cdot) - b) \end{aligned} \qquad \text{where} \qquad \delta_K(x) = \begin{cases} 0, & x \in K \\ \infty, & x \notin K \end{cases}.$$

Optimization and Inclusion Problems

General Convex Optimization Problem (GCOP)

$$\min \ \mathbf{f}(\mathbf{x}) + \mathbf{h_1}(\mathbf{x}) + \mathbf{h_2}(\mathbf{x}).$$

Conic Optimization Problem (COP)

PRIMAL		DUAL	
min	$\langle \mathbf{c}, \mathbf{x} \rangle$	max	$\langle \mathbf{b}, \mathbf{y} \rangle + \langle \mathbf{d}, \mathbf{s} \rangle$
s.t	$\mathbf{A}\mathbf{x} - \mathbf{b} \in \mathbf{K}$	s.t	$\mathbf{A}^*\mathbf{y} + \mathbf{B}^*\mathbf{s} = \mathbf{c}$
	$\mathbf{B}\mathbf{x} - \mathbf{d} \in \mathbf{L}$		$\mathbf{y} \in \mathbf{L}*, \mathbf{s} \in \mathbf{K}^*.$

COP is a special case of GCOP [Option 2]:

$$\begin{aligned} \mathbf{f}(\cdot) &= \langle \mathbf{c}, \cdot \rangle \\ \mathbf{h_1}(\cdot) &= \delta_{\mathbf{K}}(\mathbf{A}(\cdot) - \mathbf{b}) \\ \mathbf{h_2}(\cdot) &= \delta_{\mathbf{L}}(\mathbf{B}(\cdot) - \mathbf{d}) \end{aligned} \qquad \text{where} \qquad \delta_{\mathbf{K}}(\mathbf{x}) = \begin{cases} \mathbf{0}, & \mathbf{x} \in \mathbf{K} \\ \infty, & \mathbf{x} \notin \mathbf{K} \end{cases}.$$

Optimization and Inclusion Problems

General Convex Optimization Problem (GCOP)

$$\min \; f(x) + h_1(x) + h_2(x).$$

Conic Optimization Problem (COP)

PRIMAL	DUAL
$\min \; \langle c, x \rangle$	$\max \; \langle b, y \rangle + \langle d, s \rangle$
s.t $Ax - b \in K$	s.t $A^* y + B^* s = c$
$Bx - d \in L$	$y \in L*, s \in K^*.$

COP is a special case of GCOP [Option 3]:

$$\begin{aligned} f(\cdot) &= 0 \\ h_1(\cdot) &= \langle c, \cdot \rangle + \delta_K(A(\cdot) - b) \\ h_2(\cdot) &= \delta_L(B(\cdot) - d) \end{aligned} \qquad \text{where} \qquad \delta_K(x) = \begin{cases} 0, & x \in K \\ \infty, & x \notin K \end{cases}.$$

Optimization and Inclusion Problems

Block Structured Monotone Inclusion Problem (MIP)

Find $(x,u) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that

$$0 \in [F + C \otimes D)](x,u) \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \;\equiv\; 0 \in F_x(x,u) + C(x),\; 0 \in F_u(x,u) + D(u).$$

General Convex Optimization Problem (GCOP)

$$\min \; f(x) + h_1(x) + h_2(x).$$

Conic Optimization Problem (COP)

PRIMAL		DUAL	
min	$\langle c, x\rangle$	max	$\langle b, y\rangle + \langle d, s\rangle$
s.t	$Ax - b \in K$	s.t	$A^* y + B^* s = c$
	$Bx - d \in L$		$y \in L*, s \in K^*.$

BD-HPE: Block-Decomposition Hybrid Proximal Extragradient (Quick view)

Block Structured Monotone Inclusion Problem (MIP)

Find $(\mathbf{x},\mathbf{u}) \in \mathbb{R}^{\mathbf{n_1}} \times \mathbb{R}^{\mathbf{n_2}}$ such that

$$\mathbf{0} \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D})](\mathbf{x},\mathbf{u}) \;\equiv\; \mathbf{0} \in \mathbf{F_x}(\mathbf{x},\mathbf{u}) + \mathbf{C}(\mathbf{x}),\; \mathbf{0} \in \mathbf{F_u}(\mathbf{x},\mathbf{u}) + \mathbf{D}(\mathbf{u}).$$

Proximal equations to obtain next iterates $(\mathbf{x}_+,\mathbf{u}_+)$

$$\mathbf{0} \in \lambda \left[\mathbf{F}(\mathbf{x},\mathbf{u}) + \begin{pmatrix} C(x) \\ D(u) \end{pmatrix}] + \begin{pmatrix} x \\ u \end{pmatrix} - \begin{pmatrix} x_0 \\ u_0 \end{pmatrix} \right] \longrightarrow \begin{pmatrix} x_+ \\ u_+ \end{pmatrix}$$

Inexact Proximal Point Method (Rockafellar 76):

$$\left\| \mathbf{x_k} - (\lambda_\mathbf{k}\mathbf{T} + \mathbf{I})^{-1}(\mathbf{x_{k-1}}) \right\| \le \mathbf{e_k}, \quad \sum_{\mathbf{k}=1}^{\infty} \mathbf{e_k} < \infty. \qquad [\mathbf{T} := \mathbf{F} + \mathbf{C} \otimes \mathbf{D}]$$

HPE Method (Solodov & Svaiter 99): $\mathbf{v_k} \in \mathbf{T}^{\varepsilon_\mathbf{k}}(\tilde{\mathbf{z}}_\mathbf{k})$ and

$$\|\lambda_\mathbf{k}\mathbf{v_k} + \tilde{\mathbf{z}}_\mathbf{k} - \mathbf{z_{k-1}}\|^2 + 2\lambda_\mathbf{k}\varepsilon_\mathbf{k} \le \sigma_\mathbf{k}^2 \|\tilde{\mathbf{z}}_\mathbf{k} - \mathbf{z_{k-1}}\|^2, \qquad [\mathbf{T} := \mathbf{F} + \mathbf{C} \otimes \mathbf{D}]$$

BD-HPE: Block-Decomposition Hybrid Proximal Extragradient (Quick view)

Block Structured Monotone Inclusion Problem (MIP)

Find $(\mathbf{x},\mathbf{u}) \in \mathbb{R}^{\mathbf{n_1}} \times \mathbb{R}^{\mathbf{n_2}}$ such that

$$\mathbf{0} \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D})](\mathbf{x},\mathbf{u}) \;\equiv\; \mathbf{0} \in \mathbf{F_x}(\mathbf{x},\mathbf{u}) + \mathbf{C}(\mathbf{x}),\; \mathbf{0} \in \mathbf{F_u}(\mathbf{x},\mathbf{u}) + \mathbf{D}(\mathbf{u}).$$

Proximal equations to obtain next iterates $(\mathbf{x}_+, \mathbf{u}_+)$

$$\mathbf{0} \in \lambda \left[\mathbf{F}(\mathbf{x},\mathbf{u}) + \begin{pmatrix} C(x) \\ D(u) \end{pmatrix} \right] + \begin{pmatrix} x \\ u \end{pmatrix} - \begin{pmatrix} x_0 \\ u_0 \end{pmatrix} \Bigg] \longrightarrow \begin{pmatrix} x_+ \\ u_+ \end{pmatrix}$$

Inexact Proximal Point Method (Rockafellar 76):

$$\left\| \mathbf{x_k} - (\lambda_\mathbf{k} \mathbf{T} + \mathbf{I})^{-1}(\mathbf{x_{k-1}}) \right\| \le \mathbf{e_k}, \quad \sum_{\mathbf{k}=1}^{\infty} \mathbf{e_k} < \infty. \qquad [\mathbf{T} := \mathbf{F} + \mathbf{C} \otimes \mathbf{D}]$$

HPE Method (Solodov & Svaiter 99): $\mathbf{v_k} \in \mathbf{T}^{\varepsilon_\mathbf{k}}(\tilde{\mathbf{z}}_\mathbf{k})$ and

$$\|\lambda_\mathbf{k} \mathbf{v_k} + \tilde{\mathbf{z}}_\mathbf{k} - \mathbf{z_{k-1}}\|^2 + 2\lambda_\mathbf{k} \varepsilon_\mathbf{k} \le \sigma_\mathbf{k}^2 \|\tilde{\mathbf{z}}_\mathbf{k} - \mathbf{z_{k-1}}\|^2, \qquad [\mathbf{T} := \mathbf{F} + \mathbf{C} \otimes \mathbf{D}]$$

BD-HPE: Block-Decomposition Hybrid Proximal Extragradient (Quick view)

Block Structured Monotone Inclusion Problem (MIP)

Find $(\mathbf{x},\mathbf{u}) \in \mathbb{R}^{\mathbf{n_1}} \times \mathbb{R}^{\mathbf{n_2}}$ such that

$$\mathbf{0} \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D})](\mathbf{x},\mathbf{u}) \;\equiv\; \mathbf{0} \in \mathbf{F_x}(\mathbf{x},\mathbf{u}) + \mathbf{C}(\mathbf{x}),\; \mathbf{0} \in \mathbf{F_u}(\mathbf{x},\mathbf{u}) + \mathbf{D}(\mathbf{u}).$$

Proximal equations to obtain next iterates $(\mathbf{x}_+,\mathbf{u}_+)$

$$\mathbf{0} \in \lambda \left[\mathbf{F}(\mathbf{x},\mathbf{u}) + \begin{pmatrix} C(x) \\ D(u) \end{pmatrix}] + \begin{pmatrix} x \\ u \end{pmatrix} - \begin{pmatrix} x_0 \\ u_0 \end{pmatrix} \right] \longrightarrow \begin{pmatrix} x_+ \\ u_+ \end{pmatrix}$$

Inexact Proximal Point Method (Rockafellar 76):

$$\left\| \mathbf{x_k} - (\lambda_\mathbf{k} \mathbf{T} + \mathbf{I})^{-1} (\mathbf{x_{k-1}}) \right\| \le \mathbf{e_k}, \quad \sum_{\mathbf{k}=1}^{\infty} \mathbf{e_k} < \infty. \qquad [\mathbf{T} := \mathbf{F} + \mathbf{C} \otimes \mathbf{D}]$$

HPE Method (Solodov & Svaiter 99): $\mathbf{v_k} \in \mathbf{T}^{\varepsilon_\mathbf{k}}(\tilde{\mathbf{z}}_\mathbf{k})$ and

$$\|\lambda_\mathbf{k} \mathbf{v_k} + \tilde{\mathbf{z}}_\mathbf{k} - \mathbf{z_{k-1}}\|^2 + 2\lambda_\mathbf{k} \varepsilon_\mathbf{k} \le \sigma_\mathbf{k}^2 \|\tilde{\mathbf{z}}_\mathbf{k} - \mathbf{z_{k-1}}\|^2, \qquad [\mathbf{T} := \mathbf{F} + \mathbf{C} \otimes \mathbf{D}]$$

BD-HPE: Block-Decomposition Hybrid Proximal Extragradient (Quick view)

Block Structured Monotone Inclusion Problem (MIP)

Find $(\mathbf{x},\mathbf{u}) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that

$$\mathbf{0} \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D})](\mathbf{x},\mathbf{u}) \;\equiv\; \mathbf{0} \in \mathbf{F_x}(\mathbf{x},\mathbf{u}) + \mathbf{C}(\mathbf{x}),\; \mathbf{0} \in \mathbf{F_u}(\mathbf{x},\mathbf{u}) + \mathbf{D}(\mathbf{u}).$$

BD-HPE: Decoupling to (approx.) obtain intermediate iterates $(\tilde{\mathbf{x}}_+, \tilde{\mathbf{u}}_+)$

$$\left.\begin{array}{lll} \mathbf{0} \in \lambda[\mathbf{F_x}(\mathbf{x},\mathbf{u_0}) + \mathbf{C}(\mathbf{x})] + \mathbf{x} - \mathbf{x_0} & \longrightarrow \quad \tilde{x}_+ & (\text{and } \tilde{c}_+) \\ \mathbf{0} \in \lambda[\mathbf{F_u}(\tilde{\mathbf{x}}_+,\mathbf{u}) + \mathbf{D}(\mathbf{u})] + \mathbf{u} - \mathbf{u_0} & \longrightarrow \quad \tilde{u}_+ & (\text{and } \tilde{d}_+) \end{array}\right\} \longrightarrow (x_+, u_+)$$

[extragradient step]

HPE Method (Solodov & Svaiter 99): $\mathbf{v_k} \in \mathbf{T}^{\varepsilon_k}(\tilde{\mathbf{z}}_k)$ and

$$\|\lambda_k \mathbf{v_k} + \tilde{\mathbf{z}}_k - \mathbf{z}_{k-1}\|^2 + 2\lambda_k \varepsilon_k \le \sigma_k^2 \|\tilde{\mathbf{z}}_k - \mathbf{z}_{k-1}\|^2, \qquad [\mathbf{T} := \mathbf{F} + \mathbf{C} \otimes \mathbf{D}]$$

$$\tilde{\mathbf{v}}_k := \mathbf{F}(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k) + (\tilde{\mathbf{c}}_k, \tilde{\mathbf{d}}_k) \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D}]^{\varepsilon_k}(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k).$$

BD-HPE: Block-Decomposition Hybrid Proximal Extragradient (Quick view)

General Convex Optimization Problem (GCOP)

$\min\ f(x)+h_1(x)+h_2(x)$. Find $(x,u)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ such that

$$0\in u+\nabla f(x)+\partial h_1(x),\quad 0\in -x+\partial h_2^*(u).$$

BD-HPE: Decoupling to (approx.) obtain intermediate iterates $(\tilde{x}_+,\tilde{u}_+)$

$$\left.\begin{array}{l} 0\in\lambda[u_0+\nabla f(x)+\partial h_1(x)]+x-x_0\longrightarrow\quad \tilde{x}_+ \\ 0\in\lambda[-\tilde{x}_++\partial h_2^*(u)]+u-u_0\longrightarrow\quad \tilde{u}_+ \end{array}\right\}\longrightarrow (x_+,u_+)$$

[extragradient step]

HPE Method (Solodov & Svaiter 99): $v_k\in T^{\varepsilon_k}(\tilde{z}_k)$ and

$$\|\lambda_k v_k+\tilde{z}_k-z_{k-1}\|^2+2\lambda_k\varepsilon_k\le\sigma_k^2\|\tilde{z}_k-z_{k-1}\|^2,\qquad [T:=F+C\otimes D]$$

$$\tilde{v}_k:=F(\tilde{x}_k,\tilde{y}_k)+(\tilde{c}_k,\tilde{d}_k)\in[F+C\otimes D]^{\varepsilon_k}(\tilde{x}_k,\tilde{y}_k).$$

BD-HPE: Block-Decomposition Hybrid Proximal Extragradient (Quick view)

General Convex Optimization Problem (GCOP)

$\min\ f(x) + h_1(x) + h_2(x)$. Find $(x,u) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that

$$0 \in u + \nabla f(x) + \partial h_1(x), \quad 0 \in -x + \partial h_2^*(u).$$

BD-HPE: Solution of the decoupling to (approx.) obtain $(\tilde{x}_+, \tilde{u}_+)$

$$\left.\begin{array}{lcl} \left(I + \tilde{\lambda}\partial h_1\right)^{-1}\left(x_0 - \tilde{\lambda}\left(\nabla f(x_0) + u_0\right)\right) & \longrightarrow & \tilde{x}_+ \\ u_0 + \tilde{\lambda}\tilde{x}_+ - \tilde{\lambda}\left(I + \tilde{\lambda}^{-1}\partial h_2\right)^{-1}\left(\frac{1}{\tilde{\lambda}}u_0 + \tilde{x}_+\right) & \longrightarrow & \tilde{u}_+ \end{array}\right\} \longrightarrow (x_+, u_+)$$

[extragradient step]

HPE Method (Solodov & Svaiter 99): $v_k \in T^{\varepsilon_k}(\tilde{z}_k)$ and

$$\|\lambda_k v_k + \tilde{z}_k - z_{k-1}\|^2 + 2\lambda_k \varepsilon_k \le \sigma_k^2 \|\tilde{z}_k - z_{k-1}\|^2, \qquad [T := F + C \otimes D]$$

$$\tilde{v}_k := F(\tilde{x}_k, \tilde{y}_k) + (\tilde{c}_k, \tilde{d}_k) \in [F + C \otimes D]^{\varepsilon_k}(\tilde{x}_k, \tilde{y}_k).$$

BD-HPE: Block-Decomposition Hybrid Proximal Extragradient (Quick view)

Block Structured Monotone Inclusion Problem (MIP)

Find $(\mathbf{x},\mathbf{u}) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that

$$\mathbf{0} \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D})](\mathbf{x},\mathbf{u}) \;\equiv\; \mathbf{0} \in \mathbf{F_x}(\mathbf{x},\mathbf{u}) + \mathbf{C}(\mathbf{x}),\; \mathbf{0} \in \mathbf{F_u}(\mathbf{x},\mathbf{u}) + \mathbf{D}(\mathbf{u}).$$

BD-HPE: Decoupling to (approx.) obtain intermediate iterates $(\tilde{\mathbf{x}}_+, \tilde{\mathbf{u}}_+)$

$$\left.\begin{array}{lll} \mathbf{0} \in \lambda[\mathbf{F_x}(\mathbf{x},\mathbf{u_0}) + \mathbf{C}(\mathbf{x})] + \mathbf{x} - \mathbf{x_0} & \longrightarrow \quad \tilde{x}_+ & (\text{and } \tilde{c}_+) \\ \mathbf{0} \in \lambda[\mathbf{F_u}(\tilde{\mathbf{x}}_+,\mathbf{u}) + \mathbf{D}(\mathbf{u})] + \mathbf{u} - \mathbf{u_0} & \longrightarrow \quad \tilde{u}_+ & (\text{and } \tilde{d}_+) \end{array}\right\} \longrightarrow (x_+, u_+)$$

[extragradient step]

HPE Method (Solodov & Svaiter 99): $\mathbf{v_k} \in \mathbf{T}^{\varepsilon_k}(\tilde{\mathbf{z}}_\mathbf{k})$ and

$$\|\lambda_k \mathbf{v_k} + \tilde{\mathbf{z}}_\mathbf{k} - \mathbf{z_{k-1}}\|^2 + 2\lambda_k \varepsilon_k \le \sigma_k^2 \|\tilde{\mathbf{z}}_\mathbf{k} - \mathbf{z_{k-1}}\|^2, \qquad [\mathbf{T} := \mathbf{F} + \mathbf{C} \otimes \mathbf{D}]$$

$$\tilde{\mathbf{v}}_\mathbf{k} := \mathbf{F}(\tilde{\mathbf{x}}_\mathbf{k}, \tilde{\mathbf{y}}_\mathbf{k}) + (\tilde{\mathbf{c}}_\mathbf{k}, \tilde{\mathbf{d}}_\mathbf{k}) \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D}]^{\varepsilon_k}(\tilde{\mathbf{x}}_\mathbf{k}, \tilde{\mathbf{y}}_\mathbf{k}).$$

BD-HPE: Block-Decomposition Hybrid Proximal Extragradient (Quick view)

Block Structured Monotone Inclusion Problem (MIP)

Find $(\mathbf{x},\mathbf{u}) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that

$$\mathbf{0} \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D})](\mathbf{x},\mathbf{u}) \;\equiv\; \mathbf{0} \in \mathbf{F_x}(\mathbf{x},\mathbf{u}) + \mathbf{C}(\mathbf{x}),\; \mathbf{0} \in \mathbf{F_u}(\mathbf{x},\mathbf{u}) + \mathbf{D}(\mathbf{u}).$$

BD-HPE: Decoupling to (approx.) obtain intermediate iterates $(\tilde{\mathbf{x}}_+, \tilde{\mathbf{u}}_+)$

$$\mathbf{0} \in \lambda[\mathbf{F_x}(\mathbf{x},\mathbf{u_0}) + \mathbf{C}(\mathbf{x})] + \mathbf{x} - \mathbf{x_0} \longrightarrow \tilde{x}_+ \quad (\text{and } \tilde{c}_+)$$

$$\mathbf{0} \in \lambda[\mathbf{F_u}(\tilde{\mathbf{x}}_+,\mathbf{u}) + \mathbf{D}(\mathbf{u})] + \mathbf{u} - \mathbf{u_0} \longrightarrow \tilde{u}_+ \quad (\text{and } \tilde{d}_+)$$

$\longrightarrow (x_+, u_+)$ [extragradient step]

HPE Method (Solodov & Svaiter 99): $\mathbf{v_k} \in \mathbf{T}^{\varepsilon_k}(\tilde{\mathbf{z}}_\mathbf{k})$ and

$$\|\lambda_\mathbf{k}\mathbf{v_k} + \tilde{\mathbf{z}}_\mathbf{k} - \mathbf{z_{k-1}}\|^2 + 2\lambda_\mathbf{k}\varepsilon_\mathbf{k} \le \sigma_\mathbf{k}^2 \|\tilde{\mathbf{z}}_\mathbf{k} - \mathbf{z_{k-1}}\|^2, \qquad [\mathbf{T} := \mathbf{F} + \mathbf{C} \otimes \mathbf{D}]$$

$$\tilde{\mathbf{v}}_\mathbf{k} := \mathbf{F}(\tilde{\mathbf{x}}_\mathbf{k}, \tilde{\mathbf{y}}_\mathbf{k}) + (\tilde{\mathbf{c}}_\mathbf{k}, \tilde{\mathbf{d}}_\mathbf{k}) \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D}]^{\varepsilon_\mathbf{k}}(\tilde{\mathbf{x}}_\mathbf{k}, \tilde{\mathbf{y}}_\mathbf{k}).$$

Adaptive Extragradient Stepsize
Exploiting the HPE Error Criterion

Recall: $\tilde{\mathbf{v}}_\mathbf{k} \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D}]^{\varepsilon_\mathbf{k}}(\tilde{\mathbf{x}}_\mathbf{k}, \tilde{\mathbf{y}}_\mathbf{k})$

Theorem (Convergence of BD-HPE; Monteiro and Svaiter 2010)

For every $\mathbf{k} \in \mathbb{N}$,

$$\|\tilde{v}_k\| = O\left(\frac{1}{\min_{i \le k}\{\tilde{\lambda}_i\}\sqrt{k}}\right) \quad \text{and} \quad \varepsilon_k = O\left(\frac{1}{\min_{i \le k}\{\tilde{\lambda}_i\} k}\right),$$

if $\lambda = \tilde{\lambda}_\mathbf{k}$ *satisfies the error criterion*

$$\|\lambda \tilde{\mathbf{v}}_\mathbf{k} + (\tilde{\mathbf{x}}_\mathbf{k}, \tilde{\mathbf{y}}_\mathbf{k}) - (\mathbf{x}_{\mathbf{k}-1}, \mathbf{y}_{\mathbf{k}-1})\|^2 + 2\lambda\, \varepsilon_\mathbf{k} \le \sigma_\mathbf{k}^2 \|(\tilde{\mathbf{x}}_\mathbf{k}, \tilde{\mathbf{y}}_\mathbf{k}) - (\mathbf{x}_{\mathbf{k}-1}, \mathbf{y}_{\mathbf{k}-1})\|^2.$$

Speed-up Ingredient #1

Aggressive choice of stepsize $\lambda_k \ge \tilde{\lambda}_k$ (max. satisfying error criterion) for performing the extragradient step: adapted stepsize.

Adaptive Extragradient Stepsize

Exploiting the HPE Error Criterion

Recall: $\tilde{\mathbf{v}}_\mathbf{k} \in [\mathbf{F} + \mathbf{C} \otimes \mathbf{D}]^{\varepsilon_\mathbf{k}}(\tilde{\mathbf{x}}_\mathbf{k}, \tilde{\mathbf{y}}_\mathbf{k})$

Theorem (Convergence of BD-HPE; Monteiro and Svaiter 2010)

For every $\mathbf{k} \in \mathbb{N}$,

$$\|\tilde{v}_k\| = O\left(\frac{1}{\min_{i \le k}\{\tilde{\lambda}_i\}\sqrt{k}}\right) \quad \textit{and} \quad \varepsilon_k = O\left(\frac{1}{\min_{i \le k}\{\tilde{\lambda}_i\}k}\right),$$

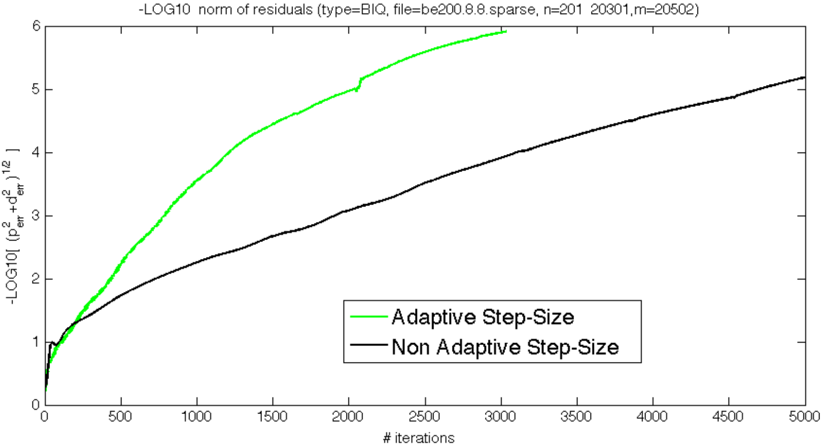
if $\lambda = \tilde{\lambda}_\mathbf{k}$ *satisfies the error criterion*

$$\|\lambda \tilde{\mathbf{v}}_\mathbf{k} + (\tilde{\mathbf{x}}_\mathbf{k}, \tilde{\mathbf{y}}_\mathbf{k}) - (\mathbf{x}_{\mathbf{k}-1}, \mathbf{y}_{\mathbf{k}-1})\|^2 + 2\lambda\, \varepsilon_\mathbf{k} \le \sigma_\mathbf{k}^2 \|(\tilde{\mathbf{x}}_\mathbf{k}, \tilde{\mathbf{y}}_\mathbf{k}) - (\mathbf{x}_{\mathbf{k}-1}, \mathbf{y}_{\mathbf{k}-1})\|^2.$$

Speed-up Ingredient #1

Aggressive choice of stepsize $\lambda_k \ge \tilde{\lambda}_k$ (max. satisfying error criterion) for performing the extragradient step: adapted stepsize.

Example: Adaptive Stepsize



Scaled Inner Product

General Convex Optimization Problem (GCOP)

$\min\ f(x)+h_1(x)+h_2(x)$. Find $(x,u)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ such that

$$0\in u+\nabla f(x)+\partial h_1(x),\ \ 0\in -x+\partial h_2^*(u).$$

Assume that x and u are in different normed spaces:

- let x be in the $X:=\mathbb{R}^n$ space;
- let u be in the $U:=\mathbb{R}^n$ space:

Given $\theta>0$, consider the inner product in X defined as

$$\langle\cdot,\cdot\rangle_\theta:=\theta^{-1}\langle\cdot,\cdot\rangle.$$

Unscaled

$$\left(I+\tilde{\lambda}\,\partial h_1\right)^{-1}\left(x_0-\tilde{\lambda}\,(\nabla f(x_0)+u_0)\right)\longrightarrow\tilde{x}_+$$

Scaled Inner Product

General Convex Optimization Problem (GCOP)

$\min\ f(x)+h_1(x)+h_2(x)$. Find $(x,u)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ such that

$$0\in u+\nabla f(x)+\partial h_1(x),\quad 0\in -x+\partial h_2^*(u).$$

Assume that x and u are in different normed spaces:

- let x be in the $X:=\mathbb{R}^n$ space;
- let u be in the $U:=\mathbb{R}^n$ space:

Given $\theta>0$, consider the inner product in X defined as

$$\langle\cdot,\cdot\rangle_\theta := \theta^{-1}\langle\cdot,\cdot\rangle.$$

Unscaled

$$\left(I+\tilde{\lambda}\,\partial h_1\right)^{-1}\left(x_0-\tilde{\lambda}\,(\nabla f(x_0)+u_0)\right)\longrightarrow \tilde{x}_+$$

Scaled Inner Product

General Convex Optimization Problem (GCOP)

$\min\ f(x)+h_1(x)+h_2(x)$. Find $(x,u)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ such that

$$0\in u+\nabla f(x)+\partial h_1(x),\ 0\in -x+\partial h_2^*(u).$$

Assume that x and u are in different normed spaces:

- let x be in the $X:=\mathbb{R}^n$ space;
- let u be in the $U:=\mathbb{R}^n$ space:

Given $\theta>0$, consider the inner product in X defined as

$$\langle\cdot,\cdot\rangle_\theta := \theta^{-1}\langle\cdot,\cdot\rangle.$$

Unscaled

$$\left(I+\tilde{\lambda}\,\partial h_1\right)^{-1}\left(x_0-\tilde{\lambda}\,(\nabla f(x_0)+u_0)\right)\longrightarrow \tilde{x}_+$$

Scaled Inner Product

General Convex Optimization Problem (GCOP)

$\min\ f(x)+h_1(x)+h_2(x)$. Find $(x,u)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ such that

$$0\in u+\nabla f(x)+\partial h_1(x),\ 0\in -x+\partial h_2^*(u).$$

Assume that x and u are in different normed spaces:

- let x be in the $X:=\mathbb{R}^n$ space;
- let u be in the $U:=\mathbb{R}^n$ space:

Given $\theta>0$, consider the inner product in X defined as

$$\langle\cdot,\cdot\rangle_\theta:=\theta^{-1}\langle\cdot,\cdot\rangle.$$

Scaled

$$\left(I+\tilde\lambda\theta\partial h_1\right)^{-1}\left(x_0-\tilde\lambda\theta(\nabla f(x_0)+u_0)\right)\longrightarrow\tilde{x}_+$$

SA-BD Method for solving GCOP

SA-BD: Scaled Adaptive Block-Decomposition

0) Let $x_0, u_0 \in \mathbb{R}^n$, , $\theta > 0$,$\sigma \in (0,1)$, $\tilde{\lambda} > 0$ be given such that $\max\left\{\tilde{\lambda} L, \tilde{\lambda}^2\right\} \le \frac{\sigma^2}{\theta}$, and set $k = 1$;

1) set $\tilde{x}_k := \left(I + \tilde{\lambda}\theta\partial h_1\right)^{-1} \left(x_{k-1} - \tilde{\lambda}\theta\left(\nabla f(x_{k-1}) + u_{k-1}\right)\right)$;

2) set $\tilde{u}_k := u_{k-1} + \tilde{\lambda}\tilde{x}_k - \tilde{\lambda}\left(I + \tilde{\lambda}^{-1}\partial h_2\right)^{-1} \left(\frac{1}{\tilde{\lambda}} u_{k-1} + \tilde{x}_k\right)$;

3) choose λ_k to be the largest $\lambda > 0$ such that

$$\left\|\lambda v_k + \begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]} + 2\lambda\varepsilon_k \le \sigma^2 \left\|\begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]}; \qquad (1)$$

4) set

$$(x_k, u_k) = (x_{k-1}, u_{k-1}) - \lambda_k v_k \qquad \text{(adaptive extragradient step)}$$

and $k \leftarrow k+1$, and go to step 1.

SA-BD Method for solving GCOP

SA-BD: Scaled Adaptive Block-Decomposition

0) Let $x_0, u_0 \in \mathbb{R}^n$, , $\theta > 0$,$\sigma \in (0,1)$, $\tilde{\lambda} > 0$ be given such that $\max\left\{\tilde{\lambda} L, \tilde{\lambda}^2\right\} \leq \frac{\sigma^2}{\theta}$, and set $k = 1$;

1) set $\tilde{x}_k := \left(I + \tilde{\lambda}\theta\partial h_1\right)^{-1}\left(x_{k-1} - \tilde{\lambda}\theta\left(\nabla f(x_{k-1}) + u_{k-1}\right)\right)$;

2) set $\tilde{u}_k := u_{k-1} + \tilde{\lambda}\tilde{x}_k - \tilde{\lambda}\left(I + \tilde{\lambda}^{-1}\partial h_2\right)^{-1}\left(\frac{1}{\tilde{\lambda}}u_{k-1} + \tilde{x}_k\right)$;

3) choose λ_k to be the largest $\lambda > 0$ such that

$$\left\|\lambda v_k + \begin{pmatrix}\tilde{x}_k\\ \tilde{u}_k\end{pmatrix} - \begin{pmatrix}x_{k-1}\\ u_{k-1}\end{pmatrix}\right\|^2_{[\theta,1]} + 2\lambda\varepsilon_k \leq \sigma^2\left\|\begin{pmatrix}\tilde{x}_k\\ \tilde{u}_k\end{pmatrix} - \begin{pmatrix}x_{k-1}\\ u_{k-1}\end{pmatrix}\right\|^2_{[\theta,1]}; \qquad (1)$$

4) set

$$(x_k, u_k) = (x_{k-1}, u_{k-1}) - \lambda_k v_k \qquad \text{(adaptive extragradient step)}$$

and $k \leftarrow k+1$, and go to step 1.

SA-BD Method for solving GCOP

SA-BD: Scaled Adaptive Block-Decomposition

0) Let $x_0, u_0 \in \mathbb{R}^n$, , $\theta > 0$,$\sigma \in (0,1)$, $\tilde{\lambda} > 0$ be given such that $\max\left\{\tilde{\lambda} L, \tilde{\lambda}^2\right\} \leq \frac{\sigma^2}{\theta}$, and set $k = 1$;

1) set $\tilde{x}_k := \left(I + \tilde{\lambda}\theta\partial h_1\right)^{-1}\left(x_{k-1} - \tilde{\lambda}\theta\left(\nabla f(x_{k-1}) + u_{k-1}\right)\right)$;

2) set $\tilde{u}_k := u_{k-1} + \tilde{\lambda}\tilde{x}_k - \tilde{\lambda}\left(I + \tilde{\lambda}^{-1}\partial h_2\right)^{-1}\left(\frac{1}{\tilde{\lambda}}u_{k-1} + \tilde{x}_k\right)$;

3) choose λ_k to be the largest $\lambda > 0$ such that

$$\left\|\lambda v_k + \begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]} + 2\lambda\varepsilon_k \leq \sigma^2 \left\|\begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]}; \quad (1)$$

4) set

$$(x_k, u_k) = (x_{k-1}, u_{k-1}) - \lambda_k v_k \qquad \text{(adaptive extragradient step)}$$

and $k \leftarrow k+1$, and go to step 1.

SA-BD Method for solving GCOP

SA-BD: Scaled Adaptive Block-Decomposition

0) Let $x_0, u_0 \in \mathbb{R}^n$, , $\theta > 0$,$\sigma \in (0,1)$, $\tilde{\lambda} > 0$ be given such that $\max\left\{\tilde{\lambda} L, \tilde{\lambda}^2\right\} \leq \frac{\sigma^2}{\theta}$, and set $k = 1$;

1) set $\tilde{x}_k := \left(I + \tilde{\lambda}\theta\partial h_1\right)^{-1}\left(x_{k-1} - \tilde{\lambda}\theta\left(\nabla f(x_{k-1}) + u_{k-1}\right)\right)$;

2) set $\tilde{u}_k := u_{k-1} + \tilde{\lambda}\tilde{x}_k - \tilde{\lambda}\left(I + \tilde{\lambda}^{-1}\partial h_2\right)^{-1}\left(\frac{1}{\tilde{\lambda}} u_{k-1} + \tilde{x}_k\right)$;

3) choose λ_k to be the largest $\lambda > 0$ such that

$$\left\|\lambda v_k + \begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]} + 2\lambda\varepsilon_k \leq \sigma^2 \left\|\begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]}; \qquad (1)$$

4) set

$$(x_k, u_k) = (x_{k-1}, u_{k-1}) - \lambda_k v_k \qquad \text{(adaptive extragradient step)}$$

and $k \leftarrow k+1$, and go to step 1.

SA-BD Method for solving GCOP

SA-BD: Scaled Adaptive Block-Decomposition

0) Let $x_0, u_0 \in \mathbb{R}^n$, , $\theta > 0$,$\sigma \in (0,1)$, $\tilde{\lambda} > 0$ be given such that $\max\left\{\tilde{\lambda} L, \tilde{\lambda}^2\right\} \leq \frac{\sigma^2}{\theta}$, and set $k = 1$;

1) set $\tilde{x}_k := \left(I + \tilde{\lambda}\theta\partial h_1\right)^{-1}\left(x_{k-1} - \tilde{\lambda}\theta\left(\nabla f(x_{k-1}) + u_{k-1}\right)\right)$;

2) set $\tilde{u}_k := u_{k-1} + \tilde{\lambda}\tilde{x}_k - \tilde{\lambda}\left(I + \tilde{\lambda}^{-1}\partial h_2\right)^{-1}\left(\frac{1}{\tilde{\lambda}} u_{k-1} + \tilde{x}_k\right)$;

3) choose λ_k to be the largest $\lambda > 0$ such that

$$\left\|\lambda v_k + \begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]} + 2\lambda\varepsilon_k \leq \sigma^2 \left\|\begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]}; \qquad (1)$$

4) set

$$(x_k, u_k) = (x_{k-1}, u_{k-1}) - \lambda_k v_k \qquad \text{(adaptive extragradient step)}$$

and $k \leftarrow k+1$, and go to step 1.

SA-BD Method for solving GCOP

SA-BD: Scaled Adaptive Block-Decomposition

0) Let $x_0, u_0 \in \mathbb{R}^n$, , $\theta > 0$,$\sigma \in (0,1)$, $\tilde{\lambda} > 0$ be given such that $\max\left\{\tilde{\lambda} L, \tilde{\lambda}^2\right\} \leq \frac{\sigma^2}{\theta}$, and set $k = 1$;

1) set $\tilde{x}_k := \left(I + \tilde{\lambda}\theta\partial h_1\right)^{-1}\left(x_{k-1} - \tilde{\lambda}\theta\left(\nabla f(x_{k-1}) + u_{k-1}\right)\right)$;

2) set $\tilde{u}_k := u_{k-1} + \tilde{\lambda}\tilde{x}_k - \tilde{\lambda}\left(I + \tilde{\lambda}^{-1}\partial h_2\right)^{-1}\left(\frac{1}{\tilde{\lambda}} u_{k-1} + \tilde{x}_k\right)$;

3) choose λ_k to be the largest $\lambda > 0$ such that

$$\left\|\lambda v_k + \begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]} + 2\lambda\varepsilon_k \leq \sigma^2 \left\|\begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]}; \qquad (1)$$

4) set

$$(x_k, u_k) = (x_{k-1}, u_{k-1}) - \lambda_k v_k \qquad \text{(adaptive extragradient step)}$$

and $k \leftarrow k+1$, and go to step 1.

SA-BD Method for solving GCOP

SA-BD: Scaled Adaptive Block-Decomposition

0) Let $x_0, u_0 \in \mathbb{R}^n$, , $\theta > 0$,$\sigma \in (0,1)$, $\tilde{\lambda} > 0$ be given such that $\max\left\{\tilde{\lambda} L, \tilde{\lambda}^2\right\} \leq \frac{\sigma^2}{\theta}$, and set $k = 1$;

1) set $\tilde{x}_k := \left(I + \tilde{\lambda}\theta\partial h_1\right)^{-1}\left(x_{k-1} - \tilde{\lambda}\theta\left(\nabla f(x_{k-1}) + u_{k-1}\right)\right)$;

2) set $\tilde{u}_k := u_{k-1} + \tilde{\lambda}\tilde{x}_k - \tilde{\lambda}\left(I + \tilde{\lambda}^{-1}\partial h_2\right)^{-1}\left(\frac{1}{\tilde{\lambda}}u_{k-1} + \tilde{x}_k\right)$;

3) choose λ_k to be the largest $\lambda > 0$ such that

$$\left\|\lambda v_k + \begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]} + 2\lambda\varepsilon_k \leq \sigma^2 \left\|\begin{pmatrix} \tilde{x}_k \\ \tilde{u}_k \end{pmatrix} - \begin{pmatrix} x_{k-1} \\ u_{k-1} \end{pmatrix}\right\|^2_{[\theta,1]}; \qquad (1)$$

4) set

$$(x_k, u_k) = (x_{k-1}, u_{k-1}) - \lambda_k v_k \qquad \text{(adaptive extragradient step)}$$

and $k \leftarrow k + 1$, and go to step 1.

The resolvent operator: $(\mathbf{I} + \lambda \partial \mathbf{h})^{-1}$

Fact

$$(I + \lambda \partial h)^{-1}(x_0) \equiv \arg\min_{x \in \mathbb{R}^n} \frac{1}{2}\|x - x_0\|^2 + \lambda h(x) =: \mathrm{prox}_{\lambda, h}(x_0)$$

- $h = \langle a, \cdot \rangle \implies (I + \lambda \partial h)^{-1}(x_0) = x_0 - \lambda a.$
- $h = \delta_M \implies (I + \lambda \partial h)^{-1} = \Pi_M$, where

$$\Pi_M = \arg\min_{x \in M} \|x - x_0\|.$$

- $h = \|\cdot\|_1 \implies (I + \lambda \partial h)^{-1}(x_0) = x_0 - \Pi_{[-\lambda, \lambda]^n}(x_0)$ (shrinkage)

Residuals

General Convex Optimization Problem (GCOP)

min $f(x)+h_1(x)+h_2(x)$. Find $(x,u)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ such that

$$0\in u+\nabla f(x)+\partial h_1(x),\ 0\in -x+\partial h_2^*(u).$$

GCOP residuals (infeasibilities):

$$\varepsilon_1=\left\|-\tilde{x}+\tilde{h}_2\right\|,\quad \varepsilon_2=\left\|\tilde{u}+\nabla f(\tilde{x})+\tilde{h}_1\right\|,$$

for some

$$\tilde{h}_1\in\partial h_1(\tilde{x}),\quad \tilde{h}_2\in\partial h_2^*(\tilde{u}).$$

Residuals

Conic Optimization Problem (COP)

PRIMAL	DUAL
$\min \ \langle \mathbf{c},\mathbf{x}\rangle$	$\max \ \langle \mathbf{b},\mathbf{y}\rangle + \langle \mathbf{d},\mathbf{s}\rangle$
s.t $\mathbf{A}\mathbf{x}-\mathbf{b}\in \mathbf{K}$	s.t $\mathbf{A}^*\mathbf{y}+\mathbf{B}^*\mathbf{s}=\mathbf{c}$
$\mathbf{B}\mathbf{x}-\mathbf{d}\in \mathbf{L}$	$\mathbf{y}\in \mathbf{L}*, \mathbf{s}\in \mathbf{K}^*.$

COP residuals (infeasibilities) [general]:

$$\varepsilon_1 = \|\tilde{\mathbf{x}}-\tilde{\mathbf{w}}\|, \quad \varepsilon_2 = \|\mathbf{A}^*\tilde{\mathbf{y}}+\mathbf{B}^*\tilde{\mathbf{s}}-\mathbf{c}\|,$$

for some $\tilde{\mathbf{w}}\in\{\mathbf{x}:\mathbf{A}\mathbf{x}-\mathbf{b}\in\mathbf{K}\}$.

Residuals

Conic Optimization Problem (COP)

PRIMAL		DUAL	
min	$\langle \mathbf{c}, \mathbf{x} \rangle$	max	$\langle \mathbf{b}, \mathbf{y} \rangle + \langle \mathbf{d}, \mathbf{s} \rangle$
s.t	$\mathbf{A}\mathbf{x} - \mathbf{b} \in \mathbf{K}$	s.t	$\mathbf{A}^*\mathbf{y} + \mathbf{B}^*\mathbf{s} = \mathbf{c}$
	$\mathbf{B}\mathbf{x} - \mathbf{d} \in \mathbf{L}$		$\mathbf{y} \in \mathbf{L}_*, \mathbf{s} \in \mathbf{K}^*.$

Satisfied every iteration for $x = \tilde{x}$ and $\tilde{u} = -A^*y$!

COP residuals (infeasibilities) [general]:

$$\varepsilon_1 = \|\tilde{\mathbf{x}} - \tilde{\mathbf{w}}\|, \quad \varepsilon_2 = \|\mathbf{A}^*\tilde{\mathbf{y}} + \mathbf{B}^*\tilde{\mathbf{s}} - \mathbf{c}\|,$$

for some $\tilde{\mathbf{w}} \in \{\mathbf{x} : \mathbf{A}\mathbf{x} - \mathbf{b} \in \mathbf{K}\}$.

Residuals

Conic Optimization Problem (COP)

PRIMAL		DUAL	
min	$\langle \mathbf{c}, \mathbf{x} \rangle$	max	$\langle \mathbf{b}, \mathbf{y} \rangle + \langle \mathbf{d}, \mathbf{s} \rangle$
s.t	$\mathbf{A}\mathbf{x} - \mathbf{b} \in \mathbf{K}$	s.t	$\mathbf{A}^*\mathbf{y} + \mathbf{B}^*\mathbf{s} = \mathbf{c}$
	$\mathbf{B}\mathbf{x} - \mathbf{d} \in \mathbf{L}$		$\mathbf{y} \in \mathbf{L}*, \mathbf{s} \in \mathbf{K}^*.$

Satisfied every iteration for $x = \tilde{x}$ and $\tilde{u} = -A^*y$!

COP residuals (infeasibilities) [general]:

$$\varepsilon_1 = \|\tilde{\mathbf{x}} - \tilde{\mathbf{w}}\|, \quad \varepsilon_2 = \|\mathbf{A}^*\tilde{\mathbf{y}} + \mathbf{B}^*\tilde{\mathbf{s}} - \mathbf{c}\|,$$

for some $\tilde{\mathbf{w}} \in \{\mathbf{x} : \mathbf{A}\mathbf{x} - \mathbf{b} \in \mathbf{K}\}$.

Residuals

COP residuals (infeasibilities) [general]:

$$\varepsilon_1 = \|\tilde{\mathbf{x}} - \tilde{\mathbf{w}}\|, \quad \varepsilon_2 = \|\mathbf{A}^*\tilde{\mathbf{y}} + \mathbf{B}^*\tilde{\mathbf{s}} - \mathbf{c}\|,$$

for some $\tilde{\mathbf{w}} \in \{\mathbf{x} : \mathbf{A}\mathbf{x} - \mathbf{b} \in \mathbf{K}\}$.

COP residuals (infeasibilities) [practical]:

$$\varepsilon_P = \frac{\mathrm{dist}\,(\mathbf{A}(\mathbf{x}) - \mathbf{b}, \mathbf{K})}{1 + \|\mathbf{b}\|}, \quad \varepsilon_D = \frac{\|\mathbf{A}^*\mathbf{y} + \mathbf{B}^*\mathbf{s} - \mathbf{c}\|}{1 + \|\mathbf{c}\|},$$

$$\mathrm{gap} = \frac{\langle \mathbf{c}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{y}\rangle - \langle \mathbf{d}, \mathbf{s}\rangle}{|\langle \mathbf{c}, \mathbf{x}\rangle| + |\langle \mathbf{b}, \mathbf{y}\rangle + \langle \mathbf{d}, \mathbf{s}\rangle| + 1}.$$

Complementarity is also satisfied every iteration:

$$\langle \tilde{\mathbf{s}}, \mathbf{B}\tilde{\mathbf{x}} - \mathbf{d}\rangle = 0, \quad \langle \tilde{\mathbf{y}}, \mathbf{A}\tilde{\mathbf{w}} - \mathbf{b}\rangle = 0.$$

Residuals

COP residuals (infeasibilities) [general]:

$$\varepsilon_1 = \|\tilde{\mathbf{x}} - \tilde{\mathbf{w}}\|, \quad \varepsilon_2 = \|\mathbf{A}^*\tilde{\mathbf{y}} + \mathbf{B}^*\tilde{\mathbf{s}} - \mathbf{c}\|,$$

for some $\tilde{\mathbf{w}} \in \{\mathbf{x} : \mathbf{A}\mathbf{x} - \mathbf{b} \in \mathbf{K}\}$.

COP residuals (infeasibilities) [practical]:

$$\varepsilon_{\mathbf{P}} = \frac{\text{dist}\,(\mathbf{A}(\mathbf{x}) - \mathbf{b}, \mathbf{K})}{1 + \|\mathbf{b}\|}, \quad \varepsilon_{\mathbf{D}} = \frac{\|\mathbf{A}^*\mathbf{y} + \mathbf{B}^*\mathbf{s} - \mathbf{c}\|}{1 + \|\mathbf{c}\|},$$

$$\text{gap} = \frac{\langle \mathbf{c}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{y}\rangle - \langle \mathbf{d}, \mathbf{s}\rangle}{|\langle \mathbf{c}, \mathbf{x}\rangle| + |\langle \mathbf{b}, \mathbf{y}\rangle + \langle \mathbf{d}, \mathbf{s}\rangle| + 1}.$$

Complementarity is also satisfied every iteration:

$$\langle \tilde{\mathbf{s}}, \mathbf{B}\tilde{\mathbf{x}} - \mathbf{d}\rangle = \mathbf{0}, \quad \langle \tilde{\mathbf{y}}, \mathbf{A}\tilde{\mathbf{w}} - \mathbf{b}\rangle = \mathbf{0}.$$

Residuals

COP residuals (infeasibilities) [general]:

$$\varepsilon_1 = \|\tilde{\mathbf{x}} - \tilde{\mathbf{w}}\|, \quad \varepsilon_2 = \|\mathbf{A}^*\tilde{\mathbf{y}} + \mathbf{B}^*\tilde{\mathbf{s}} - \mathbf{c}\|,$$

for some $\tilde{\mathbf{w}} \in \{\mathbf{x} : \mathbf{A}\mathbf{x} - \mathbf{b} \in \mathbf{K}\}$.

COP residuals (infeasibilities) [practical]:

$$\varepsilon_P = \frac{\operatorname{dist}(\mathbf{A}(\mathbf{x}) - \mathbf{b}, \mathbf{K})}{1 + \|\mathbf{b}\|}, \quad \varepsilon_D = \frac{\|\mathbf{A}^*\mathbf{y} + \mathbf{B}^*\mathbf{s} - \mathbf{c}\|}{1 + \|\mathbf{c}\|},$$

$$\text{gap} = \frac{\langle \mathbf{c}, \mathbf{x}\rangle - \langle \mathbf{b}, \mathbf{y}\rangle - \langle \mathbf{d}, \mathbf{s}\rangle}{|\langle \mathbf{c}, \mathbf{x}\rangle| + |\langle \mathbf{b}, \mathbf{y}\rangle + \langle \mathbf{d}, \mathbf{s}\rangle| + 1}.$$

NOT NEEDED!

Complementarity is also satisfied every iteration:

$$\langle \tilde{\mathbf{s}}, \mathbf{B}\tilde{\mathbf{x}} - \mathbf{d}\rangle = \mathbf{0}, \quad \langle \tilde{\mathbf{y}}, \mathbf{A}\tilde{\mathbf{w}} - \mathbf{b}\rangle = \mathbf{0}.$$

DSA-BD method (Monteiro,Ortiz and Svaiter 2012)

DSA-BD: Dynamic Scaled Adaptive Block Decomposition

DSA-BD adds the following two ingredients:

Speed-up Ingredient #2

Given $\mathbf{x_0}, \mathbf{u_0} \in \mathbb{R}^\mathbf{n}$, set θ such that $\varepsilon_1, \varepsilon_2 = \mathbf{O(1)}$ for the first iteration } initialization.

Speed-up Ingredient #3

Every $\bar{k}$ iterations, properly change θ if $\gamma^{-1} \le \varepsilon_1/\varepsilon_2 \le \gamma$ is NOT satisfied for a given scalar $\gamma > 1$: dynamic scaled inner product.

DSA-BD method (Monteiro,Ortiz and Svaiter 2012)

DSA-BD: Dynamic Scaled Adaptive Block Decomposition

DSA-BD adds the following two ingredients:

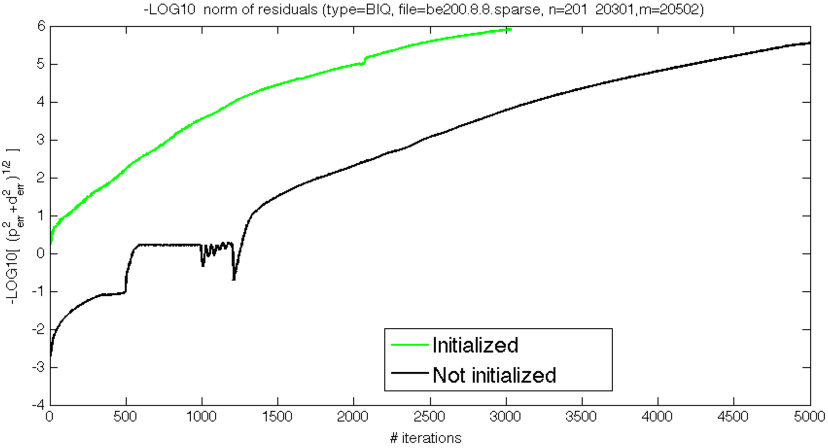
Speed-up Ingredient #2

Given $\mathbf{x_0}, \mathbf{u_0} \in \mathbb{R}^{\mathbf{n}}$, set θ such that $\varepsilon_1, \varepsilon_2 = \mathbf{O(1)}$ for the first iteration } initialization.

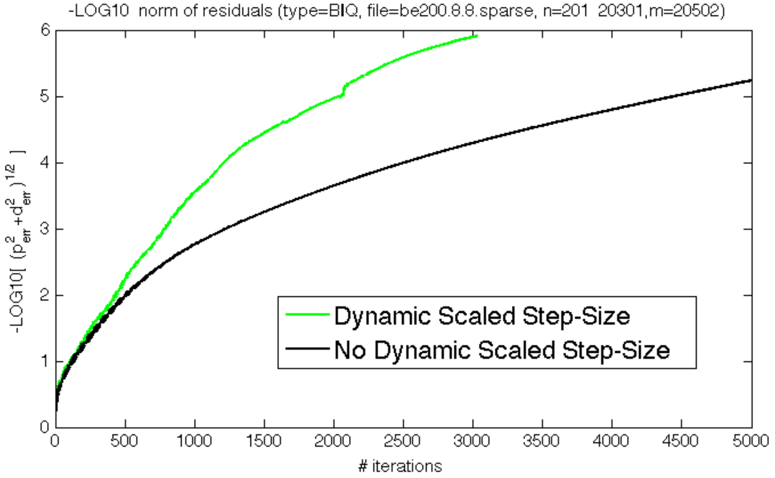
Speed-up Ingredient #3

Every $\bar{k}$ iterations, properly change θ if $\gamma^{-1} \leq \varepsilon_1/\varepsilon_2 \leq \gamma$ is NOT satisfied for a given scalar $\gamma > 1$: dynamic scaled inner product.

Example: Initialization

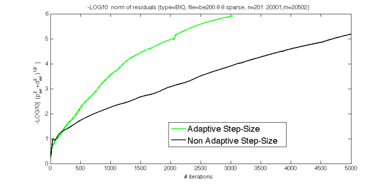


Example: Dynamic scaled inner product

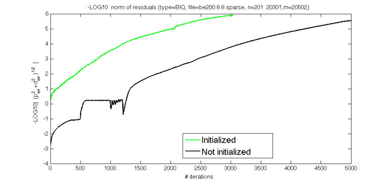


DSA-BD method: Application of BD-HPE with 4 main modifications

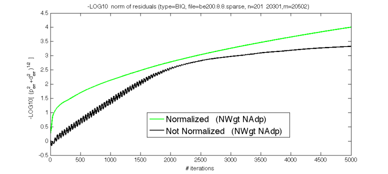
Adaptive step-size



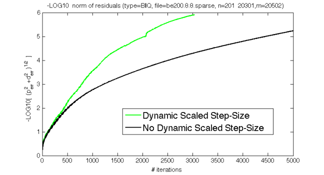
Initialization



Normalization



Dynamic scaled inner product



Instances tested

SDP:

$$\min\{\langle \mathbf{c},\mathbf{x}\rangle : \mathbf{A}(\mathbf{x}) = \mathbf{b},\, \mathbf{x}\in \mathbf{K}\}, \qquad \textbf{(primal)}$$

$$\max\{\langle \mathbf{b},\mathbf{y}\rangle : \mathbf{A}^*(\mathbf{y}) + \mathbf{s} = \mathbf{c},\, \mathbf{s}\in \mathbf{K}^*\} \qquad \textbf{(dual)}$$

$$\mathbf{f}(\cdot) = \langle \mathbf{c},\cdot\rangle$$
$$\mathbf{h_1}(\cdot) = \delta_{\mathbf{K}}(\cdot)$$
$$\mathbf{h_2}(\cdot) = \delta_{\{0\}}(\mathbf{A}(\cdot) - \mathbf{b})$$

Classes of SDPs solved:

- RAND: Sparse SDP randomly generated (Povh, Rendl and Wiegele; 2006).
- FAP: SDP relaxation of frequency assignment problems.
- QAP: SDP relaxation of quadratic assignment problems.
- BIQ: SDP relaxation of binary integer quadratic problems.
- THETA: SDP relaxation of maximum stable set problems.

Specialized implementation example

THETA. SDP relaxation of maximum stable set problems:

Given a graph G with edge set E, we know the stability number $\alpha(G)$ is bounded by $\alpha(G) \leq \theta_+(G) \leq \theta(G)$, where

$$\theta(G) \quad = \quad \max\left\{\left\langle ee^t, x\right\rangle : \langle E_{ij}, x\rangle = 0, (i,j) \in E, \langle I, x\rangle = 1, x \succeq 0\right\},$$

$E_{ij} = e_i e_j^T + e_j e_i^T$ and e_i denotes column i of the identity matrix I.

We can define the blocks of the GCOP [$\min\ \mathbf{f(x) + h_1(x) + h_2(x)}$] as:

$$\begin{aligned} \mathbf{f}(\cdot) &= \langle \mathbf{ee^t}, \cdot\rangle \\ \mathbf{h_1}(\cdot) &= \delta_{\mathbf{S^n_+}}(\cdot) \text{ or } \mathbf{h_1}(\cdot) = \delta_{\mathbf{S^n_+} \cap \{\mathbf{x}:\langle \mathbf{I},\mathbf{x}\rangle = \mathbf{1}\}}(\cdot) \\ \mathbf{h_2}(\cdot) &= \delta_{\{\mathbf{x}:\langle \mathbf{I},\mathbf{x}\rangle = \mathbf{1}, \langle \mathbf{E_{ij}},\mathbf{x}\rangle = \mathbf{0}\}}(\cdot) \end{aligned}$$

Specialized implementation example

THETA. SDP relaxation of maximum stable set problems:

Given a graph G with edge set E, we know the stability number $\alpha(G)$ is bounded by $\alpha(G) \le \theta_+(G) \le \theta(G)$, where

$$\theta(G) \quad = \quad \max\left\{\left\langle ee^t, x\right\rangle : \langle E_{ij}, x\rangle = 0, (i,j) \in E, \langle I, x\rangle = 1, x \succeq 0\right\},$$

$E_{ij} = e_i e_j^T + e_j e_i^T$ and e_i denotes column i of the identity matrix I.

We can define the blocks of the GCOP [$\min\ \mathbf{f(x)} + \mathbf{h_1(x)} + \mathbf{h_2(x)}$] as:

$$\begin{aligned} \mathbf{f}(\cdot) &= \langle \mathbf{ee^t}, \cdot\rangle \\ \mathbf{h_1}(\cdot) &= \delta_{\mathbf{S^n_+}}(\cdot) \text{ or } \mathbf{h_1}(\cdot) = \delta_{\mathbf{S^n_+} \cap \{\mathbf{x}:\langle \mathbf{I},\mathbf{x}\rangle = \mathbf{1}\}}(\cdot) \\ \mathbf{h_2}(\cdot) &= \delta_{\{\mathbf{x}:\langle \mathbf{I},\mathbf{x}\rangle=\mathbf{1}, \langle \mathbf{E_{ij}},\mathbf{x}\rangle = \mathbf{0}\}}(\cdot) \end{aligned}$$

Numerical Results I

Methods

Methods Compared

- BP: Boundary Point AL type method by Povh, Rendl and Wiegele. (2006).
- DSA-BD: Block-decomposition type method (2011).
- SDPNAL: Newton-CG AL type method by Zhao, Sun and Toh (2010).
- SDPAD: Alternating direction AL type method by WEN, GOLDFARB and YIN (2009).

Complementarity is satisfied ($\langle \mathbf{x}, \mathbf{s} \rangle = \mathbf{0}$) for all methods except SDPAD, and the stopping criterion is

$$\max\{\varepsilon_P, \varepsilon_D\} \leq \bar{\varepsilon} \quad \text{or} \quad \max\{\varepsilon_P, \varepsilon_D, \text{gap}\} \leq \bar{\varepsilon}$$

for a given tolerance $\bar{\varepsilon} > 0$ (e.g., $\bar{\varepsilon} = 10^{-4}, 10^{-5}, 10^{-6}$).

Numerical Results II

Pure SDPs

Table: Comparison on pure SDP problems for accuracy $\bar{\varepsilon} = 10^{-6}$

Problem Instance	$n_s \mid m$	$\max\{\varepsilon_P, \varepsilon_D\}$ BP	DSA-BD	SDPNAL	Time BP	DSA-BD	SDPNAL
RAND-0.6k40k	600|40000	9.5 -7	9.8 -7	7.2 -7	87	**41**	68
RAND-0.7k50k	700|50000	9.8 -7	9.7 -7	7.7 -7	140	**65**	88
RAND-0.6k50k	600|50000	9.6 -7	9.6 -7	6.1 -7	82	**42**	122
RAND-0.5k50k	500|50000	9.8 -7	9.6 -7	9.5 -7	89	**66**	100
RAND-0.6k60k	600|60000	9.5 -7	9.2 -7	9.2 -7	96	**57**	101
RAND-0.8k70k	800|70000	1.0 -6	9.8 -7	7.2 -7	196	**80**	131
RAND-0.7k70k	700|70000	9.6 -7	9.9 -7	6.3 -7	126	**63**	127
RAND-0.7k90k	700|90000	9.8 -7	9.5 -7	8.3 -7	202	**145**	192
RAND-0.9k100k	900|100000	1.0 -6	1.0 -6	6.3 -7	263	**102**	200
RAND-0.8k100k	800|100000	9.8 -7	9.6 -7	6.8 -7	203	**113**	250
RAND-1.0k100k	1000|100000	9.7 -7	9.7 -7	7.1 -7	450	**137**	240
RAND-0.8k110k	800|110000	9.9 -7	9.8 -7	3.8 -7	252	**165**	265
RAND-0.9k140k	900|140000	9.3 -7	9.6 -7	9.2 -7	384	**264**	348
RAND-1.0k150k	1000|150000	9.9 -7	9.7 -7	8.4 -7	459	**194**	394

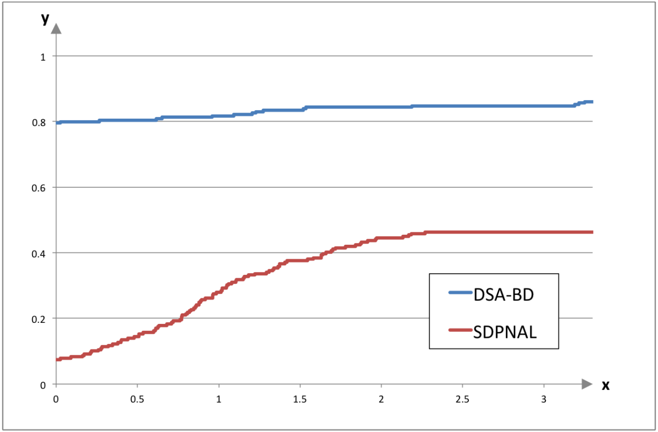
Table: DSA-BD vs. SDPNAL on mixed SDPs

Problem		$\max\{\varepsilon_P, \varepsilon_D\}$		Time	
Instance	$n_s; n_l \mid m$	DSA-BD	SDPNAL	DSA-BD	SDPNAL
BIQ-be150.8.1	151;11476|11627	1.0 -5	8.6 -6	23	100
		1.0 -6	8.2 -7	32	140
BIQ-be250.8	251;31626|31877	1.0 -5	1.6 -5*	107	444*
		1.0 -6	1.6 -5*	176	444*
BIQ-be250.10	251;31626|31877	1.0 -5	4.8 -5*	108	387*
		1.0 -6	4.8 -5*	230	387*
BIQ-bqp500-9	501;125751|126252	1.0 -5	5.7 -5*	889	2195*
		1.0 -6	5.7 -5*	1187	2195*
FAP-fap09	174;14025|15225	9.7 -6	8.8 -6	15	15
		9.9 -7	9.4 -7	19	19
FAP-fap25	2118;311044|322924	1.0 -5	7.4 -6	7444	14517
		9.9 -7	5.1 -6*	23572	29557*
FAP-fap36	4110;1112293|1154467	1.0 -5	8.5 -6	49617	29485
		9.8 -7	5.6 -6*	148855	72063*
QAP-chr20a	400;80200|80828	1.0 -4	4.1 -4*	522	2133*
		1.0 -5	4.1 -4*	1109	2133*
QAP-tai25a	625;195625|196598	1.0 -4	1.4 -4*	471	1035*
		1.0 -5	1.4 -4*	3688	1035*
QAP-bur26h	676;228826|229877	9.9 -5	3.4 -5	195	1552
		1.0 -5	2.5 -5*	1284	4419*
QAP-tai30a	900;405450|406843	1.0 -4	1.1 -4*	1059	2452*
		1.0 -5	1.1 -4*	8223	2452*

Table: DSA-BD vs. SDPAD on mixed SDPs

Problem		$\max\{\varepsilon_P, \varepsilon_D, \text{gap}\}$		Iterations (time)	
Instance	$n_s \mid m$	DSA-BD	SDPNAL	DSA-BD	SDPAD
BIQ-be150.8.9	151 | 11627	9.9 -6	1.0 -5	1348	1704
		9.9 -7	9.9 -7	2155	5793
BIQ-bqp250-6	251 | 31877	1.0 -5	1.0 -5	2141	2789
		1.0 -6	1.0 -6	3339	7066
BIQ-gka5f	501 | 126252	9.7 -6	1.0 -5	1921	6472
		9.8 -7	1.0 -6	3736	8851
BIQ-bqp500-1	501 | 126252	1.0 -5	1.0 -5	2962	6720
		1.0 -6	1.0 -6	3867	8961
FAP-fap09	174 | 15225	1.0 -5	9.9 -6	216	332
		1.0 -6	1.0 -6	370	505
FAP-fap25	2118 | 322924	1.0 -5	8.7 -6	1245 (4087)	6973
		9.8 -7	9.6 -7	3662 (11673)	9313
FAP-fap36	4110 | 1154467	1.0 -5	9.9 -6	870 (14560)	5588
		1.0 -6	9.6 -7	3095 (51379)	8716
THETA-theta102	500 | 37467	9.8 -7	9.1 -7	191	256
THETA-hamming-9-5-6	512 | 53761	8.9 -7	9.7 -7	1413	1154

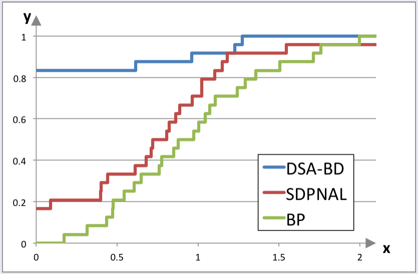
Performance Profile Plot for 229 Instances $\bar{\varepsilon} = 10^{-6}$



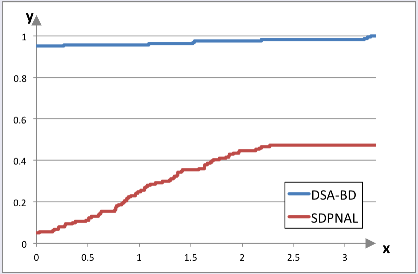
(Dolan and More 2008) A point (x, y) in the performance profile curve of a method indicates that $(100y)\%$ of the instances tested can be solved by the method in at most

$$2^x \times (\text{the cputime of any other competing method}).$$

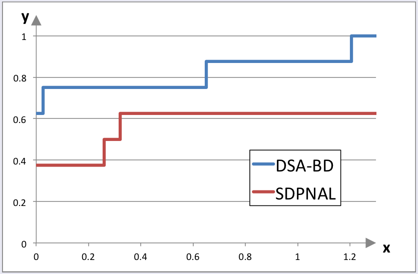
24 Random SDPs with $\bar{\varepsilon} = 10^{-6}$



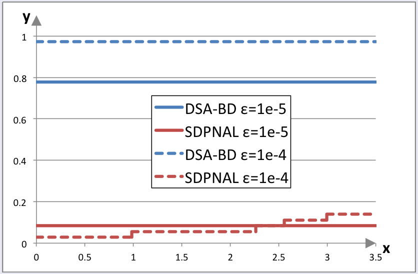
161 BIQ problems with $\bar{\varepsilon} = 10^{-6}$



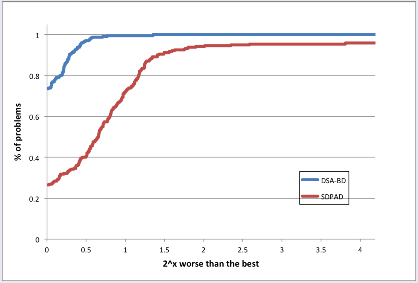
8 FAPs with $\bar{\varepsilon} = 10^{-6}$



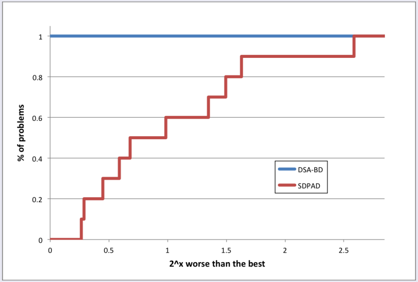
36 QAPs with $\bar{\varepsilon} = 10^{-4}, 10^{-5}$



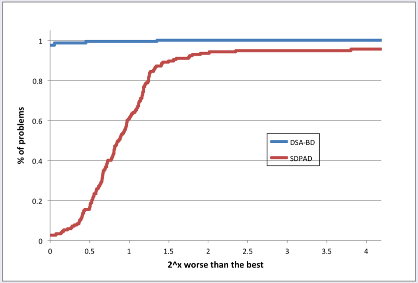
242 SDP instances $\bar{\varepsilon} = 10^{-6}$



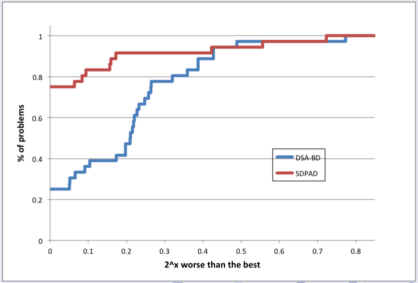
10 FAPs with $\bar{\varepsilon} = 10^{-6}$



155 BIQ problems with $\bar{\varepsilon} = 10^{-6}$



36 THETA problems with $\bar{\varepsilon} = 10^{-4}$



Conclusions:

- We have a nice framework of a method originally intended for general inclusion problems (MIPs) to general convex optimization problems .
- In order to make the algorithm competitive we had to make use of various modifications:
  - Dynamic scaled adapted step-size (DSA).
  - Good block definitions and suitable initializations.
- BD type methods are not only simple but can solve large scale SDPs very efficiently.

Further Issues:

- How does it perform on other classes of problems, e.g., convex optimization problems with $h \neq \delta_K$?
- Find more structure specific implementations (e.g., combining cones with affine constraints).
- Dynamic choice of the order of solving each block.

Conclusions:

- We have a nice framework of a method originally intended for general inclusion problems (MIPs) to general convex optimization problems .
- In order to make the algorithm competitive we had to make use of various modifications:
  - Dynamic scaled adapted step-size (DSA).
  - Good block definitions and suitable initializations.
- BD type methods are not only simple but can solve large scale SDPs very efficiently.

Further Issues:

- How does it perform on other classes of problems, e.g., convex optimization problems with $h \neq \delta_K$?
- Find more structure specific implementations (e.g., combining cones with affine constraints).
- Dynamic choice of the order of solving each block.

Thank you ...

Find SDP application paper at Optimization-Online #3032:
http://www.optimization-online.org/DB_HTML/2011/05/3032.html
Code (SDP only):
http://www.isye.gatech.edu/~cod3/

- A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. J. of Math. Imaging and Vision, 40 (2010), pp. 120-145.
- J. Malick, J. Povh, F. Rendl and A. Wiegele. *Regularization methods for semidefinite programming*. SIAM J. Optim., 20 (2009), pp. 336-356.
- R. D. C. Monteiro and B. F. Svaiter. *On the complexity of the hybrid proximal projection method for the iterates and the ergodic mean*. SIAM J. Optim., 20 (2010), pp. 2755-2787. on Optimization).
- R. D. C. Monteiro and B. F. Svaiter: *Iteration-complexity of block-decomposition algorithms and the alternating minimization augmented Lagrangian method*, working paper, School of ISyE, Georgia Tech, USA, August 2010 (submitted to Mathematical Programming).
- R.D.C. Monteiro, C. Ortiz and B.F. Svaiter Implementation of a block-decomposition algorithm for solving large-scale conic semidefinite programming problems, preprint, 2011.
- J. Povh, F. Rendl, and A. Wiegele, *A boundary point method to solve semidefinite programs, Computing*, 78 (2006), pp. 277–286.
- M. V. Solodov and B. F. Svaiter. *A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator*. Set-Valued Anal., 7(4):323–345, 1999.
- Z. Wen, D. Goldfarb and W. Yin. *Alternating direction augmented Lagrangian methods for semidefinite programming*. Math. Prog. Computation, 2 (2010), pp. 203-230.
- X.-Y. Zhao, D. Sun, and K.-C. Toh. *A Newton-CG augmented Lagrangian method for semidefinite programming*. SIAM J. Optim., 20 (2010), pp. 1737–1765.