An Operator Perspective on PDHG with Application to Quadratic Programming

Tao Jiang
In collaboration with Walaa Moursi (UWaterloo), Stephen Vavasis (UWaterloo)

School of ORIE,
Cornell University

June 2, 2023

Primal-Dual Hybrid Gradient (PDHG)

Given the following convex problem, write its saddle form as follows

$$\min_{x\in\mathbb{R}^n} f(Ax) + g(x) \qquad \Leftrightarrow \qquad \min_{x\in\mathbb{R}^n}\sup_{y\in\mathbb{R}^m} L(x,y) := \langle y, Ax\rangle - f^*(y) + g(x),$$

where f, g are convex and lower semicontinuous.

The PDHG algorithm is to alternate proximal method in x, y with approximate extragradient:

$$\begin{pmatrix} x^+ \\ y^+ \end{pmatrix} = \begin{pmatrix} \mathrm{Prox}_{\tau g}(x - \tau A^* y) \\ \mathrm{Prox}_{\sigma f^*}(y + \sigma A \underbrace{(2x^+ - x)}_{=x^+ + (x^+ - x)}) \end{pmatrix} \qquad \text{(PDHG)}$$

$$\Leftrightarrow \qquad \begin{pmatrix} x - x^+ \\ y - y^+ \end{pmatrix} \in \begin{pmatrix} \tau(A^* y + \partial g(x^+)) \\ \sigma(Ax^+ + (x^+ - x) + \partial f^*(y^+)) \end{pmatrix} \in \begin{pmatrix} \tau \cdot \frac{\partial L(x^+, y)}{\partial x} \\ \sigma \cdot \frac{\partial L(x^+ + (x^+ - x), y^+)}{\partial y^+} \end{pmatrix}$$

History of PDHG

- Esser et al., Pock, Cremers, Bischof and Chambolle proposed PDHG at the same time.
- Attouch, Briceño-Arias and Combettes introduced a similar framework with different splitting.
- O'Connor and Vandenberghe showed that PDHG is equivalent to Douglas-Rachford iteration.
- Applegate, Díaz, Lu, Lubin et al found that, among first-order methods, PDHG appears to be the best in practice for LP.

Applegate et al. (work presented by Mateo earlier) focus on PDHG for LP:

- Characterize the behavior of PDHG in LP;
- Detect infeasibility using PDHG iterates for LP.

Can we characterize the behavior of PDHG and detect infeasibility for other problem classes?

Key Contributions

Can we characterize the behavior of PDHG and detect infeasibility for other problem classes?

We partially answer the question through the lens of operator theory and convex optimization:

General Convex Problems:

(i) Offer insights of the PDHG operator and its iterative behavior;
(ii) Ongoing: seek problem structures that allow full characterization of the behavior of PDHG and infeasibility detection.

Quadratic Programming (QP):

(i) Fully characterize the behavior of PDHG;
(ii) Detect infeasibility and establish certificates from PDHG iterates.

PDHG for Convex Problems

Recall the convex problem where f, g are convex, l.s.c. and $A : \mathbb{R}^n \to \mathbb{R}^m$.

$$\min_{x\in\mathbb{R}^n} f(Ax) + g(x) \qquad \Leftrightarrow \qquad \min_{x\in\mathbb{R}^n} \sup_{y\in\mathbb{R}^m} L(x,y) := \langle y, Ax\rangle - f^*(y) + g(x),$$

We derive the operator form of PDHG update as follows:

$$\begin{pmatrix} x^+ \\ y^+ \end{pmatrix} = \begin{pmatrix} \mathrm{Prox}_{\tau g}(x - \tau A^* y) \\ \mathrm{Prox}_{\sigma f^*}(y + \sigma A(2x^+ - x)) \end{pmatrix} = T\begin{pmatrix} x \\ y \end{pmatrix} = \left(M + \begin{pmatrix} \partial g \\ \partial f^* \end{pmatrix} + S\right)^{-1} M \begin{pmatrix} x \\ y \end{pmatrix},$$

where $S := \begin{pmatrix} 0 & A^* \\ -A & 0 \end{pmatrix}$ and $M := \begin{pmatrix} \frac{1}{\tau}\,\mathrm{Id} & -A^* \\ -A & \frac{1}{\sigma}\,\mathrm{Id} \end{pmatrix}$.

Define $v = \begin{pmatrix} v_x \\ v_y \end{pmatrix}$ to be the minimal M-norm element in $\overline{\mathrm{ran}}\,(\mathrm{Id} - T)$:

$$\begin{aligned} \min \quad & \|v\|_M \\ \text{s.t.} \quad & v \in \overline{\mathrm{ran}}\,(\mathrm{Id} - T), \end{aligned}$$

Interpretation of $\operatorname{ran}(\operatorname{Id} - T)$

$$\min_{x\in\mathbb{R}^n}\sup_{y\in\mathbb{R}^m} L(x,y) := \langle y, Ax\rangle - f^*(y) + g(x), \qquad \begin{pmatrix} x^+ \\ y^+ \end{pmatrix} = T\begin{pmatrix} x \\ y \end{pmatrix} = \left(M + \begin{pmatrix} \partial g \\ \partial f^* \end{pmatrix} + S\right)^{-1} M \begin{pmatrix} x \\ y \end{pmatrix},$$

We could measure the progress of the PDHG algorithm by monitoring

$$\begin{pmatrix} x_k - x_{k+1} \\ y_k - y_{k+1} \end{pmatrix} = (\operatorname{Id} - T)\begin{pmatrix} x_k \\ y_k \end{pmatrix} \in \operatorname{ran}(\operatorname{Id} - T).$$

- If $0 \in \operatorname{ran}(\operatorname{Id} - T)$, then T admits fixed points and there is hope for $\{(x_k, y_k)\}_{k\in\mathbb{N}}$ to converge.
- If $0 \notin \operatorname{ran}(\operatorname{Id} - T)$, then T does not have a fixed point and
  (i) $\{(x_k, y_k)\}_{k\in\mathbb{N}}$ diverges to infinity in norm;
  (ii) the primal-dual problem does not have a solution.

Remark

For QP, the PDHG operator T has a fixed point ($0 \in \operatorname{ran}(\operatorname{Id} - T)$) iff KKT admits a solution.

Asymptotic Behavior of (x_k, y_k)

The observations of PDHG iterates so far are not quite quantifiable. However, we can analyze the asymptotic behavior of $\{(x_k, y_k)\}_{k\in\mathbb{N}}$ with more assumptions on T:

Fact

If T is firmly nonexpansive, then

$$\begin{pmatrix} x_k - x_{k+1} \\ y_k - y_{k+1} \end{pmatrix} \to \begin{pmatrix} v_x \\ v_y \end{pmatrix},$$

where $v = \operatorname{argmin}_{u\in\overline{\operatorname{ran}}\,(\mathrm{Id}-T)} \|u\|_M$ *is the minimal norm element of* $\overline{\operatorname{ran}}\,(\mathrm{Id}-T)$.

- If $v_x = 0$, then $\{x_k\}_{k\in\mathbb{N}}$ converges.
- If $v_x \neq 0$, then $x_k - x_{k+1} - v_x \to 0$, hence
  - (i) $\{x_k\}_{k\in\mathbb{N}}$ diverges to infinity in norm;
  - (ii) the primal-dual problem does not have a solution.
- The same result holds for v_y.

Remark

v can be interpreted as a "gap" vector that measures how far the primal-dual problem is from having a solution.

PDHG Operator T

Definition (Firm nonexpansiveness)

An operator T is firmly nonexpansive if

$$(\forall x, y \in \mathbb{R}^d) \qquad \|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leq \|x - y\|^2.$$

Fact (Great things about firmly nonexpansive operators)

Suppose $T\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$ is firmly nonexpansive. Let $\{(x_k, y_k)\}_{k\in\mathbb{N}}$ be a sequence generated by $\begin{pmatrix} x^+ \\ y^+ \end{pmatrix} = T \begin{pmatrix} x \\ y \end{pmatrix}$.

(a) *The sequence $\{(x_k, y_k)\}_{k\in\mathbb{N}}$ satisfies*
  (i) (Pazy): $\frac{1}{k}\begin{pmatrix} x_k \\ y_k \end{pmatrix} \to -\begin{pmatrix} v_x \\ v_y \end{pmatrix}$ *and*
  (ii) (Bruck–Reich): $\begin{pmatrix} x_k - x_{k+1} \\ y_k - y_{k+1} \end{pmatrix} \to \begin{pmatrix} v_x \\ v_y \end{pmatrix}$

(b) *Further assume* $\mathrm{Fix}(v + T) \neq \emptyset$ *(equivalently $v \in \mathrm{ran}\,(\mathrm{Id} - T)$). The sequence $(x_k + kv_x, y_k + kv_y)_{k\in\mathbb{N}}$ is Fejér monotone[1] with respect to $\mathrm{Fix}(v + T)$, hence bounded.*

---

[1] $\{x_k\}_{k\in\mathbb{N}}$ is Fejér monotone with respect to C if $(\forall x \in C)(\forall k \in \mathbb{N}) \quad \|x_{k+1} - x\| \leq \|x_k - x\|$

PDHG Operator T

$$\begin{pmatrix} x^+ \\ y^+ \end{pmatrix} = \begin{pmatrix} \mathrm{Prox}_{\tau g}(x - \tau A^* y) \\ \mathrm{Prox}_{\sigma f^*}(y + \sigma A(2x^+ - x)) \end{pmatrix} = T \begin{pmatrix} x \\ y \end{pmatrix} = \left(M + \begin{pmatrix} \partial g \\ \partial f^* \end{pmatrix} + S \right)^{-1} M \begin{pmatrix} x \\ y \end{pmatrix}.$$

Claim

If $\tau\sigma\|A\|_2^2 < 1$, then $T = \left(M + \begin{pmatrix} \partial g \\ \partial f^* \end{pmatrix} + S \right)^{-1} M \colon \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ is firmly nonexpansive w.r.t. $\|\cdot\|_M$, where $M := \begin{pmatrix} \frac{1}{\tau}\mathrm{Id}_X & -A^* \\ -A & \frac{1}{\sigma}\mathrm{Id}_Y \end{pmatrix} \succ 0$ and $S := \begin{pmatrix} 0 & A^* \\ -A & 0 \end{pmatrix}$.

Proof.

Key fact: If F is maximally monotone, then $(\mathrm{Id} + F)^{-1}$ is firmly nonexpansive.

$$T = \left(M + \begin{pmatrix} \partial g \\ \partial f^* \end{pmatrix} + S \right)^{-1} M = (\mathrm{Id} + M^{-1}(\partial F + S))^{-1} M^{-1} M = (\mathrm{Id} + \underbrace{\overbrace{M^{-1}}^{\text{positive definite}} \Bigg(\overbrace{\begin{pmatrix} \partial g \\ \partial f^* \end{pmatrix} + S}^{\text{maximally monotone}} \Bigg)}_{\text{maximally monotone w.r.t. } M\text{-norm}})^{-1}.$$

$\square$

$\overline{\mathrm{ran}}\,(\mathrm{Id} - T)$ AND ITS MINIMAL M-NORM ELEMENT v

Recall that $v = \mathrm{argmin}_{u \in \overline{\mathrm{ran}}\,(\mathrm{Id} - T)} \|u\|_M$ encodes feasibility information and the iterative behavior of the firmly nonexpansive PDHG operator $T = (\mathrm{Id} + M^{-1}(\partial F + S))^{-1}$.

The vector v and $\overline{\mathrm{ran}}\,(\mathrm{Id} - T)$ enjoy the following properties:

(i) $\mathrm{ran}\,(\mathrm{Id} - T)$ is nearly convex, which implies $\overline{\mathrm{ran}}\,(\mathrm{Id} - T)$ is convex;
(ii) If $\mathrm{ran}\,(\mathrm{Id} - T)$ is closed, then $\mathrm{ran}\,(\mathrm{Id} - T) = \overline{\mathrm{ran}}\,(\mathrm{Id} - T)$ is convex and $v \in \mathrm{ran}\,(\mathrm{Id} - T)$;
(iii) $\mathrm{ran}\,(\mathrm{Id} - T) = M^{-1}\left(\mathrm{ran}\left(\begin{pmatrix} \partial g(x) \\ \partial f^*(y) \end{pmatrix} + S\right)\right)$;
(iv) $\overline{\mathrm{ran}}\,(\mathrm{Id} - T) = M^{-1}\left(\overline{\mathrm{ran}}\left(\begin{pmatrix} \partial g(x) \\ \partial f^*(y) \end{pmatrix} + S\right)\right)$.

Example (i) Let K be a closed convex cone in $\mathbb{R}^m$:

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & g(x) \\ \text{subject to} & Ax - b \in K, \end{array} \qquad (1)$$

$$f = \iota_K(\cdot - b) \quad , \qquad f^* = \iota_{K^\ominus}(\cdot) + \langle b, \cdot \rangle$$

$$\begin{pmatrix} \partial g(x) \\ \partial f^*(y) \end{pmatrix} = \partial g(x) \times (\mathcal{N}_{K^\ominus}(y) + b)$$

$$S = \begin{pmatrix} 0 & A^* \\ -A & 0 \end{pmatrix}$$

Example (ii) In addition, let $H \in \mathbb{R}^n \to \mathbb{R}^n$ be linear, monotone and self-adjoint.

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & \frac{1}{2}\langle x, Hx \rangle + \langle c, x \rangle \\ \text{s.t.} & Ax - b \in K, \end{array} \qquad (2)$$

$$\begin{pmatrix} \partial g(x) \\ \partial f^*(y) \end{pmatrix} = (Hx + c) \times (\mathcal{N}_{K^\ominus}(y) + b)$$

$$S = \begin{pmatrix} 0 & A^* \\ -A & 0 \end{pmatrix}$$

QUADRATIC PROGRAMMING

Consider the following QP and the corresponding PDHG update with step sizes $\tau = \sigma = 1$

$$\begin{aligned} \min \quad & \tfrac{1}{2}x^T H x + c^T x \\ \text{s.t.} \quad & Ax - b \le 0, \end{aligned} \qquad \text{(PQP)} \qquad\qquad \begin{aligned} x^+ &:= (H + \mathrm{Id})^{-1}(x - A^T y - c), \\ y^+ &:= \mathrm{Proj}_{\mathbb{R}^m_+}\left(y + A(2x^+ - x) - b\right). \end{aligned} \qquad \text{(PDHG)}$$

The operator form of PDHG update is simply

$$\begin{pmatrix} x^+ \\ y^+ \end{pmatrix} = T_{\mathrm{QP}} \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} (H + \mathrm{Id})^{-1}(x - A^T y - c) \\ \mathrm{Proj}_{\mathbb{R}^m_+}\left(y + A(2(H + \mathrm{Id})^{-1}(x - A^T y - c) - x) - b\right) \end{pmatrix}. \tag{3}$$

Define $v = \begin{pmatrix} v_x \\ v_y \end{pmatrix}$ to be the minimal M-norm element in $\overline{\mathrm{ran}}\,(\mathrm{Id} - T_{QP})$:

$$\begin{aligned} \min \quad & \|v\|_M \\ \text{s.t.} \quad & v \in \overline{\mathrm{ran}}\,(\mathrm{Id} - T_{QP}), \end{aligned}$$

PDHG for QP: $\overline{\text{ran}}\,(\text{Id} - T_{QP})$ and minimal M-norm element v

Recall that $v = \text{argmin}_{u \in \overline{\text{ran}}\,(\text{Id} - T_{QP})} \|u\|_M$ encodes feasibility information and the iterative behavior of the firmly nonexpansive PDHG operator T_{QP}.
The set $\text{ran}\,(\text{Id} - T_{QP})$ satisfies the following properties

Lemma

(i) $\text{ran}\,(\text{Id} - T_{QP})$ *is a union of finitely many polyhedral*[2] *sets.*

(ii) $\text{ran}\,(\text{Id} - T_{QP})$ *is convex and closed (so* $\overline{\text{ran}}\,(\text{Id} - T_{QP}) = \text{ran}\,(\text{Id} - T_{QP})$*). In fact,* $\text{ran}\,(\text{Id} - T_{QP})$ *is polyhedral.*

(iii)

$$\text{ran}\,(\text{Id} - T_{QP}) = \left\{ (u_x, u_y) : \quad \begin{array}{l} u_x - Hu_x = -Hw + A^T y + c, \\ u_y \leq y, \\ u_y \leq Aw + b. \end{array} \right\}.$$

As a consequence:

$$\begin{array}{rl} \min & \|v\|_M \\ \text{s.t.} & v \in \text{ran}\,(\text{Id} - T_{QP}), \end{array} \quad \Leftrightarrow \quad \begin{array}{rl} \min & v_x^T v_x - v_y^T A v_x + v_y^T v_y \\ \text{s.t.} & v_x - Hv_x = -Hw + A^T y + c, \\ & v_y \leq y, \\ & v_y \leq Aw + b. \end{array}$$

[2]Let $C \subseteq X$. We say that C is *polyhedral* if C is the intersection of finitely many halfspaces.

PDHG for QP: Infeasibility Detection

$$\begin{array}{ll} \min & \frac{1}{2}x^T Hx + c^T x \\ \text{s.t.} & Ax - b \le 0, \end{array} \qquad \text{(PQP)}$$

$$\begin{array}{ll} \min & \frac{1}{2}z^T Hz + b^T y \\ \text{s.t.} & A^T y + c = Hz, \\ & y \ge 0. \end{array} \qquad \text{(DQP)}$$

Theorem

(PQP) is infeasible if and only if $v_y \neq 0$, and in this case, v_y is an infeasibility certificate for (PQP).

Proof.

Goal: Prove that $v_y \neq 0 \Leftrightarrow A^T v_y = 0, b^T v_y > 0$ (Theorem follows using Farkas lemma).

Step 1: Write the minimal norm problem that finds (v_x, v_y) and its Lagrangian dual:

$$\begin{array}{ll} \min & \|v\|_M^2 = v_x^T v_x - v_y^T A v_x + v_y^T v_y \\ \text{s.t.} & v_x - Hv_x = -Hw + A^T y + c, \\ & v_y \le y, \\ & v_y \le Aw + b. \end{array}$$

$$\begin{array}{ll} \max & -\lambda^T \lambda/2 + \xi^T A\lambda - \xi^T \xi/2 - c^T \lambda - b^T \xi \\ \text{s.t.} & H\lambda - A^T \xi = 0, \\ & A\lambda + \mu = 0, \\ & \mu \ge 0, \\ & \xi \ge 0. \end{array}$$

Step 2: KKT condition implies $A^T v_y = 0$ and $b^T v_y = v_y^T v_y$.

Step 3: $v_y \neq 0 \Leftrightarrow b^T v_y > 0 \Leftrightarrow$ (PQP) is infeasible (by Farkas Lemma).

PDHG for QP: Infeasibility Detection

$$\begin{array}{ll} \min & \frac{1}{2}x^T Hx + c^T x \\ \text{s.t.} & Ax - b \le 0, \end{array} \qquad \text{(PQP)}$$

$$\begin{array}{ll} \min & \frac{1}{2}z^T Hz + b^T y \\ \text{s.t.} & A^T y + c = Hz, \\ & y \ge 0. \end{array} \qquad \text{(DQP)}$$

Theorem

(DQP) is infeasible if and only if $v_x \neq 0$, and in this case, v_x is an infeasibility certificate for (DQP).

Proof.

Goal: Prove that $v_x \neq 0 \Leftrightarrow Av_x \le 0, (Hz - c)^T v_x > 0\ \forall z$ (rest follows from Farkas lemma).

Step 1: Write the minimal norm problem that finds (v_x, v_y) and its Lagrangian dual:

$$\begin{array}{ll} \min & \|v\|_M = v_x^T v_x - v_y^T A v_x + v_y^T v_y \\ \text{s.t.} & v_x - Hv_x = -Hw + A^T y + c, \\ & v_y \le y, \\ & v_y \le Aw + b. \end{array}$$

$$\begin{array}{ll} \max & -\lambda^T \lambda / 2 + \xi^T A \lambda - \xi^T \xi / 2 - c^T \lambda - b^T \xi \\ \text{s.t.} & H\lambda - A^T \xi = 0, \\ & A\lambda + \mu = 0, \\ & \mu \ge 0, \\ & \xi \ge 0. \end{array}$$

Step 2: KKT condition implies $Av_x \le 0$ and $(Hz - c)^T v_x \ge v_x^T v_x$ for any z.

Step 3: $v_x \neq 0 \Leftrightarrow (Hz - c)^T v_x > 0\ \forall z \Leftrightarrow$ (DQP) is infeasible (by Farkas Lemma).

PDHG for QP: Dynamic Behavior

Theorem

Let $\{(x_k, y_k)\}_{k\in\mathbb{N}}$ *denote the sequence generated by PDHG update* $\binom{x^+}{y^+} = T_{\text{QP}}\binom{x}{y}$*. Then* $\exists\alpha \geq 0$

$$(x_k + kv_x, y_k + kv_y) \to (x^*, y^*) \in \alpha v + \text{Fix}(v + T_{QP}),$$

where $v = \text{argmin}_{u\in\text{ran}\,(\text{Id}\,-T_{QP})} \|u\|_M$ *is the minimal M-norm element of* $\text{ran}\,(\text{Id}\,-T_{QP})$.

Remark

Analogous results are only known for any f.n.e. affine operator and PDHG operator in LP. It is unclear what other operator structure also admits such convergence behavior.

Consequently, v fully characterizes the behavior of PDHG iterates as follows:

(i) If $v_x = 0$, $v_y = 0$, then $(x_k, y_k) \to (x^*, y^*)$, which is the primal-dual solution.

(ii) If $v_x \neq 0$, $v_y \neq 0$, then (x_k, y_k) diverges along the ray $\{-\alpha(v_x, v_y)\}_{\alpha\geq 0}$.

(iii) If $v_x = 0$, $v_y \neq 0$, then (x_k, y_k) diverges along the ray $\{-\alpha(0, v_y)\}_{\alpha\geq 0}$.

(iv) If $v_x \neq 0$, $v_y = 0$, then (x_k, y_k) diverges along the ray $\{-\alpha(v_x, 0)\}_{\alpha\geq 0}$.

PDHG for QP: Dynamic Behavior

Theorem

Let $\{(x_k, y_k)\}_{k\in\mathbb{N}}$ *denote the sequence generated by PDHG update* $\binom{x^+}{y^+} = T_{\mathrm{QP}}\binom{x}{y}$*. Then* $\exists \alpha \geq 0$

$$(x_k + kv_x, y_k + kv_y) \to (x^*, y^*) \in \alpha v + \mathrm{Fix}(v + T_{QP}),$$

where $v = \operatorname{argmin}_{u\in \operatorname{ran}(\mathrm{Id} - T_{QP})} \|u\|_M$ *is the minimal M-norm element of* $\operatorname{ran}(\mathrm{Id} - T_{QP})$.

Proof.

Goal: Prove $(\exists K \in \mathbb{N}, \alpha \geq 0)$, $\{(x_{k+K} + (k+K)v_x, y_{k+K} + (k+K)v_y)\}_{k\in\mathbb{N}} \to \alpha v + \mathrm{Fix}(v + T_{QP})$.

Step 1: $(\forall k \in \mathbb{N})$, $\begin{pmatrix} x_{k+K} + kv_x \\ y_{k+K} + kv_y \end{pmatrix} = (v + T_{\mathrm{QP}})^k \begin{pmatrix} x_K \\ y_K \end{pmatrix}$, and $\mathrm{Fix}(v + T_{\mathrm{QP}}) \neq \emptyset$.

Step 2: $\begin{pmatrix} x_{k+K} + (k+K)v_x \\ y_{k+K} + (k+K)v_y \end{pmatrix} \to (v + T_{\mathrm{QP}})^k \begin{pmatrix} x_K \\ y_K \end{pmatrix} + Kv \to \mathrm{Fix}(v + T_{QP}) + Kv$

Step 3: Since $\mathrm{Fix}(v + T_{QP}) = R_- \cdot v + \mathrm{Fix}(v + T_{\mathrm{QP}})$, there exists $\alpha > 0$ such that

$$\mathrm{Fix}(v + T_{QP}) + Kv - \alpha v = \mathrm{Fix}(v + T_{QP}).$$

PDHG for QP: Summary

Given quadratic programming and PDHG update,

$$\begin{array}{rl} \min & \frac{1}{2}x^T H x + c^T x \\ \text{s.t.} & Ax - b \le 0, \end{array} \qquad \text{(PQP)} \qquad \begin{array}{rl} x^+ & := (H + \mathrm{Id})^{-1}(x - A^T y - c), \\ y^+ & := \mathrm{Proj}_{\mathbb{R}^m_+}\left(y + A(2x^+ - x) - b\right), \end{array} \qquad \text{(PDHG)}$$

we observe the following four scenarios:

(i) $(x_k, y_k) \to (x^*, y^*)$: $v_x = 0, v_y = 0$;
both problems are feasible; (x^*, y^*) is an optimal primal-dual solution.

(ii) $(x_k + kv_x, y_k + kv_y) \to (x^*, y^*)$: $v_x \neq 0, v_y \neq 0$;
both problems are infeasible; v_x, v_y are infeasibility certificates for (DQP), (PQP) respectively.

(iii) $(x_k, y_k + kv_y) \to (x^*, y^*)$: $v_x = 0, v_y \neq 0$;
(PQP) is infeasible; v_y is an infeasibility certificate for (PQP).

(iv) $(x_k + kv_x, y_k) \to (x^*, y^*)$: $v_x \neq 0, v_y = 0$;
(DQP) is infeasible; v_x is an infeasibility certificate for (DQP).

Example: Really Simple SOCP

Consider the following SOCP:

$$\begin{array}{ll} \min & \begin{pmatrix} 0 \\ c \end{pmatrix}^T \begin{pmatrix} x_1 \\ \bar{x} \end{pmatrix} \\ \text{s.t.} & \frac{1}{2}x_1 = r, \\ & \begin{pmatrix} x_1 \\ \bar{x} \end{pmatrix} \in C_2^{d+1}. \end{array} \qquad \text{(PSOCP)}$$

$$\begin{array}{ll} \max & ry \\ \text{s.t.} & \begin{pmatrix} 0 \\ c \end{pmatrix} - y \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} \succeq 0 \end{array} \qquad \text{(DSOCP)}$$

The following observations hold:

(i)

$$\operatorname{ran}(\mathrm{Id} - T_{SOCP}) = \left\{ (u_x, u_y) : \begin{array}{l} u_x \preceq y \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ c \end{pmatrix}, \\ w \preceq u_x, \\ u_y = \frac{1}{2}w_1 + r, \end{array} \right\}.$$

which is closed and convex.

(ii) $\mathrm{Fix}(v + T) \neq \emptyset$, equivalently, $v \in \operatorname{ran}(\mathrm{Id} - T)$.

(iii) (PSOCP) is infeasible iff $v_y \neq 0$, and in this case, v_y is an infeasibility certificate for (PSOCP).

(iv) The sequence $(x_k + kv_x, y_k + kv_y)_{k\in\mathbb{N}}$ is Fejér monotone w.r.t. $\mathrm{Fix}(v + T)$, hence bounded.

Open Problems and Future Directions

- Under what condition is $\{(x_k + kv_x, y_k + kv_y)\}_{k\in\mathbb{N}}$ convergent?
- Under what condition is $v \in \operatorname{ran}(\operatorname{Id} - T)$?