# Bridging Primal and Primal-Dual Analyses: From Error Bounds to Quadratic Growth

**Huikang Liu**[*] Jiajin Li[†]

## 1 Introduction

We consider the following general convex optimization:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) \\ \text{subject to} \quad & Ax \le b. \end{aligned} \tag{1}$$

where $G(x) = (g_1(x), \cdots, g_m(x))$, both $f$ and $g_i, i \in [m]$ are smooth and convex functions, and $h(x)$ is a convex but nonsmooth function whose proximal mapping is easy to compute. We consider its primal-dual formulation:

$$\min_x \max_{y\ge 0} \mathcal{L}(x,y) = f(x) + h(x) + G(x)^T y. \tag{2}$$

Let $z = (x, y) \in \mathcal{Z} = \mathbb{R}^n \times \mathbb{R}^m_+$ and $\mathcal{Z}^\star$ denote the set of the optimal solutions to problem (2), i.e., the set of solutions satisfying the KKT system. Let

$$\mathcal{X}^\star := \arg\min_x \max_{y\ge 0} \mathcal{L}(x,y), \qquad \mathcal{Y}^\star := \arg\max_{y\ge 0} \min_x \mathcal{L}(x,y). \tag{3}$$

Then, according to [4, Lemma 36.2], we know that a saddle point $(x^\star, y^\star) \in \mathcal{Z}^\star$ if and only if $x^\star \in \mathcal{X}^\star$ and $y^\star \in \mathcal{Y}^\star$. That is, $\mathcal{Z}^\star = \mathcal{X}^\star \times \mathcal{Y}^\star$.

## 2 Error Bound

We assume that the following error-bound condition holds:

**Assumption 1.** *There exists some constant $\tau > 0$ such that, for any $x \in \mathcal{X}$ and $y^\star \in \mathcal{Y}^\star$ with $\|(x, y^\star)\| \le R$, we have*

$$\tau \cdot \textit{dist}(x, \mathcal{X}^\star) \le \|[G(x)]_+\| + \|\nabla f(x) + y^\star \nabla G(x)^T\|_2. \tag{4}$$

**Lemma 1.** *Assumption 1 holds for the case $f(x) = h(Kx) + c^T x$ where $h$ is stronly convex and $G(x) = Ax - b$.*

The KKT optimality conditions is given by

$$\begin{aligned} \operatorname{prox}_{\frac{1}{L}h}\left(x - \frac{1}{L}(\nabla f(x) + \nabla G(x)^T y)\right) - x = 0, \\ G(x) \le 0, \\ G(x)^T y = 0. \end{aligned} \tag{KKT}$$

It is straight-forward to check that the complementary slack is indeed equivalent to the following condition:

$$y - [y + \eta G(x)]_+ = 0, \text{ for all } \eta > 0. \tag{5}$$

[*]Shanghai Jiao Tong University
[†]University of British Columbia

Preprint.

**Definition 1.** *We call the following term as the scaled KKT residual of* (1) *at* $z \in \mathcal{Z}$*:*

$$F_\xi(z) = F_\xi(x,y) = \begin{bmatrix} [G(x)]_+ \\ \nabla f(x) + \nabla G(x)^T y \\ y - [y + \frac{1}{\xi}G(x)]_+ \end{bmatrix} \tag{6}$$

Through this paper, we assume the following error bound condition holds:

**Proposition 1.** *There exists some constant* $\gamma_\xi > 0$ *such that, for any* $z \in \mathcal{Z}$ *with* $\|z\| \le R$*, we have*

$$\gamma_\xi \cdot \textit{dist}(z, \mathcal{Z}^\star) \le \|F_\xi(z)\|. \tag{7}$$

The smoothed duality gap, proposed in [3], is defined below:

**Definition 2.** *For any* $\xi > 0$ *and* $z, \dot{z} \in \mathcal{Z}$*, we define the smoothed duality gap at* $z$ *centered at* $\dot{z}$ *as*

$$G_\xi(z; \dot{z}) = \max_{\hat{z}=(\hat{x},\hat{y})\in\mathcal{Z}} \left\{ \mathcal{L}(x,\hat{y}) - \mathcal{L}(\hat{x},y) - \frac{\xi}{2}\|\hat{z} - \dot{z}\|^2 \right\}. \tag{8}$$

## 3 From Error Bound to Quadratic Growth

The following simple lemma implies that we can split the (primal-dual) smoothed duality gap into the summation of the primal and dual parts, either of which has a simplified form.

**Lemma 2.** *For any* $z \in \mathcal{Z}$ *and* $z^\star = (x^\star, y^\star) \in \mathcal{Z}^\star$*, the smoothed duality gap at* $(x,y)$ *can be written as*

$$G_\xi((x,y);z^\star) = G_\xi((x,y^\star);z^\star) + G_\xi((x^\star,y);z^\star).$$

*Proof.* The proof follows from the definition of the smoothed duality gap.

$$\begin{aligned} G_\xi((x,y);z^\star) &= \max_{\hat{x},\hat{y}\ge 0} \left\{ \mathcal{L}(x,\hat{y}) - \mathcal{L}(\hat{x},y) - \frac{\xi}{2}\|\hat{x}-x^\star\|^2 - \frac{\xi}{2}\|\hat{y}-y^\star\|^2 \right\} \\ &= \max_{\hat{x},\hat{y}\ge 0} \left\{ \mathcal{L}(x,\hat{y}) - \mathcal{L}(x^\star,y^\star) + \mathcal{L}(x^\star,y^\star) - \mathcal{L}(\hat{x},y) - \frac{\xi}{2}\|\hat{x}-x^\star\|^2 - \frac{\xi}{2}\|\hat{y}-y^\star\|^2 \right\} \\ &= \max_{\hat{y}\ge 0} \left\{ \mathcal{L}(x,\hat{y}) - \mathcal{L}(x^\star,y^\star) - \frac{\xi}{2}\|\hat{y}-y^\star\|^2 \right\} + \max_{\hat{x}} \left\{ \mathcal{L}(x^\star,y^\star) - \mathcal{L}(\hat{x},y) - \frac{\xi}{2}\|\hat{x}-x^\star\|^2 \right\} \end{aligned} \tag{9}$$

According to the optimality condition of $(x^\star, y^\star)$, we have

$$\mathcal{L}(x^\star,y^\star) = \min_{\tilde{x}} \mathcal{L}(\tilde{x},y^\star) = \min_{\tilde{x}} \left\{ \mathcal{L}(\tilde{x},y^\star) + \frac{\xi}{2}\|\tilde{x}-x^\star\|^2 \right\},$$

which impies that

$$\begin{aligned} &\max_{\hat{y}\ge 0} \left\{ \mathcal{L}(x,\hat{y}) - \mathcal{L}(x^\star,y^\star) - \frac{\xi}{2}\|\hat{y}-y^\star\|^2 \right\} \\ =&\max_{\tilde{x},\hat{y}\ge 0} \left\{ \mathcal{L}(x,\hat{y}) - \mathcal{L}(\tilde{x},y^\star) - \frac{\xi}{2}\|\tilde{x}-x^\star\|^2 - \frac{\xi}{2}\|\hat{y}-y^\star\|^2 \right\} \\ =&G_\xi((x,y^\star);z^\star). \end{aligned} \tag{10}$$

Simlarly, we could show that

$$\max_{\hat{x}} \left\{ \mathcal{L}(x^\star,y^\star) - \mathcal{L}(\hat{x},y) - \frac{\xi}{2}\|\hat{x}-x^\star\|^2 \right\} = G_\xi((x^\star,y);z^\star). \tag{11}$$

So we complete the proof. □

**Theorem 1.** *For any* $\xi > 0$*, the smoothed duality gap of* (2) *satisfies quadratic growth on* $\mathcal{Z}$*. Namely, it holds for any* $z \in \mathcal{Z}$ *with* $\|z\| \le R$ *that*

$$G_\xi(z;z^\star) \ge \alpha_\xi \cdot \textit{dist}^2(z, \mathcal{Z}^\star),$$

*where* $\alpha_\xi = \gamma_\xi^2 \cdot \min\left\{ \frac{\xi}{4}, \frac{1}{4\xi}, \frac{1}{2L(1+\sqrt{m}R)} \right\}$.

**Lemma 3.** *For any $z \in \mathcal{Z}$ and $z^\star = (x^\star, y^\star) \in \mathcal{Z}^\star$, the smoothed duality gap at $(x, y^\star)$ equals to*

$$G_\xi((x,y^\star);z^\star) \geq \gamma_\xi^2 \cdot \min\left\{\frac{\xi}{4}, \frac{1}{4\xi}, \frac{1}{2L(1+\|y^\star\|_1)}\right\} \cdot dist(x, \mathcal{X}^\star)^2.$$

*Proof.* According to (10), we have

$$\begin{aligned}
G_\xi((x,y^\star);z^\star) &= \max_{\hat{y}\geq 0}\left\{\mathcal{L}(x,\hat{y}) - \mathcal{L}(x^\star,y^\star) - \frac{\xi}{2}\|\hat{y}-y^\star\|^2\right\} \\
&= \max_{\hat{y}\geq 0}\left\{f(x) + h(x) + G(x)^T\hat{y} - \frac{\xi}{2}\|\hat{y}-y^\star\|^2\right\} - \mathcal{L}(x^\star,y^\star) \\
&= \max_{\hat{y}\geq 0}\left\{-\frac{\xi}{2}\left\|\hat{y} - y^\star - \frac{1}{\xi}G(x)\right\|^2\right\} + \frac{\xi}{2}\left\|\frac{1}{\xi}G(x)\right\|^2 + f(x) + h(x) + G(x)^Ty^\star - \mathcal{L}(x^\star,y^\star) \\
&= \frac{\xi}{2}\left(\left\|\frac{1}{\xi}G(x)\right\|^2 - \left\|\left[y^\star + \frac{1}{\xi}G(x)\right]_-\right\|^2\right) + \mathcal{L}(x,y^\star) - \mathcal{L}(x^\star,y^\star)
\end{aligned} \tag{12}$$

where the last equality holds because the optimal $\hat{y} = \left[y^\star + \frac{1}{\xi}G(x)\right]_+$. On the one hand, the fact that $y^\star \geq 0$ implies that $\left[y^\star + \frac{1}{\xi}G(x)\right]_- \leq \left[\frac{1}{\xi}G(x)\right]_-$, so we have

$$\left\|\frac{1}{\xi}G(x)\right\|^2 - \left\|\left[y^\star + \frac{1}{\xi}G(x)\right]_-\right\|^2 \geq \left\|\frac{1}{\xi}G(x)\right\|^2 - \left\|\left[\frac{1}{\xi}G(x)\right]_-\right\|^2 = \left\|\left[\frac{1}{\xi}G(x)\right]_+\right\|^2.$$

On the other hand, it is easy to check that

$$\begin{aligned}
&\left\|\frac{1}{\xi}G(x)\right\|^2 - \left\|\left[y^\star + \frac{1}{\xi}G(x)\right]_-\right\|^2 \\
=&\left\|\frac{1}{\xi}G(x) + \left[y^\star + \frac{1}{\xi}G(x)\right]_-\right\|^2 - 2\left(\frac{1}{\xi}G(x) + \left[y^\star + \frac{1}{\xi}G(x)\right]_-\right)^T\left[y^\star + \frac{1}{\xi}G(x)\right]_- \\
=&\left\|\left[y^\star + \frac{1}{\xi}G(x)\right]_+ - y^\star\right\|^2 - 2\left(\left[y^\star + \frac{1}{\xi}G(x)\right]_+ - y^\star\right)^T\left[y^\star + \frac{1}{\xi}G(x)\right]_- \\
=&\left\|\left[y^\star + \frac{1}{\xi}G(x)\right]_+ - y^\star\right\|^2 + 2y^{\star T}\left[y^\star + \frac{1}{\xi}G(x)\right]_- \\
\geq&\left\|\left[y^\star + \frac{1}{\xi}G(x)\right]_+ - y^\star\right\|^2
\end{aligned} \tag{13}$$

where the second equality holds because of the fact

$$\frac{1}{\xi}G(x) + \left[y^\star + \frac{1}{\xi}G(x)\right]_- = \left(y^\star + \frac{1}{\xi}G(x)\right) + \left[y^\star + \frac{1}{\xi}G(x)\right]_- - y^\star = \left[y^\star + \frac{1}{\xi}G(x)\right]_+ - y^\star,$$

the third equality holds because the product of the positive part and negative part should be 0, and the last ineqality holds because $y^\star \geq 0$ and $\left[y^\star + \frac{1}{\xi}G(x)\right]_- \geq 0$. Next, note that the function $L(x, y^\star)$ is convex w.r.t. $x$ and $x^\star$ is one of the global optimal solutions, and the $L$-smoothness of both $f$ and $G$ implies that $\mathcal{L}(x, y^\star)$ is $L(1 + \|y^\star\|_1)$-smooth. So we have

$$\mathcal{L}(x,y^\star) - \mathcal{L}(x^\star,y^\star) \geq \frac{L}{2}\left\|\text{prox}_{\frac{1}{L}h}\left(x - \frac{1}{L}(\nabla f(x) + \nabla G(x)^Ty^\star)\right) - x\right\|^2$$

Therefore, we have

$$
\begin{aligned}
G_\xi((x,y^\star);z^\star) \geq & \frac{\xi}{4}\left\|\left[y^\star+\frac{1}{\xi}G(x)\right]_+ - y^\star\right\|^2 + \frac{1}{4\xi}\|[G(x)]_+\|^2 + \frac{1}{2L(1+\|y^\star\|_1)}\|\nabla f(x)+\nabla G(x)^T y^\star\|^2 \\
\geq & \min\left\{\frac{\xi}{4}, \frac{1}{4\xi}, \frac{1}{2L(1+\|y^\star\|_1)}\right\}\|F_\xi(x,y^\star)\|^2 \\
\geq & \gamma_\xi^2 \cdot \min\left\{\frac{\xi}{4}, \frac{1}{4\xi}, \frac{1}{2L(1+\|y^\star\|_1)}\right\} \cdot \mathrm{dist}(x,\mathcal{X}^\star)^2.
\end{aligned} \tag{14}
$$

where the last inequality holds by applying Proposition 1 to $z=(x,y^\star)$. ☐

**Lemma 4.** *For any $z\in\mathcal{Z}$ and $z^\star=(x^\star,y^\star)\in\mathcal{Z}^\star$, the smoothed duality gap at $(x^\star,y)$ is equal to*

$$G_\xi((x^\star,y);z^\star) \geq \gamma_\xi^2\cdot\min\left\{\frac{1}{\xi}, \frac{1}{2L(1+\|y\|_1)}\right\}\cdot dist(y,\mathcal{Y}^\star)^2.$$

*Proof.* According to (11), we have

$$
\begin{aligned}
G_\xi((x^\star,y);z^\star) &= \max_{\hat{x}}\left\{\mathcal{L}(x^\star,y^\star)-\mathcal{L}(\hat{x},y)-\frac{\xi}{2}\|\hat{x}-x^\star\|^2\right\} \\
&= f(x^\star)+h(x^\star)-\min_{\hat{x}}\left\{f(\hat{x})+h(\hat{x})+G(\hat{x})^T y+\frac{\xi}{2}\|\hat{x}-x^\star\|^2\right\}
\end{aligned} \tag{15}
$$

The $L$-smoothness of both $f$ and $G$ implies that

$$f(\hat{x})+h(\hat{x})+G(\hat{x})^T y \leq f(x^\star)+G(x^\star)^T y+\left(\nabla f(x^\star)+\nabla G(x^\star)^T y\right)^T(\hat{x}-x^\star)+\frac{L(1+\|y\|_1)}{2}\|\hat{x}-x^\star\|^2+h(\hat{x})$$

Assume $y$ is bounded and $M=L(1+\|y\|_1)+\xi$, we have

$$
\begin{aligned}
G_\xi((x^\star,y);z^\star) &\geq f(x^\star)+h(x^\star)-\min_{\hat{x}}\left\{f(x^\star)+G(x^\star)^T y+\left(\nabla f(x^\star)+\nabla G(x^\star)^T y\right)^T(\hat{x}-x^\star)+\frac{M}{2}\|\hat{x}-x^\star\|^2+h(\hat{x})\right\} \\
&= \frac{M}{2}\left\|\mathrm{prox}_{\frac{1}{M}h}\left(x^\star-\frac{1}{M}(\nabla f(x^\star)+\nabla G(x^\star)^T y)\right)-x\right\|^2 - G(x^\star)^T y.
\end{aligned} \tag{16}
$$

We complete the proof by

$$-G(x^\star)^T y \geq \frac{1}{\xi}\left\|y-\left[y+\frac{1}{\xi}G(x^\star)\right]_+\right\|^2$$

which can be checked by entrywise discuss whether $y_i+\frac{1}{\xi}G_i(x^\star)\geq 0$ or not. ☐

## 4 Application: Fisher Equilibrium Problem

In Fisher's market model, the players are divided into two sets: producers and consumers. Consumer $i, i\in C$ has given money endowment $w_i$ to spend and buys goods to maximize their individual utility functions; producer $j, j\in P$, sells its good for money. The price equilibrium is an assignment of prices to goods so that when every consumer buys a maximal bundle of goods then the market clears, meaning that all the money is spent and all the goods are sold. A convex optimization formulation is considered in the literature:

$$
\begin{aligned}
\max_x \quad & \sum_{i\in C} w_i \log\left(\sum_{j\in P} u_{ij}x_{ij}\right) \\
\text{s.t.} \quad & \sum_{i\in C} x_{ij} = p_j,\ \forall j\in P \\
& x_{ij}\geq 0,\ \forall i,j
\end{aligned} \tag{17}
$$

where $p_j > 0$ denotes the unit of producer $j$'s good, $u_{ij} \geq 0$ is the given utility coefficient of player $i$ for producer $j$'s good, and the varible $x_{ij}$ represents the amount of good bought from producer $j$ by comsumer $i$.

Existing results in literature:

- Sublinear convergence rate $(1/k)$ is guaranteed both exact PDHG [1] and linear approximation PDHG [2].
- According to [3], exact PDHG has linear convergence rate if the smoothed quadratic growth property holds.
-

## References


[1] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, 40:120–145, 2011.

[2] Yunmei Chen, Guanghui Lan, and Yuyuan Ouyang. Optimal primal-dual methods for a class of saddle point problems. *SIAM Journal on Optimization*, 24(4):1779–1814, 2014.

[3] Olivier Fercoq. Quadratic error bound of the smoothed gap and the restarted averaged primal-dual hybrid gradient. *arXiv preprint arXiv:2206.03041*, 2022.

[4] R. TYRRELL ROCKAFELLAR. Princeton University Press, 1970.