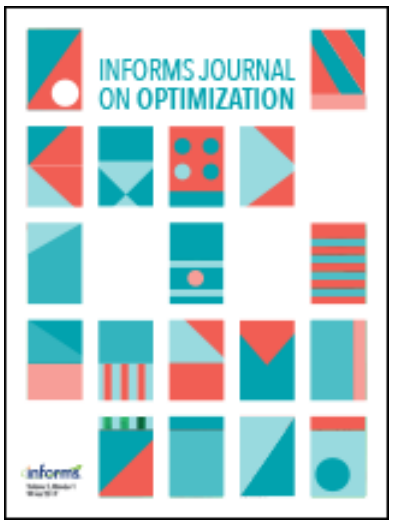


INFORMS Journal on Optimization

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Comparison-Based Algorithms for One-Dimensional Stochastic Convex Optimization

Xi Chen, Qihang Lin, Zizhuo Wang

To cite this article:
Xi Chen, Qihang Lin, Zizhuo Wang (2020) Comparison-Based Algorithms for One-Dimensional Stochastic Convex Optimization. INFORMS Journal on Optimization 2(1):34-56. https://doi.org/10.1287/ijoo.2019.0022

http://pubsonline.informs.org/journal/ijoo

INFORMS JOURNAL ON OPTIMIZATION
Vol. 2, No. 1, Winter 2020, pp. 34–56
ISSN 1091-9856 (print), ISSN 1526-5528 (online)

Comparison-Based Algorithms for One-Dimensional Stochastic Convex Optimization

Xi Chen,[a] Qihang Lin,[b] Zizhuo Wang[c,d]

[a] Stern School of Business, New York University, New York, New York 10012; [b] Tippie College of Business, University of Iowa, Iowa City, Iowa 52245; [c] Department of Industrial and Systems Engineering, University of Minnesota, Minneapolis, Minnesota 55455; [d] Institute for Data and Decision Analytics, The Chinese University of Hong Kong, 518172 Shenzhen, China

Contact: xchen3@stern.nyu.edu (XC); qihang-lin@uiowa.edu, http://orcid.org/0000-0003-2943-3267 (QL), zwang@umn.edu, http://orcid.org/0000-0003-0828-7280 (ZW)

Received: March 2, 2018
Revised: August 17, 2018; January 23, 2019
Accepted: March 31, 2019
Published Online in Articles in Advance: January 27, 2020

https://doi.org/10.1287/ijoo.2019.0022

Abstract. Stochastic optimization finds a wide range of applications in operations research and management science. However, existing stochastic optimization techniques usually require the information of random samples (e.g., demands in the newsvendor problem) or the objective values at the sampled points (e.g., the lost sales cost), which might not be available in practice. In this paper, we consider a new setup for stochastic optimization, in which the decision maker has access to only comparative information between a random sample and two chosen decision points in each iteration. We propose a comparison-based algorithm (CBA) to solve such problems in one dimension with convex objective functions. Particularly, the CBA properly chooses the two points in each iteration and constructs an unbiased gradient estimate for the original problem. We show that the CBA achieves the same convergence rate as the optimal stochastic gradient methods (with the samples observed). We also consider extensions of our approach to multidimensional quadratic problems as well as problems with nonconvex objective functions. Numerical experiments show that the CBA performs well in test problems.

Funding: X. Chen was supported by the National Science Foundation [IIS-1845444], the Alibaba Innovation Research Award, and the Bloomberg Data Science Research Grant.
Supplemental Material: The e-companion is available at https://doi.org/10.1287/ijoo.2019.0022.

Keywords: stochastic optimization • comparison-based algorithm • stochastic gradient

1. Introduction

In this paper, we consider the following stochastic optimization problem:

$$\min_{\ell \le x \le u} H(x) = \mathbb{E}_{\xi}[h(x,\xi)], \tag{1}$$

where $-\infty \le \ell \le u \le +\infty$ and ξ is a random variable.[1] This problem has many applications and is fundamental for stochastic optimization. For example,

1. If $h(x,\xi) = (x-\xi)^2$ with $\ell = -\infty$ and $u = +\infty$, then the problem is to find the expectation of ξ.
2. If $h(x,\xi) = \mathfrak{h}\cdot(x-\xi)^+ + \mathfrak{b}\cdot(\xi-x)^+$, then the problem is the classical newsvendor problem with unit holding cost $\mathfrak{h}$ and unit back-order cost $\mathfrak{b}$. It can also be viewed as the problem of finding the $\mathfrak{b}/(\mathfrak{h}+\mathfrak{b})$th quantile of ξ when $\ell = -\infty$ and $u = +\infty$. Furthermore, one can consider a more general version of this problem in which

$$h(x,\xi) = \begin{cases} \mathfrak{h}_+(x,\xi) & \text{if } x \ge \xi \\ \mathfrak{h}_-(x,\xi) & \text{if } x < \xi. \end{cases}$$

This problem can be viewed as a newsvendor problem with general holding and back-order costs (see, e.g., Halman et al. 2012 and references therein for discussions of this problem in which $\mathfrak{h}_+(x,\xi)$ is a general holding cost function and $\mathfrak{h}_-(x,\xi)$ is a general back-order cost function). It can also be viewed as a single-period appointment scheduling problem with general waiting and overtime costs (see, e.g., Gupta and Denton 2008) or a staffing problem with general underage and overage costs (see, e.g., Kolker 2017).

3. If $h(x,\xi) = -x\cdot 1(\xi \ge x)$, then the problem can be viewed as an optimal pricing problem in which x is the price set by the seller, ξ is the valuation of each customer, and $h(x,\xi)$ is the negative of the revenue obtained from the customer (a customer purchases at price x if and only if the customer's valuation ξ is greater than or equal to x).

In many practical situations, the distribution of ξ, whose cumulative distribution function (c.d.f.) is denoted by $F(\cdot)$, is unknown a priori. Existing stochastic optimization techniques for (1) usually require sampling from the distribution of ξ and use random samples to update the decision x toward optimality. In the existing methods, it is assumed that either the random samples of ξ are fully observed or the objective value $h(x, \xi)$ under a decision x and random samples ξ can be observed. However, such information may not always be available in practice (see Examples 1–3).

In this paper, we investigate whether having full sample information is always critical for solving one-dimensional stochastic convex optimization problems. We realize that, in some cases in which full samples are not observed, a comparative relation between a chosen decision variable and a random sample may still be accessible. This motivates us to study stochastic optimization with only the presence of comparative information. Specifically, given a decision x, a sample ξ is drawn from the underlying distribution, and we assume that we only have information about whether ξ is greater than or less than (or equals to) x. In addition, after knowing the comparative relation between x and ξ, we further assume that we can choose another point z and obtain information about whether ξ is greater than or less than (or equals to) z. Such a z is not a decision variable but a randomly sampled point. We show that, in fact, having the comparative information in this way can sometimes be sufficient for solving (1). In the following, we list several scenarios in which such situations may arise:

Example 1. Suppose x represents a certain feature of a product (e.g., size or taste, etc.), ξ is the preference of each customer about that feature, and the firm selling this product would like to find out the average preference of the customers (or, equivalently, to find the optimal offering to minimize the expected customer dissatisfaction measured by $h(x, \xi) = (x - \xi)^2$). Such a firm faces a stochastic optimization problem described in the first example. In many cases, it is hard for a customer to give an exact value for the customer's preference (i.e., the exact value of the customer's ξ). However, it is quite plausible that the customer can report a comparative relation between the customer's preferred value of the feature and the actual value of the feature of the product presented to the customer (e.g., whether the product should be larger or whether the taste should be saltier). Furthermore, it is possible to ask one customer to compare the customer's preferred value with two different values of the same feature of the product, for example, by giving the customer two different samples. Moreover, the second sample may be given in a customer satisfaction survey, and the customer will not count the second sample toward its (dis) satisfaction value. Therefore, such a scenario fits the setting described.

Example 2. In a newsvendor problem, it is sometimes hard to observe the exact demand in each period because of demand censorship. In such situations, one does not have direct access to the sample point (the demand), nor does one have access to the cost in the corresponding period (the lost sales cost).[2] However, the seller usually has comparative information between the realized demand and the chosen inventory level (e.g., by observing if there is a stockout or a leftover). Moreover, by allowing the seller to make a one-time additional order in each time period (this ability is sometimes called the quick response ability for the seller; see, e.g., Cachon and Swinney 2011), it is possible that one can obtain such information at two points. In such cases, the firm faces a newsvendor problem as described in the second example, and thus, it corresponds to the setting in our problem.

Example 3. In a revenue management problem, by offering a price to each customer, the seller can observe whether the customer purchased the product, and the seller faces a stochastic optimization problem described in the third example. In practice, it is hard to ask the customer to report a true valuation of the product. However, it is possible to ask the customer in a market survey whether the customer will purchase the product at a different price. Such an example can also be extended to a divisible product case in which a customer can buy a continuous amount of a product with a maximum of one. In this case, the h function can be redefined as $h(x, \xi) = -x \min\{1, g(x, \xi)\}$, where x is the offered price, $g(x, \xi)$ is the unconstrained purchase amount of the customer, and ξ is the maximum price this customer is willing to pay for the full amount of this product (i.e., $g(x, \xi)$ is decreasing in x with $g(x, \xi) > 1$ when $x < \xi$ and $g(x, \xi) < 1$ when $x > \xi$). Such a purchase behavior can be explained by a quadratic utility function of the customer, which is often used in the literature (see, e.g., Candogan et al. 2012). For the seller, by observing whether the customer buys the full amount of the product, the seller can infer whether an offered price is greater than or less than the ξ value of this customer.[3]

In this paper, we propose an efficient algorithm to solve the described stochastic optimization problem. More precisely, we propose a stochastic approximation algorithm that only utilizes comparative information between each sample point and two chosen decision points in each iteration. We show that, by properly choosing the two points (one point has to be chosen randomly according to a specifically designed

distribution), we can obtain unbiased gradient estimates for the original problem.[4] The unbiased gradient estimates, in turn, give rise to efficient algorithms based on a standard stochastic gradient method (we review the related literature shortly). Under some mild conditions, we show that, if the original problem is convex, then our algorithm achieves a convergence rate of $O(1/\sqrt{T})$ for the objective value (where T is the number of iterations); if the original problem is strongly convex, then the convergence rate can be improved to $O(1/T)$. Moreover, the information at two points is necessary in this setting as we show that only knowing comparative information between the sample and one point in each iteration is insufficient for any algorithm to converge to the optimal solution (see Example 4). We also perform several numerical experiments using our algorithm. The experimental results show that our algorithms are indeed efficient with convergence speed in the same order compared with the case in which one has direct observations of the samples. We also extend our algorithm to a multidimensional setting with quadratic objective function, a setting with nonconvex objective function, and a setting in which multiple comparisons can be conducted in each iteration.

1.1. Literature Review

Broadly speaking, our work falls into the area of stochastic optimization, a subject on which there is vast literature. For a comprehensive review of this literature, we refer the readers to Shapiro et al. (2014). In particular, in this literature, it is usually assumed that one has access to random samples (or, alternatively, the objective values at the sampled points). Two main types of algorithms have been proposed, namely the sample average approximation (SAA) method (see, e.g., Shapiro et al. 2014) and the stochastic approximation (SA) methods (see, e.g., Robbins and Monro 1951 and Kiefer and Wolfowitz 1952). A typical SAA method collects a number of samples from the underlying distribution and uses the sample averaged objective function to approximate the expected value function. This method has been widely used in many operations management problems (see, e.g., Levi et al. 2015 and Ban and Rudin 2018). In contrast, the SA approach (e.g., the stochastic gradient descent) is usually an iterative algorithm. In each iteration, a new sample (or a small batch of new samples) is drawn, and a new iterate is computed using the new sample(s). Our work belongs to the category of SA. In the following, we focus our literature review on the stochastic approximation methods.

If the objective function is convex, then various stochastic gradient methods can guarantee, under slightly different assumptions, that the objective value of the iterates converges to the optimal value in a rate of $O(1/\sqrt{T})$ after T iterations (see, e.g., Nemirovski et al. 2009, Duchi and Singer 2009, Hu et al. 2009, Xiao 2010, Chen et al. 2012, Lan 2012, Lan and Ghadimi 2012, Rakhlin et al. 2012, Lan and Ghadimi 2013, Shamir and Zhang 2013, and Hazan and Kale 2014, Lin et al. 2014). Furthermore, this convergence rate is known to be optimal (Nemirovski and Yudin 1983). When the objective function is strongly convex, some stochastic gradient methods can obtain an improved convergence rate of $O(\log T/T)$ (Duchi and Singer 2009, Xiao 2010). More recently, several papers have further improved the convergence rate to $O(1/T)$. Among those papers, there are three different methods used: (a) the accelerated stochastic gradient method with auxiliary iterates besides the main iterate (Hu et al. 2009, Chen et al. 2012, Lan 2012, Lan and Ghadimi 2012, Lan and Ghadimi 2013, Lin et al. 2014), (b) averaging the historical solutions (Rakhlin et al. 2012, Shamir and Zhang 2013), and (c) the multistage stochastic gradient method that periodically restarts (Hazan and Kale 2014). Again, the convergence rate of $O(1/T)$ has been shown to be optimal by Nemirovski and Yudin (1983) for strongly convex problems.

Apart from the setting of minimizing a single objective function, stochastic gradient methods can also be applied to online learning problems (for a comprehensive review, see Shalev-Shwartz 2012) in which a sequence of functions is presented to a decision maker who needs to provide a solution sequentially to each function with the goal of minimizing the total regret. It is known that the stochastic gradient methods can obtain a regret of $O(\sqrt{T})$ after T decisions if the functions presented are convex. Moreover, this regret has been shown to be optimal (Cesa-Bianchi and Lugosi 2006). When the functions are strongly convex, Zinkevich (2003), Duchi and Singer (2009), Xiao (2010), and Duchi et al. (2011) show that the regret can be further improved to $O(\log T)$.

To distinguish our work from these examples, we note that all of these works have assumed that either the sample is directly accessible (one can observe the value of each sample) or the objective value corresponding to the decision variable and the current sample is accessible. In either case, it is easy to obtain an estimate of the gradient of the objective function. In contrast, in our case, we do not have access to the sample or the objective value. Instead, we only have comparative information between each sample and two chosen decision points. Indeed, as we discuss in the next section, one of the main challenges in our problem is to use this very limited information to construct an unbiased gradient for the original problem and then further use it to find the

optimal solution. Our contribution is to show that the same order of convergence rate can still be achieved under this setting with less information.

2. Main Results

In this paper, we make the following assumptions:

Assumption 1. *The random variable ξ follows a continuous distribution.*

Assumption 2. *For each ξ, $h(x,\xi)$ is continuously differentiable with respect to x on $[\ell,\xi)$ and $(\xi,u]$ with the derivative denoted by $h'_x(x,\xi)$. Furthermore, for any $x\in[\ell,u]$, $h'_-(x):=\lim_{z\to x-}h'_x(x,z)$ and $h'_+(x):=\lim_{z\to x+}h'_x(x,z)$ exist and are finite.*

Assumption 3. *For any $x\in[\ell,u]$, $x\neq\xi$, $h''_{x,\xi}(x,\xi)=\frac{\partial^2 h(x,\xi)}{\partial\xi\partial x}$ exists.*

Assumption 4. *The function $H(x)$ in (1) is differentiable and μ-convex on $[\ell,u]$ for some $\mu\geq 0$, namely*

$$H(x_2)\geq H(x_1)+H'(x_1)(x_2-x_1)+\frac{\mu}{2}(x_2-x_1)^2,\quad \forall x_1,x_2\in[\ell,u]. \tag{2}$$

Moreover, $\mathbb{E}_\xi(h'_x(x,\xi))=H'(x)$ for all $x\in[\ell,u]$.

Assumption 5. Either of the following statements is true:

a. *There exists a constant K_1 such that $\mathbb{E}_\xi(h'_x(x,\xi))^2\leq K_1^2$ for any $x\in[\ell,u]$.*
b. *$H'(x)$ is L-Lipschitz continuous on $[\ell,u]$. Furthermore, there exists a constant K_2 such that*

$$\mathbb{E}_\xi(h'_x(x,\xi)-H'(x))^2\leq K_2^2,\quad \forall x\in[\ell,u].$$

Now, we make several comments on these assumptions. The first assumption that ξ is continuously distributed is mainly for ease of discussion. In fact, all of our results continue to hold as long as, with probability one, for all iterates x in our algorithm, $\mathbb{P}(\xi=x)=0$. We revisit this assumption in Section 6. Assumptions 2–4 are some regularity assumptions on the functions h and H. Particularly, the last point of Assumption 4 is satisfied under many cases, for example, when $h'_x(x,\xi)$ is continuous in ξ and ξ is supported on a finite set (Widder 1990) or when $h(x,\xi)$ is convex in x for each ξ (by the monotone convergence theorem). When (2) holds with $\mu>0$, we call H a *μ-strongly convex* function. The last assumption states that the partial derivative $h'_x(x,\xi)$ has uniformly bounded second-order moment or variance. This is used to guarantee that the step in each iteration in our algorithm has bounded variance, which is a common assumption in the stochastic approximation literature (see, e.g., Duchi and Singer 2009, Nemirovski et al. 2009, and Lan and Ghadimi 2013).

In addition, Assumption 1 is not hard to satisfy in our examples mentioned earlier. Specifically, Example 1 satisfies Assumption 1 when ξ is continuously distributed and has finite variance. Example 2 satisfies Assumption 1 when ξ is continuously distributed and the cost functions are linear. When the cost functions are nonlinear ($\mathfrak{h}_+$ and $\mathfrak{h}_-$, respectively), it satisfies Assumption 1 if both $\mathfrak{h}_+$ and $\mathfrak{h}_-$ are second-order continuously differentiable on their respective domains, have bounded first-order derivatives (for example, when x and ξ are restricted to finite intervals), and h is convex in x. Example 3 satisfies Assumption 1 under the divisible case when ξ is a continuous random variable and the expected revenue function $x\mathbb{E}_\xi\min\{g(x,\xi),1\}$ is concave on $[\ell,u]$, which holds, for example, when $g(x,\xi)$ is a piecewise linear function and when the range $[\ell,u]$ is small.[5]

In the following, we propose a comparison-based algorithm (CBA) to solve (1). Let Ξ denote the support of ξ with $-\infty\leq\underline{s}:=\inf\{\Xi\}\leq\bar{s}:=\sup\{\Xi\}\leq+\infty$. The algorithm requires specification of two functions, $f_-(x,z)$ and $f_+(x,z)$, which need to satisfy the following conditions.

Condition 1. *$f_-(x,z)=0$ for all $z\geq x$ and $f_-(x,z)>0$ for all $\underline{s}\leq z<x$. In addition, for all x, we have $\int_{-\infty}^{x-}f_-(x,z)dz=1$.*

Condition 2. *$f_+(x,z)=0$ for all $z\leq x$ and $f_+(x,z)>0$ for all $\bar{s}\geq z>x$. In addition, for all x, we have $\int_{x+}^{\infty}f_+(x,z)dz=1$.*

Condition 3. *There exists a constant K_3 such that $\int_{\underline{s}}^{x-}\frac{F(z)(h''_{x,z}(x,z))^2}{f_-(x,z)}dz\leq K_3$ and $\int_{x+}^{\bar{s}}\frac{(1-F(z))(h''_{x,z}(x,z))^2}{f_+(x,z)}dz\leq K_3$ for all $x\in[\ell,u]$, where $F(\cdot)$ is the c.d.f. of ξ.*

Note that, for any given $x\in[\ell,u]$, $f_-(x,z)$ and $f_+(x,z)$ essentially define two density functions of z on $(-\infty,x]$ and $[x,+\infty)$. (We discuss how to choose $f_-(\cdot,\cdot)$ and $f_+(\cdot,\cdot)$ in Section 3.) Also, we note that $\underline{s}$ and $\bar{s}$ need not be known in advance. If one is unsure about $\underline{s}$ ($\bar{s}$, respectively), then one can choose a sufficiently small (large) value, or just choose $\underline{s}=-\infty$ ($\bar{s}=+\infty$). Condition 3 is a technical condition and may not be straightforward to

verify at the first glance. However, in Section 3, we show that, under mild conditions (e.g., ξ has a light tail and $h''_{x,z}(x,z)$ is uniformly bounded), it is not hard to choose the functions $f_-(x,z)$ and $f_+(x,z)$ such that Condition 3 is satisfied (we leave the detailed discussions in Section 3). Next, in Algorithm 1, we describe the detailed procedure of the CBA.

Algorithm 1 (CBA)

1. **Initialization.** Set $t = 1$, $x_1 \in [\ell, u]$. Define η_t for all $t \geq 1$. Set the maximum number of iterations T. Choose functions $f_-(x,z)$ and $f_+(x,z)$ that satisfy Conditions 1–3.
2. **Main iteration.** Sample ξ_t from the distribution of ξ. If $\xi_t = x_t$, then resample ξ_t until it does not equal x_t. (This step always terminates in a finite number of steps as long as ξ is not deterministic.)

   a. If $\xi_t < x_t$, then generate z_t from a distribution on $(-\infty, x_t]$ with probability density function (p.d.f.) $f_-(x_t, z_t)$. Set

$$g(x_t, \xi_t, z_t) = \begin{cases} h'_-(x_t), & \text{if } z_t < \xi_t, \\ h'_-(x_t) - \frac{h''_{x,z}(x_t,z_t)}{f_-(x_t,z_t)}, & \text{if } z_t \geq \xi_t. \end{cases} \tag{3}$$

   b. If $\xi_t > x_t$, then generate z_t from a distribution on $[x_t, +\infty)$ with p.d.f. $f_+(x_t, z_t)$. Set

$$g(x_t, \xi_t, z_t) = \begin{cases} h'_+(x_t), & \text{if } z_t > \xi_t, \\ h'_+(x_t) + \frac{h''_{x,z}(x_t,z_t)}{f_+(x_t,z_t)}, & \text{if } z_t \leq \xi_t. \end{cases} \tag{4}$$

   Let

$$x_{t+1} = \mathrm{Proj}_{[\ell,u]}\big(x_t - \eta_t g(x_t, \xi_t, z_t)\big) = \max\big(\ell, \min\big(u, x_t - \eta_t g(x_t, \xi_t, z_t)\big)\big). \tag{5}$$

3. **Termination.** Stop when $t \geq T$. Otherwise, let $t \leftarrow t + 1$ and go back to step 2.
4. **Output.** $\mathbf{CBA}(x_1, T, \{\eta_t\}_{t=1}^T) = \bar{x}_T = \frac{1}{T}\sum_{t=1}^T x_t$.

In iteration t of the CBA, the solution x_t is updated based on two random samples. First, a sample ξ_t is drawn from the distribution of ξ. In contrast to existing stochastic gradient methods, the CBA does not require exactly observing ξ_t but only needs to know whether $\xi_t < x_t$ or $\xi_t > x_t$. Based on the result of the comparison between ξ_t and x_t, a second sample z_t is drawn from the density function $f_+(x_t, z)$ or $f_-(x_t, z)$. An unbiased stochastic gradient, $g(x_t, \xi_t, z_t)$, of $H(x_t)$ is then constructed and used to update x_t with the standard gradient descent step.

We note that the output of Algorithm 1 is the average of the historical solutions $\bar{x}_T = \frac{1}{T}\sum_{t=1}^T x_t$. This is because the convergence of objective value is established based on $\bar{x}_T$. However, Algorithm 1 can be applied to the online learning setting in which one can use the solution x_t as the decision in each stage t and obtain the desired expected total regret (see Propositions 2–4). We have the following proposition about the stochastic gradient $g(x_t, \xi_t, z_t)$ in the CBA.

Proposition 1. *Suppose $f_-(x,z)$ and $f_+(x,z)$ satisfy Conditions 1–3 and Assumption 1 holds. Then*

1. $\mathbb{E}_z g(x, \xi, z) = h'_x(x, \xi)$ *for all* $x \in [\ell, u]$, $x \neq \xi$.
2. $\mathbb{E}_{z,\xi} g(x, \xi, z) = H'(x)$ *for all* $x \in [\ell, u]$.
3. *If Assumption 5(a) holds, then* $\mathbb{E}_{z,\xi}(g(x,\xi,z))^2 \leq G^2 := K_1^2 + 2K_3$. *If Assumption 5(b) holds, then* $\mathbb{E}_{z,\xi}(g(x,\xi,z) - H'(x))^2 \leq \sigma^2 := K_2^2 + 2K_3$.

Proof of Proposition 1. First, we consider the case in which $\xi < x$. We have

$$\mathbb{E}_z g(x, \xi, z) = h'_-(x) - \int_\xi^{x-} h''_{x,z}(x,z)dz = h'_x(x, \xi).$$

Similarly, when $\xi > x$,

$$\mathbb{E}_z g(x, \xi, z) = h'_+(x) + \int_{x+}^{\xi} h''_{x,z}(x,z)dz = h'_x(x, \xi).$$

Thus, the first conclusion of the proposition is proved. The second conclusion of the proposition follows from Assumption 1 (which ensures $\xi = x$ is a zero-measure event) and Assumption 4.

Next, we show the first part of the third conclusion when Assumption 5(a) is true. If $\xi < x$, then we have

$$\begin{aligned}\mathbb{E}_z(g(x,\xi,z))^2 &= \int_{-\infty}^{x-}(h'_-(x))^2 f_-(x,z)dz + \int_{\xi}^{x-}\left(-2h'_-(x)\frac{h''_{x,z}(x,z)}{f_-(x,z)} + \left(\frac{h''_{x,z}(x,z)}{f_-(x,z)}\right)^2\right)f_-(x,z)dz\\ &= (h'_-(x))^2 - 2h'_-(x)(h'_-(x) - h'_x(x,\xi)) + \int_{\xi}^{x-}\frac{(h''_{x,z}(x,z))^2}{f_-(x,z)}dz\\ &\le (h'_x(x,\xi))^2 + \int_{\xi}^{x-}\frac{(h''_{x,z}(x,z))^2}{f_-(x,z)}dz, \end{aligned} \tag{6}$$

where the last inequality is because $a^2 + b^2 \ge 2ab$ for any a, b. By similar arguments, if $\xi > x$, then

$$\mathbb{E}_z(g(x,\xi,z))^2 \le (h'_x(x,\xi))^2 + \int_{x+}^{\xi}\frac{(h''_{x,z}(x,z))^2}{f_+(x,z)}dz.$$

These two inequalities and Assumption 5(a) further imply

$$\begin{aligned}\mathbb{E}_{z,\xi}(g(x,\xi,z))^2 &\le K_1^2 + \int_{\underline{s}}^{x-}\left(\int_{\xi}^{x-}\frac{(h''_{x,z}(x,z))^2}{f_-(x,z)}dz\right)dF(\xi) + \int_{x+}^{\bar{s}}\left(\int_{x+}^{\xi}\frac{(h''_{x,z}(x,z))^2}{f_+(x,z)}dz\right)dF(\xi)\\ &= K_1^2 + \int_{\underline{s}}^{x-}\frac{F(z)(h''_{x,z}(x,z))^2}{f_-(x,z)}dz + \int_{x+}^{\bar{s}}\frac{(1-F(z))(h''_{x,z}(x,z))^2}{f_+(x,z)}dz\\ &\le K_1^2 + 2K_3, \end{aligned} \tag{7}$$

where the interchanging of integrals in the equality is justified by Tonelli's theorem and the last inequality is due to Condition 3.

Next, we show the second part of the third conclusion when Assumption 5(b) is true. If $\xi < x$, then, following the similar analysis as in (6), we have

$$\mathbb{E}_z(g(x,\xi,z) - h'_x(x,\xi))^2 = \mathbb{E}_z(g(x,\xi,z))^2 - (h'_x(x,\xi))^2 \le \int_{\xi}^{x-}\frac{(h''_{x,z}(x,z))^2}{f_-(x,z)}dz.$$

Similarly, if $\xi > x$, then

$$\mathbb{E}_z(g(x,\xi,z) - h'_x(x,\xi))^2 \le \int_{x+}^{\xi}\frac{(h''_{x,z}(x,z))^2}{f_+(x,z)}dz.$$

By using the same argument as in (7), we have

$$\mathbb{E}_{z,\xi}(g(x,\xi,z) - h'_x(x,\xi))^2 \le 2K_3.$$

Finally, we note that,

$$\mathbb{E}_{z,\xi}(g(x,\xi,z) - H'(x))^2 = \mathbb{E}_{z,\xi}(g(x,\xi,z) - h'_x(x,\xi))^2 + \mathbb{E}_{\xi}(h'_x(x,\xi) - H'(x))^2.$$

Therefore, when Assumption 5(b) holds, we have $\mathbb{E}_{z,\xi}(g(x,\xi,z) - H'(x))^2 \le K_2^2 + 2K_3$. Thus, the proposition holds. □

Proposition 1 shows that, in the CBA, the gradient estimate $g(x,\xi,z)$ is an unbiased estimate of the true gradient at x and can be utilized as a stochastic gradient of $H(x)$. Note that such an unbiased gradient is generated without accessing the sample ξ itself nor the value of the objective function $h(x,\xi)$ at the sampled point. The only information used in generating the unbiased gradient is comparative information between the sample and two points.

Using the unbiased stochastic gradient, we can characterize the convergence results of the CBA in the following propositions.

Proposition 2. *Suppose $\mu = 0$. Let G^2 and σ^2 be defined as in Proposition 1 and x^* be any optimal solution to (1).*

- *If Assumption 5(a) holds, then, by choosing $\eta_t = \frac{1}{\sqrt{T}}$, the CBA ensures that*

$$\mathbb{E}(H(\bar{x}_T) - H(x^*)) \le \frac{(x_1 - x^*)^2}{2\sqrt{T}} + \frac{G^2}{2\sqrt{T}} \qquad \text{and} \qquad \sum_{t=1}^{T}\mathbb{E}(H(x_t) - H(x^*)) \le \frac{\sqrt{T}}{2}(x_1 - x^*)^2 + \frac{\sqrt{T}G^2}{2}.$$

If, in addition, u and ℓ are finite, then by choosing $\eta_t = \frac{1}{\sqrt{t}}$, the CBA ensures that

$$\mathbb{E}(H(\bar{x}_T) - H(x^*)) \le \frac{(u-\ell)^2}{2\sqrt{T}} + \frac{G^2}{\sqrt{T}} \qquad \text{and} \qquad \sum_{t=1}^{T} \mathbb{E}(H(x_t) - H(x^*)) \le \frac{\sqrt{T}}{2}(u-\ell)^2 + \sqrt{T}G^2.$$

- *If Assumption 5(b) holds, then, by choosing $\eta_t = \frac{1}{L+\sqrt{T}}$, the CBA ensures that*

$$\mathbb{E}(H(\bar{x}_T) - H(x^*)) \le \frac{L+\sqrt{T}}{2T}(x_1 - x^*)^2 + \frac{H(x_1) - H(x^*)}{T} + \frac{\sigma^2}{2\sqrt{T}}, \text{and}$$
$$\sum_{t=1}^{T} \mathbb{E}(H(x_t) - H(x^*)) \le \frac{L+\sqrt{T}}{2}(x_1 - x^*)^2 + H(x_1) - H(x^*) + \frac{\sqrt{T}\sigma^2}{2}.$$

If, in addition, u and ℓ are finite, then, by choosing $\eta_t = \frac{1}{L+\sqrt{t}}$, the CBA ensures that

$$\mathbb{E}(H(\bar{x}_T) - H(x^*)) \le \frac{(u-\ell)^2}{2\sqrt{T}} + \frac{H(x_1) - H(x^*)}{T} + \frac{L(x_1 - x^*)^2}{2T} + \frac{\sigma^2}{\sqrt{T}}, \text{and}$$
$$\sum_{t=1}^{T} \mathbb{E}(H(x_t) - H(x^*)) \le \frac{\sqrt{T}}{2}(u-\ell)^2 + H(x_1) - H(x^*) + \frac{L}{2}(x_1 - x^*)^2 + \sqrt{T}\sigma^2.$$

Proposition 2 gives the performance of the CBA when $\mu = 0$. When $\mu = 0$ and u and/or ℓ are infinite, the step size η_t is chosen to be a constant depending on the total number of iterations T in order to achieve the optimal convergence rate (i.e., $\eta_t = \frac{1}{\sqrt{T}}$ when Assumption 5(a) holds and $\eta_t = \frac{1}{L+\sqrt{T}}$ when Assumption 5(b) holds). Therefore, one needs to determine the total number of iterations T before running the optimization algorithm for the computation of the step size. When $\mu = 0$ and u and ℓ are finite, the step size η_t can be chosen as a decreasing sequence in t (i.e., $\eta_t = \frac{1}{\sqrt{t}}$ when Assumption 5(a) holds and $\eta_t = \frac{1}{L+\sqrt{t}}$ when Assumption 5(b) holds). In such a case, one does not need to prespecify the total number of iterations T. The convergence result when Assumption 5(a) holds has a proof similar to Nemirovski et al. (2009) and Duchi and Singer (2009) except using our comparison-based stochastic gradient. Also, the convergence result when Assumption 5(b) holds is largely built upon the results in Lan (2012). The detailed proof of the proposition is given in Appendix A in the e-companion.

Next, we have a further result when $\mu > 0$.

Proposition 3. *Suppose $\mu > 0$. Let G^2 and σ^2 be defined as in Proposition 1 and x^* be any optimal solution to (1).*

- *If Assumption 5(a) holds, then, by choosing $\eta_t = \frac{1}{\mu t}$, the CBA ensures that*

$$\mathbb{E}(H(\bar{x}_T) - H(x^*)) \le \frac{G^2}{2\mu}\frac{\log T + 1}{T} \qquad \text{and} \qquad \sum_{t=1}^{T} \mathbb{E}(H(x_t) - H(x^*)) \le \frac{G^2}{2\mu}(\log T + 1).$$

- *If Assumption 5(b) holds, then, by choosing $\eta_t = \frac{1}{\mu t + L}$, the CBA ensures that*

$$\mathbb{E}(H(\bar{x}_T) - H(x^*)) \le \frac{\sigma^2}{2\mu}\frac{\log T + 1}{T} + \frac{H(x_1) - H(x^*)}{T} + \frac{L(x_1 - x^*)^2}{2T}, \text{and}$$
$$\sum_{t=1}^{T} \mathbb{E}(H(x_t) - H(x^*)) \le \frac{\sigma^2}{2\mu}(\log T + 1) + H(x_1) - H(x^*) + \frac{L(x_1 - x^*)^2}{2}.$$

Proposition 3 gives the performance of the CBA when $\mu > 0$. The convergence rate of $O(\log T/T)$ when $\mu > 0$ is better than the convergence rate of $O(1/\sqrt{T})$ when $\mu = 0$. In such case, one does not need to know T in advance and can always choose the step size as a decreasing sequence in t. Again, the detailed proof of the proposition is given in Appendix A in the e-companion.

According to Proposition 3, when $\mu > 0$, the convergence rate of the CBA is $O(\log T/T)$. In the following, we improve the convergence rate when $\mu > 0$ to $O(1/T)$ using a restarting method first proposed by Hazan and Kale (2014). In Hazan and Kale (2014), the authors show that the restarting method works when Assumption 5(a) holds. In this paper, we extend the result by showing that the restarting method can also obtain the $O(1/T)$ rate if Assumption 5(b) holds. According to Nemirovski and Yudin (1983), no algorithm can achieve a convergence rate better than $O(1/T)$; thus, we have obtained the best possible convergence rates in those settings.

We now describe the restarting method in Algorithm 2, which we refer to as the multistage comparison-based algorithm (MCBA).

Algorithm 2 (MCBA)

1. Initialize the number of stages $K \geq 1$, the starting solution $\hat{x}^1$. Set $k = 1$.
2. Let T_k be the number of iterations in stage k and η_t^k be the step length in iteration t of the CBA in stage k for $t = 1, 2, \ldots, T_k$.
3. Let $\hat{x}^{k+1} = \mathbf{CBA}(\hat{x}^k, T_k, \{\eta_t^k\}_{t=1}^{T_k})$.
4. Stop when $k \geq K$. Otherwise, let $k \leftarrow k+1$ and go back to step 2.
5. Output $\hat{x}^{K+1}$.

We have the following proposition about the performance of Algorithm 2. The proof of the proposition is given in Appendix A in the e-companion.

Proposition 4. *Suppose $\mu > 0$. Let G^2 and σ^2 be defined as in Proposition 1, x^* be any optimal solution to (1), and $T = \sum_{k=1}^{K} T_k$ with T_k defined in the MCBA.*

- *If Assumption 5(a) holds, then, by choosing $\eta_t^k = \frac{1}{2^{k+1}\mu}$ and $T_k = 2^{k+3}$, the MCBA ensures that*

$$\mathbb{E}(H(\hat{x}_{K+1}) - H(x^*)) \leq \frac{16(H(\hat{x}_1) - H(x^*) + G^2/\mu)}{T}.$$

- *If Assumption 5(b) holds, then, by choosing $\eta_t^k = \frac{1}{2^{k+1}\mu + L}$ and $T_k = 2^{k+3} + 4$, the MCBA ensures that*

$$\mathbb{E}(H(\hat{x}_{K+1}) - H(x^*)) \leq \frac{32(H(\hat{x}_1) - H(x^*) + L(\hat{x}_1 - x^*)^2/2 + \sigma^2/\mu)}{T}.$$

Both Propositions 3 and 4 give the convergence rates of the CBA and MCBA for strongly convex problems (i.e., $\mu > 0$). The convergence rate of MCBA is $O(1/T)$, which improves the $O(\log T/T)$ convergence rate of the CBA by a factor of $\log T$. The intuition for this difference is that the output $\bar{x}_T$ of the CBA is the average of all historical solutions so that its quality is reduced by the earlier solutions (i.e., x_t with a small t), which are far from the optimal solution. On the contrary, the MCBA restarts the CBA periodically with a better initial solution (i.e., $\hat{x}_k$) for each restart. As a result, the output of the MCBA is the average of historical solutions only in the last (Kth) call of the CBA, which does not involve the earlier solutions and, thus, has a higher quality. This is the main reason for the MCBA to have a better solution after the same number of iterations or, equivalently, a better convergence rate than the CBA. However, the CBA is easier to implement as it does not require periodic restart as needed in the MCBA. Moreover, the theoretical convergence of the MCBA requires strong convexity in the problem while the CBA converges without strong convexity requirement (see Proposition 2).[6]

By Proposition 2–4, we have shown that, under some mild assumptions (Assumption 1), if one has access to comparative information between each sample ξ_t and *two* points, then one can still find the optimal solution to (1), and the convergence speed is in the same order as when one can observe the actual value of the sample (or the objective value at the sampled point). One natural question is whether the same convergence result can be achieved by only having comparative information between each sample ξ_t and *one* point. The next example gives a negative answer to this question, showing that it is impossible to always find the optimal solution in this case even if one allows the algorithm to be a randomized one. Thus, it verifies the necessity of having comparative information at two points in each iteration (for each sample).

Example 4. Let $h(x, \xi) = (x - \xi)^2$, $\ell = -1$, and $u = 1$. In this case, the optimization problem (1) is to find the projection of the expected value of ξ onto the interval $[-1, 1]$. Suppose there are two underlying distributions for ξ. In the first case, ξ follows a uniform distribution on $[-3, -2]$ or $[2, 3]$, each with probability 0.5. In the second case, ξ follows a uniform distribution on $[-3, -2]$ or $[3, 4]$, each with probability 0.5. In the following, we denote the distributions corresponding to the first and second cases by $F_1(\cdot)$ and $F_2(\cdot)$, respectively. It is easy to verify that, in the first case, the optimal solution to (1) is $x^* = 0$, and in the second case, the optimal solution to (1) is $x^* = 0.25$. And it is also easy to verify that these settings (both cases) satisfy Assumption 1.

Now we consider any algorithm that only utilizes the comparative information between ξ_t and one decision point x_t in each iteration (however, the point has to be chosen between $[\ell, u]$ because we can modify $h(x, \xi)$ such that it is undefined or ∞ on $x \notin [\ell, u]$). Suppose the algorithm maps x_t and the comparative information between ξ_t and the chosen decision point to a distribution of x_{t+1} (thus, we allow randomized algorithm).

Note that, for any $x \in [\ell, u]$, $F_1(x) = F_2(x) = 0.5$. In other words, there is a 0.5 probability that $\xi > x$ and a 0.5 probability that $\xi < x$ no matter whether ξ is drawn from $F_1(\cdot)$ or $F_2(\cdot)$. For any algorithm, the distribution of each x_t is the same under either case. Thus, no algorithm can return the optimal solution in both cases. In other words, no algorithm can guarantee to solve (1) with comparative information between each sample and only one point, even under Assumption 1. □

3. Choice of f_- and f_+

In the CBA, one important step is the specification of the two sets of density functions $f_-(x,z)$ and $f_+(x,z)$. In the last section, we only said that f_- and f_+ need to satisfy Conditions 1–3 but did not give any specific examples. Nor did we discuss what are good choices of f_- and f_+. In this section, we address this issue by first showing several examples of f_- and f_+ that could be useful in practice and then discussing the effect of choices of f_- and f_+ on the efficiency of the algorithms. We start with the following examples of choices of f_- and f_+.

Example 5 (Uniform Sampling Distribution). Suppose the support of ξ is known to be contained in a finite interval $[\underline{s}, \bar{s}]$ and the optimal decision x^* is known to be within a finite interval $[\ell, u]$. (Without loss of generality, we assume $[\underline{s}, \bar{s}] \subseteq [\ell, u]$. Otherwise, we can expand $[\ell, u]$ to contain $[\underline{s}, \bar{s}]$.) And we assume $h''_{x,z}(x,z)$ is uniformly bounded on $[\ell, u] \times [\underline{s}, \bar{s}]$, $x \neq z$. Then we can set both $f_-(x,z)$ and $f_+(x,z)$ to be uniformly distributed, that is, for $x \in (\ell, u)$,

$$f_-(x,z) = \begin{cases} \frac{1}{x-\ell} & \ell \le z < x \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad f_+(x,z) = \begin{cases} \frac{1}{u-x} & x < z \le u \\ 0 & \text{otherwise.} \end{cases}$$

When $x = \ell$, we can set f_- to be a uniform distribution on $[\ell - 1, \ell]$, and when $x = u$, we can set f_+ to be a uniform distribution on $[u, u+1]$. It is not hard to verify that this set of choices satisfy Conditions 1–3.

Example 6 (Exponential Sampling Distribution). Suppose the support of ξ is $\mathbb{R}$ or unknown, and ξ follows a light tail distribution (more precisely, there exists a constant $\bar{\lambda} > 0$ such that $\lim_{t\to\infty} e^{\bar{\lambda}t}\mathbb{P}(|\xi| > t) = 0$). Moreover, x is constrained on a finite interval $[\ell, u]$, and we assume $h''_{x,z}(x,z)$ is uniformly bounded on $[\ell, u] \times \mathbb{R}$, $x \neq z$. Then we can choose f_- and f_+ to be exponential distributions. More precisely, we can choose

$$f_-(x,z) = \begin{cases} 0 & z \ge x \\ \lambda_- \exp(-\lambda_-(x-z)) & z < x \end{cases} \quad \text{and} \quad f_+(x,z) = \begin{cases} \lambda_+ \exp(-\lambda_+(z-x)) & z > x \\ 0 & z \le x, \end{cases}$$

where $0 < \lambda_-, \lambda_+ < \bar{\lambda}$ are two parameters one can adjust. Apparently, under this choice of f_- and f_+, Conditions 1–2 are satisfied. For Condition 3, we note that, by the light tail assumption, there exists a constant C such that $e^{-\bar{\lambda}t}F(t) \le C$ and $e^{\bar{\lambda}t}(1 - F(t)) \le C$ for all t. Therefore, we have

$$\int_{-\infty}^{x-} \frac{F(z)}{f_-(x,z)}dz = \frac{e^{\lambda_- x}}{\lambda_-}\int_{-\infty}^{x} e^{-\lambda_- z}F(z)dz \le \frac{Ce^{\lambda_- x}}{\lambda_-}\int_{-\infty}^{x} e^{(\bar{\lambda}-\lambda_-)z}dz \le \frac{Ce^{\bar{\lambda}x}}{(\bar{\lambda}-\lambda_-)\lambda_-},$$
$$\int_{x+}^{\infty} \frac{1-F(z)}{f_+(x,z)}dz = \frac{e^{-\lambda_+ x}}{\lambda_+}\int_{x}^{\infty} e^{\lambda_+ z}(1-F(z))dz \le \frac{Ce^{-\lambda_+ x}}{\lambda_+}\int_{x}^{\infty} e^{-(\bar{\lambda}-\lambda_+)z}dz \le \frac{Ce^{-\bar{\lambda}x}}{(\bar{\lambda}-\lambda_+)\lambda_+}.$$

Combined with the uniform boundedness of $h''_{x,z}(x,z)$, Condition 3 also holds in this case.

Next, we discuss the effect of choosing different f_- and f_+ on the efficiency of the algorithm and the optimal choices of f_- and f_+. First we note that, by Propositions 2–4, the choice of f_- and f_+ does not affect the asymptotic convergence rate of the algorithms as long as they satisfy Conditions 1–3. All that they affect is the constant in the convergence results, which depends on $\mathbb{E}_{z,\xi}(g(x,\xi,z))^2$ or $\mathbb{E}_{z,\xi}(g(x,\xi,z) - H'(x))^2$ (depending on whether Assumption 5(a) or (b) holds). However, by Proposition 1, $\mathbb{E}_{z,\xi}(g(x,\xi,z) - H'(x))^2 = \mathbb{E}_{z,\xi}(g(x,\xi,z))^2 - (H'(x))^2$, that is, the two terms only differ by a constant that does not depend on the choice of f_- and f_+. Therefore, in what follows, we focus on choosing f_- and f_+ to minimize $\mathbb{E}_{z,\xi}(g(x,\xi,z))^2$.

By (6), for any $\xi < x$, we have

$$\mathbb{E}_z(g(x,\xi,z))^2 = 2h'_-(x)h'_x(x,\xi) - (h'_-(x))^2 + \int_{\xi}^{x-} \frac{(h''_{x,z}(x,z))^2}{f_-(x,z)}dz.$$

By a similar argument, for $\xi > x$, we have

$$\mathbb{E}_z(g(x,\xi,z))^2 = 2h'_-(x)h'_x(x,\xi) - (h'_-(x))^2 + \int_{x+}^{\xi} \frac{(h''_{x,z}(x,z))^2}{f_+(x,z)}dz.$$

Further taking expectation over ξ, we have

$$\begin{aligned}
&\mathbb{E}_{z,\xi}(g(x,\xi,z))^2 \\
&= 2h'_-(x)H'(x) - (h'_-(x))^2 + \int_{-\infty}^{x-}\left(\int_{\xi}^{x-}\frac{(h''_{x,z}(x,z))^2}{f_-(x,z)}dz\right)dF(\xi) + \int_{x+}^{\infty}\left(\int_{x+}^{\xi}\frac{(h''_{x,z}(x,z))^2}{f_+(x,z)}dz\right)dF(\xi) \\
&= 2h'_-(x)H'(x) - (h'_-(x))^2 + \int_{-\infty}^{x-}\frac{F(z)(h''_{x,z}(x,z))^2}{f_-(x,z)}dz + \int_{x+}^{\infty}\frac{(1-F(z))(h''_{x,z}(x,z))^2}{f_+(x,z)}dz.
\end{aligned}$$

By the Cauchy–Schwarz inequality and Condition 1, we have

$$\int_{-\infty}^{x-}\frac{F(z)(h''_{x,z}(x,z))^2}{f_-(x,z)}dz = \int_{-\infty}^{x-}\frac{F(z)(h''_{x,z}(x,z))^2}{f_-(x,z)}dz \cdot \int_{-\infty}^{x-} f_-(x,z)dz \geq \left(\int_{-\infty}^{x}\sqrt{F(z)}|h''_{x,z}(x,z)|dz\right)^2.$$

And the equality holds only if $f_-(x,z) = C_-\sqrt{F(z)}|h''_{x,z}(x,z)|$ for all $z < x$ for some $C_- > 0$. Therefore, if $\sqrt{F(z)}|h''_{x,z}(x,z)|$ is integrable on $(-\infty, x]$, then the optimal choice for $f_-(x,z)$ is

$$f_-(x,z) = \frac{\sqrt{F(z)}|h''_{x,z}(x,z)|}{\int_{-\infty}^{x}\sqrt{F(z)}|h''_{x,z}(x,z)|dz}, \quad \forall z < x. \tag{8}$$

Similarly, if $\sqrt{1-F(z)}|h''_{x,z}(x,z)|$ is integrable on $[x,\infty)$, then the optimal choice for $f_+(x,z)$ is

$$f_+(x,z) = \frac{\sqrt{1-F(z)}|h''_{x,z}(x,z)|}{\int_{x}^{\infty}\sqrt{1-F(z)}|h''_{x,z}(x,z)|dz}, \quad \forall z > x. \tag{9}$$

Now, we use an example to illustrate these results. Suppose $h(x,\xi) = (x-\xi)^2$ and $F(z)$ is a uniform distribution on $[a,b]$. Then the optimal choice of f_- and f_+ are

$$f_-(x,z) = \frac{3(z-a)^{1/2}}{2(x-a)^{3/2}},\ \forall a \leq z < x \leq b \quad \text{and} \quad f_+(x,z) = \frac{3(b-z)^{1/2}}{2(b-x)^{3/2}},\ \forall a \leq x < z \leq b.$$

Similarly, when $h(x,\xi) = (x-\xi)^2$ and $F(z)$ is a normal distribution $\mathcal{N}(a,b^2)$, the optimal choice of f_- and f_+ are (it is easy to show that the integrals on the bottom are finite)

$$f_-(x,z) = \frac{\sqrt{\Phi\left(\frac{z-a}{b}\right)}}{\int_{-\infty}^{x}\sqrt{\Phi\left(\frac{z-a}{b}\right)}dz} \quad \forall z < x \quad \text{and} \quad f_+(x,z) = \frac{\sqrt{1-\Phi\left(\frac{z-a}{b}\right)}}{\int_{x}^{\infty}\sqrt{1-\Phi\left(\frac{z-a}{b}\right)}dz} \quad \forall z > x,$$

where $\Phi(\cdot)$ is the c.d.f. of a standard normal distribution.

However, in many cases, the optimal choice of f_- and f_+ may not exist because either (1) $\sqrt{F(z)}|h''_{x,z}(x,z)|$ or $\sqrt{1-F(z)}|h''_{x,z}(x,z)|$ is not integrable or (2) the integration is zero (e.g., in the case in which $h(x,\xi)$ is piecewise linear in x and ξ). In those cases, either the optimal f_- and f_+ are not attainable or the choice of f_- and f_+ does not matter (e.g., in the piecewise linear case).[7] Moreover, finding the optimal f_- and f_+ essentially needs the knowledge of the distribution of ξ, which is not known in advance. Therefore, one can only use an approximate (or prior) distribution of ξ to calculate the distribution.[8] In addition, sampling from the distributions described usually involves much more computational effort than sampling from a uniform distribution or an exponential distribution, the overhead of which may well overshadow the improvement of the convergence speed. Therefore, in practice, choosing a heuristic sampling distribution f_- and f_+ may be preferable, such as the uniform or exponential distributions described earlier in this section. Indeed, as we see later in Section 5, using uniform or exponential distributions leads to efficient solutions in our test cases.

4. Extensions

In this section, we discuss a few extensions of our comparison-based algorithms. In particular, in Section 4.1, we extend our discussions to multidimensional problems with quadratic objective functions. In Section 4.2, we consider the case in which the objective function is not a convex function. In Section 4.3, we consider the case in which multiple samples can be drawn and multiple comparisons can be conducted in each iteration. We also consider a case in which categorical results (depending on the difference between the decision point and the sampled point) instead of binary results can be obtained from each comparison in Appendix C in the e-companion. Overall, we show that our proposed ideas can still be applied in those settings (with some variations).

4.1. Multidimensional Convex Quadratic Problem

In this section, we extend our model to a multidimensional convex quadratic problem and propose a stochastic optimization algorithm for such a setting based on comparative information. Specifically, we consider the following stochastic convex optimization problem:

$$\min_{x\in\mathcal{X}} H(x) = \mathbb{E}_\xi\left[h(x,\xi) := \frac{1}{2}(x-\xi)^\top Q(x-\xi)\right], \tag{10}$$

where $x \in \mathbb{R}^d$, $\xi \in \mathbb{R}^d$ is a random variable, Q is a positive definite matrix, and $\mathcal{X}$ is a closed convex set in $\mathbb{R}^d$.

We denote the gradient of H by $\nabla H(x) := \mathbb{E}Q(x-\xi)$, the directional derivative of $h(x,\xi)$ along a direction $u \in \mathbb{R}^d$ by $\nabla_u h(x,\xi) := u^T Q(x-\xi)$, and the gradient of $h(x,\xi)$ with respect to x by $\nabla h(x,\xi) := Q(x-\xi)$. The following assumption is made in this section:

Assumption 6. *There exists a constant K_4 such that $\mathbb{E}\|\xi - \mathbb{E}\xi\|_2^2 \le K_4^2$.*

This assumption simply requires that ξ has a finite variance, which is not hard to satisfy in practice.

Suppose we can generate a random vector u in $\mathbb{R}^d$ that satisfies $\mathbb{E}(uu^\top) = I_d$, where I_d is the $d \times d$ identity matrix. We have $\mathbb{E}_u u\nabla_u h(x,\xi) = \mathbb{E}_u uu^\top Q(x-\xi) = \nabla h(x,\xi)$. Hence, to construct an unbiased stochastic gradient for $H(x)$, we only need to construct an unbiased stochastic estimation for $u^\top Q(x-\xi)$ and multiply it to u. In the following, we show that this can be done by first comparing $x+zu$ and $x-zu$ (in the value of $h(\cdot,\xi)$) with a random positive number z and then comparing the better one between $x+zu$ and $x-zu$ with x. In other words, we can still construct a stochastic gradient for $H(x)$ in (10) using two comparisons.

Similar to Example 1, this problem may still represent a problem of finding the average features of a group of customers. In particular, in such problems, x may represent d features (e.g., size, taste, etc.) of a product, and ξ is the preference of a random customer on those features. The firm would like to find the optimal features to minimize the expected customer dissatisfaction, which is measured by $h(x,\xi)$ in (10). Suppose the current product is x. To implement the two comparisons, the firm can generate a random change u and a random level z and then ask a customer for the customer's preference between two new products $x+zu$ and $x-zu$ and the customer's preference between the better new product and the current product.

We call this method the comparison-based algorithm for a quadratic problem (CBA-QP) and provide its details in Algorithm 3.

In CBA-QP, we need to specify a density function $f(z)$ on $[0,+\infty)$, which needs to satisfy the following condition.

Condition 4. *There exists a constant K_5 such that, for any $x \in \mathcal{X}$,*

$$\int_0^{+\infty} \frac{\text{Prob}\left(\frac{\lambda_{\max}(Q)\|x-\xi\|_2}{\lambda_{\min}(Q)\sqrt{d}} \ge z\right)}{f(z)} dz = \iint_0^{\frac{\lambda_{\max}(Q)\|x-\xi\|_2}{\lambda_{\min}(Q)\sqrt{d}}} \frac{1}{f(z)} dz dF(\xi) \le K_5,$$

where $\lambda_{\max}(Q)$ and $\lambda_{\min}(Q)$ are the largest and smallest eigenvalues of Q, respectively.

We give two examples of choices of $f(z)$ such that it satisfies Condition 4.

Example 7 (Uniform Sampling Distribution)**.** Suppose Ξ (the support of ξ) and the feasible set $\mathcal{X}$ are both bounded. The quantity $R := \frac{\lambda_{\max}(Q)}{\lambda_{\min}(Q)\sqrt{d}} \max_{x\in\mathcal{X},\xi\in\Xi}\|x-\xi\|_2$ is finite. We can set $f(z)$ to be the density function of a uniform distribution on $[0,R]$, that is,

$$f(z) = \begin{cases} 1/R & 0 \le z \le R \\ 0 & \text{otherwise.} \end{cases}$$

With this choice, we have

$$\iint_0^{\frac{\lambda_{\max}(Q)\|x-\xi\|_2}{\lambda_{\min}(Q)\sqrt{d}}} \frac{1}{f(z)} dz dF(\xi) \le R^2,$$

and thus Condition 4 is satisfied with $K_5 = R^2$.

Example 8 (Exponential Sampling Distribution)**.** Suppose ξ follows a light tail distribution with $\mathbb{E}\exp(\|\xi\|_2) \le \bar{\sigma}$ for some $\bar{\sigma} > 0$. Moreover, suppose the feasible set $\mathcal{X}$ is bounded. Then we can choose $f(z)$ to be an exponential distribution, that is,

$$f(z) = \begin{cases} 0 & z < 0 \\ \lambda \exp(-\lambda z) & z \ge 0, \end{cases} \tag{11}$$

where λ can be any positive constant less than $c := \frac{\sqrt{d}\lambda_{\min}(Q)}{\lambda_{\max}(Q)}$.

With this choice, we have

$$\int_0^{\frac{\lambda_{\max}(Q)\|x-\xi\|_2}{\lambda_{\min}(Q)\sqrt{d}}} \frac{1}{f(z)} dz \le \frac{1}{\lambda^2} \exp\left(\frac{\lambda_{\max}(Q)\|x-\xi\|_2 \lambda}{\lambda_{\min}(Q)\sqrt{d}}\right) \le \frac{1}{\lambda^2} \exp\left(\frac{\lambda \max_{x\in\mathcal{X}} \|x\|_2}{c}\right) \exp\left(\frac{\lambda\|\xi\|_2}{c}\right).$$

By the light tail assumption and the concavity of $(\cdot)^{\lambda/c}$, we can use Jensen's inequality to show that $\mathbb{E}\exp\left(\frac{\lambda\|\xi\|_2}{c}\right) \le \left[\mathbb{E}\exp(\|\xi\|_2)\right]^{\frac{\lambda}{c}} \le \bar{\sigma}^{\frac{\lambda}{c}}$. Using this inequality, we further have

$$\iint_0^{\frac{\lambda_{\max}(Q)\|x-\xi\|_2}{\lambda_{\min}(Q)\sqrt{d}}} \frac{1}{f(z)} dz dF(\xi) \le \frac{\bar{\sigma}^{\frac{\lambda}{c}}}{\lambda^2} \exp\left(\frac{\lambda \max_{x\in\mathcal{X}} \|x\|_2}{c}\right),$$

and thus Condition 4 is satisfied with $K_5 = \frac{\bar{\sigma}^{\frac{\lambda}{c}}}{\lambda^2} \exp\left(\frac{\lambda}{c} \max_{x\in\mathcal{X}} \|x\|_2\right)$.

Algorithm 3 (CBA-QP)

1. **Initialization.** Set $t = 1$, $x_1 \in \mathcal{X} \subset \mathbb{R}^d$. Define η_t for all $t \ge 1$. Set the maximum number of iterations T.
Choose a density function $f(z)$ on $[0, +\infty)$. Let $\mathcal{Q}$ be the uniform distribution on a sphere in $\mathbb{R}^d$ with radius $\sqrt{d}$. Note that $\mathbb{E}uu^T = I_d$ for u following distribution $\mathcal{Q}$.
2. **Main iteration.** Sample ξ_t from the distribution of ξ. Sample u_t from $\mathcal{Q}$. Sample z_t from $f(z)$.
   a. If $h(x_t + z_t u_t, \xi_t) < h(x_t - z_t u_t, \xi_t)$, set

$$g(x_t, \xi_t, u_t, z_t) = \begin{cases} 0, & \text{if } h(x_t + z_t u_t, \xi_t) > h(x_t, \xi_t), \\ -\frac{1}{2}\frac{u_t^T Q u_t}{f(z_t)} u_t, & \text{if } h(x_t + z_t u_t, \xi_t) \le h(x_t, \xi_t). \end{cases}$$

   b. If $h(x_t + z_t u_t, \xi_t) \ge h(x_t - z_t u_t, \xi_t)$, set

$$g(x_t, \xi_t, u_t, z_t) = \begin{cases} 0, & \text{if } h(x_t - z_t u_t, \xi_t) > h(x_t, \xi_t), \\ \frac{1}{2}\frac{u_t^T Q u_t}{f(z_t)} u_t, & \text{if } h(x_t - z_t u_t, \xi_t) \le h(x_t, \xi_t). \end{cases}$$

Let

$$x_{t+1} = \mathrm{Proj}_{\mathcal{X}}\left(x_t - \eta_t g(x_t, \xi_t, u_t z_t)\right). \tag{12}$$

3. **Termination.** Stop when $t \ge T$. Otherwise, let $t \leftarrow t + 1$ and go back to step 2.
4. **Output. CBA-QP**$(x_1, T, \{\eta_t\}_{t=1}^T) = \bar{x}_T = \frac{1}{T}\sum_{t=1}^T x_t$.

The following proposition gives some properties of the stochastic gradient $g(x_t, \xi_t, u_t, z_t)$ in the CBA-QP (see Algorithm 3):

Proposition 5. *Let z, u, and g be as defined in Algorithm 3. Then the following properties hold:*

1. $\mathbb{E}_{z,u} g(x, \xi, u, z) = \nabla h(x, \xi)$ *for all* $x \in \mathcal{X}$.
2. $\mathbb{E}_{z,u,\xi} g(x, \xi, u, z) = \nabla H(x)$ *for all* $x \in \mathcal{X}$.
3. $\mathbb{E}_{z,u,\xi}\|g(x, \xi, u, z) - \nabla H(x)\|_2^2 \le \sigma^2 := \lambda_{\max}(Q)^2 K_4^2 + \frac{\lambda_{\max}(Q)^2 d^3 K_5}{4}$.

Proof of Proposition 5. First, we consider the case in which $h(x+zu,\xi) < h(x-zu,\xi)$, which is equivalent to $u^TQ(x-\xi) < 0$ by the definition of h and the nonnegativity of z. Note that $h(x+zu,\xi) \le h(x,\xi)$ if and only if $zu^TQ(x-\xi) + \frac{z^2}{2}u^TQu \le 0$ or, equivalently, $0 \le z \le -\frac{2u^TQ(x-\xi)}{u^TQu}$. It then follows from the definition of g that

$$\mathbb{E}_z g(x,\xi,u,z) = \int_0^{-\frac{2u^TQ(x-\xi)}{u^TQu}} -\frac{1}{2}u^TQu \cdot u dz = u^TQ(x-\xi)u = \nabla_u h(x,\xi)u.$$

Similarly, the second case $h(x+zu,\xi) \ge h(x-zu,\xi)$ occurs when $u^TQ(x-\xi) \ge 0$. The inequality $h(x-zu,\xi) \le h(x,\xi)$ holds if and only if $-zu^TQ(x-\xi) + \frac{z^2}{2}u^TQu \le 0$ or, equivalently, $0 \le z \le \frac{2u^TQ(x-\xi)}{u^TQu}$. It then follows again from the definition of g that

$$\mathbb{E}_z g(x,\xi,u,z) = \int_0^{\frac{2u^TQ(x-\xi)}{u^TQu}} \frac{1}{2}u^TQu \cdot u dz = u^TQ(x-\xi)u = \nabla_u h(x,\xi)u.$$

Therefore, in both cases, we have $\mathbb{E}_z g(x,\xi,u,z) = \nabla_u h(x,\xi)u$. Because $\mathbb{E}(uu^T) = I_d$, further taking expectation over u on both sides of this equality gives the first conclusion of the proposition.

The second conclusion can be obtained by taking expectation over ξ on both sides of the first conclusion, namely $\mathbb{E}_{z,u,\xi} g(x,\xi,u,z) = \mathbb{E}_\xi \mathbb{E}_{z,u} g(x,\xi,u,z) = \mathbb{E}_\xi \nabla h(x,\xi) = \nabla H(x)$. (This holds because h is a convex function; thus, one can apply the monotone convergence theorem.)

Next, we prove the third conclusion. The first conclusion implies that

$$\mathbb{E}_{u,z}\|g(x,\xi,u,z) - \nabla h(x,\xi)\|^2 = \mathbb{E}_{u,z}\|g(x,\xi,u,z)\|^2 - \|\nabla h(x,\xi)\|^2.$$

If $h(x+zu,\xi) < h(x-zu,\xi)$, by the definition of g, we can show that

$$\begin{aligned}
\mathbb{E}_z\|g(x,\xi,u,z)\|^2 &= \int_0^{-\frac{2u^TQ(x-\xi)}{u^TQu}} \frac{(u^TQu)^2}{4f(z)}\|u\|_2^2 dz \\
&\le \int_0^{\frac{\lambda_{\max}(Q)\|u\|_2\|x-\xi\|_2}{\lambda_{\min}(Q)\|u\|_2^2}} \frac{\lambda_{\max}(Q)^2\|u\|_2^4}{4f(z)}\|u\|_2^2 dz \\
&= \frac{\lambda_{\max}(Q)^2 d^3}{4}\int_0^{\frac{\lambda_{\max}(Q)\|x-\xi\|_2}{\lambda_{\min}(Q)\sqrt{d}}} \frac{1}{f(z)} dz,
\end{aligned}$$

where the inequality is by the Cauchy–Schwarz inequality and the second equality is because $\|u\| = \sqrt{d}$. Similarly, if $h(x+zu,\xi) \ge h(x-zu,\xi)$, then we can also show that

$$\mathbb{E}_z\|g(x,\xi,u,z) - \nabla h(x,\xi)\|^2 \le \frac{\lambda_{\max}(Q)^2 d^3}{4}\int_0^{\frac{\lambda_{\max}(Q)\|x-\xi\|_2}{\lambda_{\min}(Q)\sqrt{d}}} \frac{1}{f(z)} dz.$$

As a result, we have

$$\mathbb{E}_{u,z}\|g(x,\xi,u,z) - \nabla h(x,\xi)\|^2 \le \frac{\lambda_{\max}(Q)^2 d^3}{4}\int_0^{\frac{\lambda_{\max}(Q)\|x-\xi\|_2}{\lambda_{\min}(Q)\sqrt{d}}} \frac{1}{f(z)} dz.$$

According to this inequality and Condition 4, we have

$$\mathbb{E}_{z,u,\xi}\|g(x,\xi,u,z) - \nabla h(x,\xi)\|^2 \le \frac{\lambda_{\max}(Q)^2 d^3}{4}\iint_0^{\frac{\lambda_{\max}(Q)\|x-\xi\|_2}{\lambda_{\min}(Q)\sqrt{d}}} \frac{1}{f(z)} dz dF(\xi) \le \frac{\lambda_{\max}(Q)^2 d^3 K_5}{4}.$$

In addition, by Assumption 6, we have

$$\mathbb{E}_\xi\|\nabla h(x,\xi) - \nabla H(x)\|_2^2 = \mathbb{E}_\xi\|Q(x-\xi) - Q(x-\mathbb{E}\xi)\|_2^2 \le \lambda_{\max}(Q)^2\mathbb{E}_\xi\|\xi - \mathbb{E}\xi\|_2^2 \le \lambda_{\max}(Q)^2 K_4^2.$$

Finally, we note that

$$\mathbb{E}_{z,u,\xi}\|g(x,\xi,u,z) - \nabla H(x)\|_2^2 = \mathbb{E}_{z,u,\xi}\|g(x,\xi,u,z) - \nabla h(x,\xi)\|_2^2 + \mathbb{E}_\xi\|\nabla h(x,\xi) - \nabla H(x)\|_2^2.$$

Therefore, $\mathbb{E}_{z,u,\xi}\|g(x,\xi,u,z) - \nabla H(x)\|_2^2 \le \sigma^2 := \lambda_{\max}(Q)^2 K_4^2 + \frac{\lambda_{\max}(Q)^2 d^3 K_5}{4}$. Thus, the proposition holds. □

Based on Proposition 5, we have the following proposition about the performance of the CBA-QP.

Proposition 6. *Let σ^2 be defined as in Proposition 5. Let x^* be any optimal solution to (10) and $\mu > 0$ and $L > 0$ be the smallest and largest eigenvalues of Q, respectively. Then, by choosing $\eta_t = \frac{1}{\mu t + L}$, the CBA-QP ensures that*

$$\mathbb{E}(H(\bar{x}_T) - H(x^*)) \le \frac{\sigma^2}{2\mu}\frac{\log T + 1}{T} + \frac{H(x_1) - H(x^*)}{T} + \frac{L\|x_1 - x^*\|_2^2}{2T}, \text{ and}$$

$$\sum_{t=1}^{T} \mathbb{E}(H(x_t) - H(x^*)) \le \frac{\sigma^2}{2\mu}(\log T + 1) + H(x_1) - H(x^*) + \frac{L\|x_1 - x^*\|_2^2}{2}.$$

Similar to the CBA, we can further improve the theoretical performance of the CBA-QP using a restarting strategy as described in the MCBA. We denote the resulting restarted algorithm by MCBA-QP.

Algorithm 4 (MCBA-QP)

1. Initialize the number of stages $K \ge 1$, the starting solution $\hat{x}^1$. Set $k = 1$.
2. Let T_k be the number of iterations in stage k and η_t^k be the step length in iteration t of the CBA-QP in stage k for $t = 1, 2, \ldots, T_k$.
3. Let $\hat{x}^{k+1} = \textbf{CBA-QP}(\hat{x}^k, T_k, \{\eta_t^k\}_{t=1}^{T_k})$.
4. Stop when $k \ge K$. Otherwise, let $k \leftarrow k + 1$ and go back to step 2.
5. Output $\hat{x}^{K+1}$.

Proposition 7. *Let σ^2 be defined as in Proposition 5, $\mu > 0$ and $L > 0$ be defined as in Proposition 6, and $T = \sum_{k=1}^{K} T_k$ with T_k defined in the MCBA-QP. By choosing $\eta_t^k = \frac{1}{2^{k+1}\mu + L}$ and $T_k = 2^{k+3} + 4$, the MCBA-QP ensures that*

$$\mathbb{E}(H(\hat{x}_{K+1}) - H(x^*)) \le \frac{32(H(\hat{x}_1) - H(x^*) + L\|\hat{x}_1 - x^*\|_2^2/2 + \sigma^2/\mu)}{T}.$$

The proofs of Propositions 6 and 7 are very similar to those of the second parts of Propositions 3 and 4, respectively, and are provided in Appendix B in the e-companion.

Although we mainly focus on the quadratic problem for the multidimensional case, it is easy to see that our approach can also be applied to the multidimensional cases in which the objective function is separable with respect to each dimension while the constraint may not be separable. In particular, our method can be applied to the following problem:

$$\min_{x \in \mathcal{X}} H(x) = \mathbb{E}_\xi\left[h(x, \xi) := \sum_{i=1}^{d} h_i(x_i, \xi_i)\right],$$

where $x = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d$, $\xi = (\xi_1, \ldots, \xi_d)^\top \in \mathbb{R}^d$, $\mathcal{X}$ is a convex closed set in $\mathbb{R}^d$, and $h_i(x_i, \xi_i)$ is a function satisfying Assumption 1. In such cases, one can apply the idea of the CBA to construct an unbiased stochastic gradient of each $h_i(x_i, \xi_i)$ based on comparisons and then concatenate them into an unbiased stochastic gradient of $h(x, \xi)$ in order to apply the projected gradient step (12). The resulting algorithm has the same convergence rates as the CBA under each setting in Propositions 2–4. Note that, we do not assume $\mathcal{X}$ is separable, so we still cannot solve the preceding problem as d independent problems.

4.2. Nonconvex Problem

The focus of our study in this paper is to construct an unbiased stochastic gradient based on comparative information. Once the construction is done, one can also apply the stochastic gradient to nonconvex stochastic optimization problems. Although the stochastic gradient method can no longer guarantee to reach a global optimal solution for a nonconvex problem, the recent result by Davis and Drusvyatskiy (2018) shows that the iterative solution generated by stochastic gradient method still converges to a nearly stationary point under some conditions.

In this section, we still consider one-dimensional problem (1) but $H(x)$ is no longer convex. More specifically, we still assume Assumption 1 holds except that Assumption 4 is replaced by Assumption 4′, $H(x)$ in (1) is differentiable and ρ-*weakly convex* on $[\ell, u]$ for some $\rho > 0$, namely

$$H(x_2) \ge H(x_1) + H'(x_1)(x_2 - x_1) - \frac{\rho}{2}(x_2 - x_1)^2, \quad \forall x_1, x_2 \in [\ell, u]. \tag{13}$$

Moreover, $\mathbb{E}_\xi(h'_x(x, \xi)) = H'(x)$ for all $x \in [\ell, u]$.

For any $\lambda > 0$, the *Moreau envelope* for (1) is defined as a function

$$H_\lambda(x) := \min_{\ell \le y \le u} \left\{ H(y) + \frac{1}{2\lambda}(x-y)^2 \right\}. \tag{14}$$

By definition, as long as $\lambda < \frac{1}{\rho}$, the minimization problem (14) has a strongly convex objective function and has a unique solution, denoted by $\hat{x}$. Moreover, the function $H_\lambda(x)$ is continuously differentiable with the gradient given by $H'_\lambda(x) := \frac{1}{\lambda}(x - \hat{x})$. According to Davis and Drusvyatskiy (2018), the value of $|H'_\lambda(x)|$ measures the near stationarity of a solution $x \in [\ell, u]$ because

$$|x - \hat{x}| = \lambda |H'_\lambda(x)| \quad \text{and} \quad \operatorname{dist}(0; \partial H(\hat{x})) \le |H'_\lambda(x)|,$$

where $\partial H(\hat{x})$ is the subdifferential of H at x, that is, the set of all v satisfying $H(y) \ge H(x) + v(y-x) + o(\|y-x\|)$ as $y \to x$ and $\operatorname{dist}(x; A)$ is the nearest distance between point x and set A defined as $\operatorname{dist}(x; A) = \inf_{y \in A} \|x - y\|$. This means that, if $|H'_\lambda(x)| = \frac{1}{\lambda}|x - \hat{x}|$ is small, then the solution x is closed to a point $\hat{x}$ (because of small $|x - \hat{x}|$), which is nearly stationary (because of small $\operatorname{dist}(0; \partial H(\hat{x}))$).

According to Davis and Drusvyatskiy (2018), in order to find a solution x with a small $|H'_\lambda(x)|$, we can still use the CBA except that the output will be x_{t^*}, where t^* is a random index such that $\mathbb{P}(t^* = t) = \frac{\eta_t}{\sum_{s=1}^T \eta_s}$ for $t = 1, 2, \ldots, T$. Then, we have $\mathbb{E}|H'_\lambda(x_{t^*})|$ with $\lambda = \frac{1}{2\rho}$ converging to zero at a rate of $O(\frac{1}{\sqrt{T}})$. We state this result formally as follows.

Proposition 8 (Corollary 2.2 in Davis and Drusvyatskiy (2018))**.** *Suppose Assumption 4 is replaced by Assumption 4′ and the optimal value of* (1) *is finite. Let G^2 be defined as in Proposition 1 and t^* be a random index such that $\mathbb{P}(t^* = t) = \frac{\eta_t}{\sum_{s=1}^T \eta_s}$ for $t = 1, 2, \ldots, T$. By choosing $\eta_t = \frac{1}{\sqrt{T}}$, the CBA ensures that*

$$\mathbb{E}|H'_{\frac{1}{2\rho}}(x_{t^*})|^2 \le 2\frac{(H_{\frac{1}{2\rho}}(x_1) - \min_{\ell \le x \le u} H(x)) + \rho G^2}{\sqrt{T}}.$$

Therefore, the comparison-based algorithm can be partly applied to nonconvex problems.

Remark 1. Recently, there has been growing interest in the convergence of first-order methods when the objective function is nonconvex but satisfies the Kurdyka–Łojasiewicz (KL) property (Bolte et al. 2007, 2014; Attouch et al. 2010; Noll 2014; Karimi et al. 2016; Li et al. 2017; Xu and Yin 2017). A function $H(x)$ satisfies the KL property at a point x^* if there exists $\eta > 0$, a neighborhood U of x^*, and a concave function $\kappa : [0, \eta] \to [0, +\infty)$ such that (i) $\kappa(0) = 0$, (ii) κ is continuously differentiable on $(0, \eta)$, (iii) $\kappa'(\cdot) > 0$ on $(0, \eta)$, and (iv)

$$\kappa'(H(x) - H(x^*)) \cdot \operatorname{dist}(0; \partial H(x^*)) \ge 1$$

for any $x \in U$ that satisfies $H(x^*) < H(x) < H(x^*) + \eta$. Existing results show that, if $H(x)$ is nonconvex but satisfies the KL property, each bounded sequence generated by a first-order method converges to a stationary point x^* of $H(x)$, and the local convergence rate to x^* can be characterized by the geometry of κ near x^* (see, e.g., proposition 3 in Attouch et al. 2010).

However, almost all existing studies on first-order methods under the KL property focus on deterministic cases. Because our approaches construct stochastic gradients, most of the existing results utilizing the KL property cannot be directly applied to our problem. The only result we are aware of about stochastic gradient descent under the KL property is given by Karimi et al. (2016). However, they make stronger assumptions than the KL property on the problem. In particular, they require that the problem be unconstrained, that the aforementioned neighborhood U be the entire space $\mathbb{R}$, and that κ be in the form of $\kappa(t) = c\sqrt{t}$ for some $c > 0$. If these conditions are satisfied by $H(x)$, the CBA can also find an ϵ-optimal solution within $O(1/\epsilon)$ iterations according to theorem 4 in Karimi et al. (2016). This is because the CBA is essentially a stochastic gradient descent method except that the stochastic gradient is constructed using comparative information.

4.3. Minibatch Method with Additional Comparisons

In this section, we consider the scenario in which multiple comparisons can be conducted in each iteration. In such cases, a minibatch technique can be implemented in the CBA and MCBA to reduce the noise in the stochastic gradient and improve the performance of the algorithms. In particular, we still compare ξ_t and x_t in iteration t in Algorithm 1. In case (a), in which $\xi_t < x_t$, we generate S independent samples, denoted by z_t^s for $s = 1, 2, \ldots, S$ from a distribution on $(-\infty, x_t]$ with p.d.f. $f_-(x_t, z)$ and construct a stochastic gradient $g(x_t, \xi_t, z_t^s)$ as

in (3) with z_t replaced by z_t^s. Similarly, in case (b), in which $\xi_t > x_t$, we generate z_t^s from a distribution on $[x_t, +\infty)$ with p.d.f. $f_+(x_t, z)$ and construct $g(x_t, \xi_t, z_t^s)$ as in (4) with z_t replaced by z_t^s. After obtaining $g(x_t, \xi_t, z_t^s)$ in either case, we replace (5) in the CBA with the following two steps:

$$\bar{g}_t = \frac{1}{S}\sum_{s=1}^{S} g(x_t, \xi_t, z_t^s)$$
$$x_{t+1} = \mathrm{Proj}_{[\ell,u]}(x_t - \eta_t \bar{g}_t) = \max(\ell, \min(u, x_t - \eta_t \bar{g}_t)).$$

Here, $\bar{g}_t$ is the average gradient constructed by a minibatch, which satisfies $\mathbb{E}_{z_t^s, s=1,\ldots,S,\xi_t}\bar{g}_t = H'(x_t)$ by conclusion 2 in Proposition 1 and has a smaller noise than $g(x_t, \xi_t, z_t)$. The MCBA can also benefit from this technique by calling the CBA after the aforementioned modification.

This minibatch technique does not improve the asymptotic convergence rate of the CBA and MCBA because it does not completely eliminate the noise in the stochastic gradient. However, by reducing the noise, it improves the algorithm's performance in practice as we demonstrate in Section 5.3.[9]

5. Numerical Tests

In this section, we conduct numerical experiments to show that, although much less information is used, the proposed algorithms based only on comparative information converge at the same rate as the stochastic gradient methods. We also investigate the impact of choices of f_- and f_+ and the impact of a minibatch on the performances of the proposed methods. We implemented all algorithms in MATLAB running on a 64-bit Microsoft Windows 10 machine with a 2.70 Ghz Intel Core i7-6,820HQ CPU and 8 GB of memory.

5.1. Convergence of Objective Value

In this section, we conduct numerical experiments to test the performance of the CBA and the MCBA. We consider two objective functions:

$$(1)\quad h_1(x,\xi) = (x-\xi)^2 \qquad \text{and} \qquad (2)\quad h_2(x,\xi) = \begin{cases} (x-\xi)^2 + (x-\xi) & \text{if } \xi < x \\ 2(x-\xi)^2 + 2(\xi - x) & \text{if } \xi \geq x. \end{cases}$$

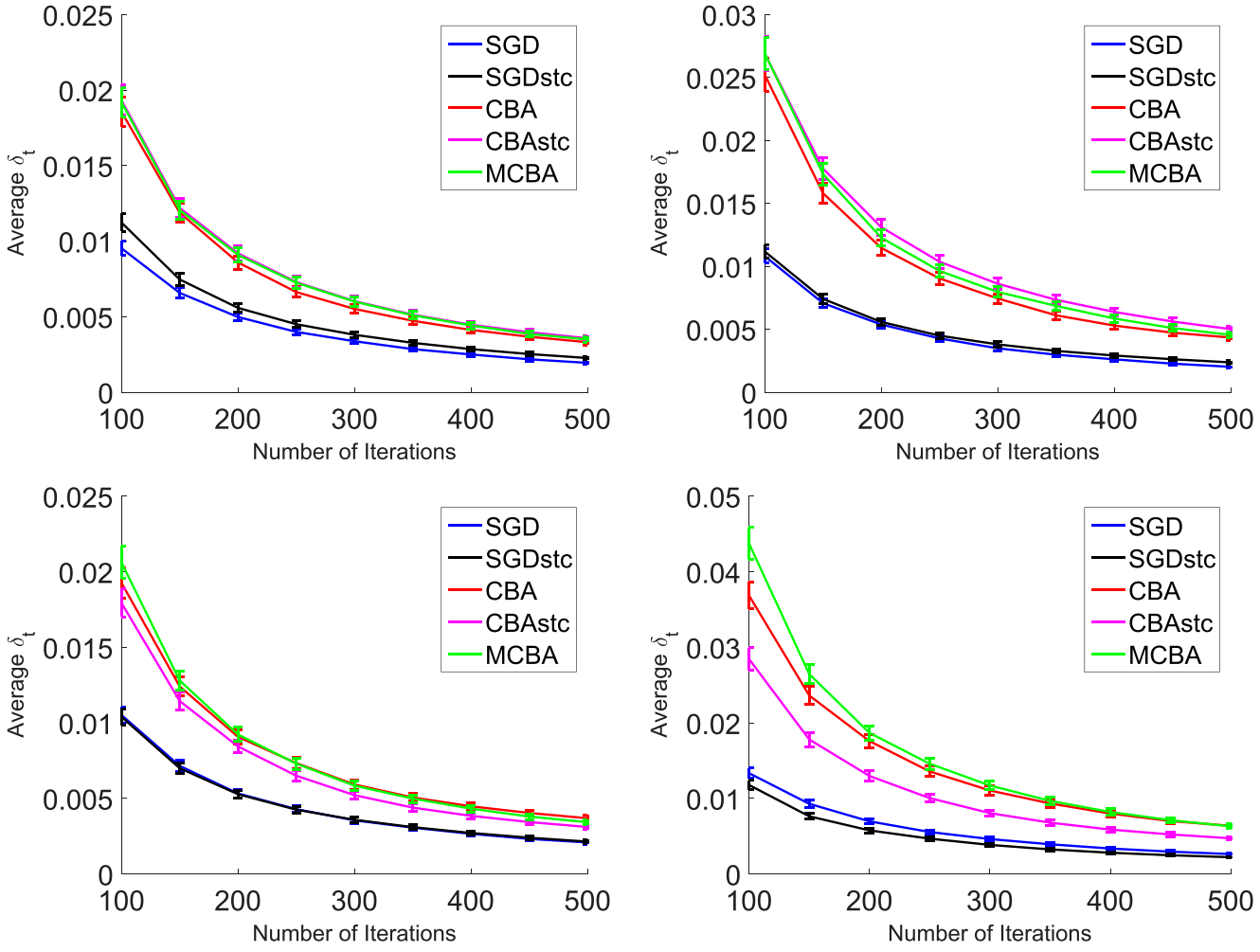
Here, the two objective functions correspond to the two examples described in the beginning with h_1 being smooth but not h_2. For each choice of $h(x,\xi)$, we consider two distributions of ξ, a uniform distribution $\mathcal{U}[50,150]$ and a normal distribution $\mathcal{N}(100,100)$. Thus, we have four cases in total. It is easy to see that, for the first objective function, the optimal solution under either underlying distribution is $x^* = 100$. For the second objective function, the optimal solution under the uniform distribution is $x^* = 108.66$, and the optimal solution under the normal distribution is $x^* = 102.82$. In all experiments, we choose the feasible set to be $[\ell, u] = [50, 150]$. For the cases in which $\xi \sim \mathcal{U}[50,150]$, we choose f_- and f_+ to be uniform distributions as in Example 5 with $\ell = 50$ and $u = 150$. For the cases in which $\xi \sim \mathcal{N}(100,100)$, we choose f_- and f_+ to be exponential distributions as in Example 6 with $\lambda_+ = \lambda_- = 2^{-4} = 0.0625$. (Later in Section 5.2 we test the effect of choosing different f_- and f_+ on the convergence speed of the algorithms.)

In each of these four settings, Assumption 1 holds with Assumption 5(a) and (b) satisfied, and $H(x)$ is strongly convex. To compare the performance of the CBA with or without utilizing the strong convexity of the problem, we test both step lengths $\eta_t = 1/\sqrt{t}$ and $\eta_t = 1/(\mu t)$ for the CBA as suggested by Propositions 2 and 3. For the MCBA, we use the step sizes $\eta_t^k = 1/(2^{k+1}\mu)$ and $T_k = 2^{k+3}$ as suggested by Proposition 4. We choose $\mu = 0.5$ in all settings.

In the following, we compare the CBA and the MCBA with the standard stochastic gradient descent (SGD) method (see Nemirovski et al. (2009) and Duchi and Singer (2009)). In the standard SGD method, it is assumed that ξ_t can be observed directly and the stochastic gradient update step $x_{t+1} = \mathrm{Proj}_{[\ell,u]}(x_t - \eta_t h_x'(x_t, \xi_t))$ is performed in each iteration. In the experiments, we apply the same step lengths in the SGD method as in the CBA. In all tests, we start from a random initial point $x_1 \sim \mathcal{U}[50,150]$ and run each algorithm for $T = 500$ iterations, and we report the average relative optimality gap $\delta_t = \frac{H(\bar{x}_t) - H(x^*)}{H(x^*)}$ over 2,000 independent trails for each t. The results are reported in Figure 1.

In Figure 1, the x-axis represents the number of iterations and the y-axis represents the average relative optimality gap for each algorithm. The curves "CBA" and "SGD" represent the results of the CBA and SGD with $\eta_t = 1/\sqrt{t}$, respectively, and the curves "CBAstc" and "SGDstc" represent the results of the CBA and SGD with $\eta_t = 1/(\mu t)$, respectively. The curve "MCBA" represents the results of the MCBA algorithm. In each of the curves, the bar at each point represents the standard error of the corresponding δ_t. As one can see, the

Figure 1. (Color online) The Convergence of the Average Relative Optimality Gap for Different Algorithms for the Four Instances



Notes. First row: $h_1(x,\xi)$; second row: $h_2(x,\xi)$. First column: $\xi \sim \mathcal{U}[50,150]$; second column: $\xi \sim \mathcal{N}(100,100)$.

standard errors are fairly small. Thus, the test results are quite stable in these numerical experiments. We also present the average computation time of all algorithms in Table 1.

From the results shown in Figure 1, we can see that all of the CBA, CBAstc, and MCBA converge quite fast in these problems. Even though they use much less information than the SGD and SGDstc methods, it takes only about twice as many iterations to get the same accuracy. As shown in Table 1, the computation time for 500 iterations is less than 0.1 seconds in each algorithm, and the time is not much different across different algorithms. (This short runtime is because of the low dimensionality of the problems.) Moreover, in our tests, the CBA, CBAstc, and MCBA have quite similar performances despite their different theoretical guarantees.[10]

Table 1. The Computation Time (in Seconds) of Different Algorithms for 500 Iterations for the Four Instances

	$h_1(x,\xi)$		$h_2(x,\xi)$	
Algorithm	$\xi \sim \mathcal{U}[50,150]$	$\xi \sim \mathcal{N}(100,100)$	$\xi \sim \mathcal{U}[50,150]$	$\xi \sim \mathcal{N}(100,100)$
SGD	0.008	0.007	0.010	0.040
SGDstc	0.007	0.007	0.009	0.038
CBA	0.011	0.018	0.013	0.049
CBAstc	0.011	0.017	0.013	0.045
MCBA	0.013	0.021	0.017	0.055

Table 2. The Computation Time (in Seconds) of the CBA for 500 Iterations When Using Different Numbers of Comparisons (S) per Iteration

	$h_1(x,\xi)$		$h_2(x,\xi)$	
S	$\xi \sim \mathcal{U}[50,150]$	$\xi \sim \mathcal{N}(100,100)$	$\xi \sim \mathcal{U}[50,150]$	$\xi \sim \mathcal{N}(100,100)$
1	0.011	0.018	0.013	0.049
2	0.015	0.022	0.020	0.053
5	0.021	0.025	0.022	0.057
10	0.026	0.033	0.029	0.064
100	0.117	0.158	0.125	0.175

Table 3. The Computation Time (in Seconds) of Different Algorithms for 2,000 Iterations for the Four Instances

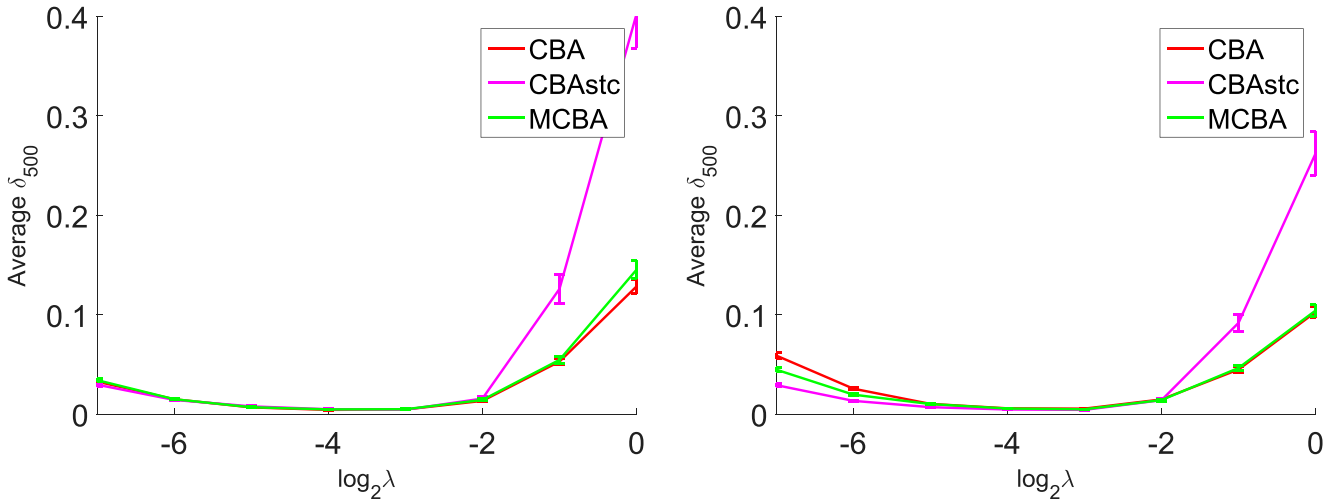
Algorithm	$d = 5$	$d = 20$
SGD	0.236	0.344
CBA-QP	0.391	0.490
MCBA-QP	0.440	0.585

5.2. Choices of f_- and f_+

In this section, we perform some additional tests to study the impact of different choices of f_- and f_+ on the performance of the CBA. We still consider the two objective functions considered in the last section, but we focus on the case in which $\xi \sim \mathcal{N}(100,100)$. First, we keep f_- and f_+ to be exponential distributions (as in Example 6) and see how the performance is affected by different values of $\lambda_- = \lambda_+ = \lambda$. The results are shown in Figure 2.

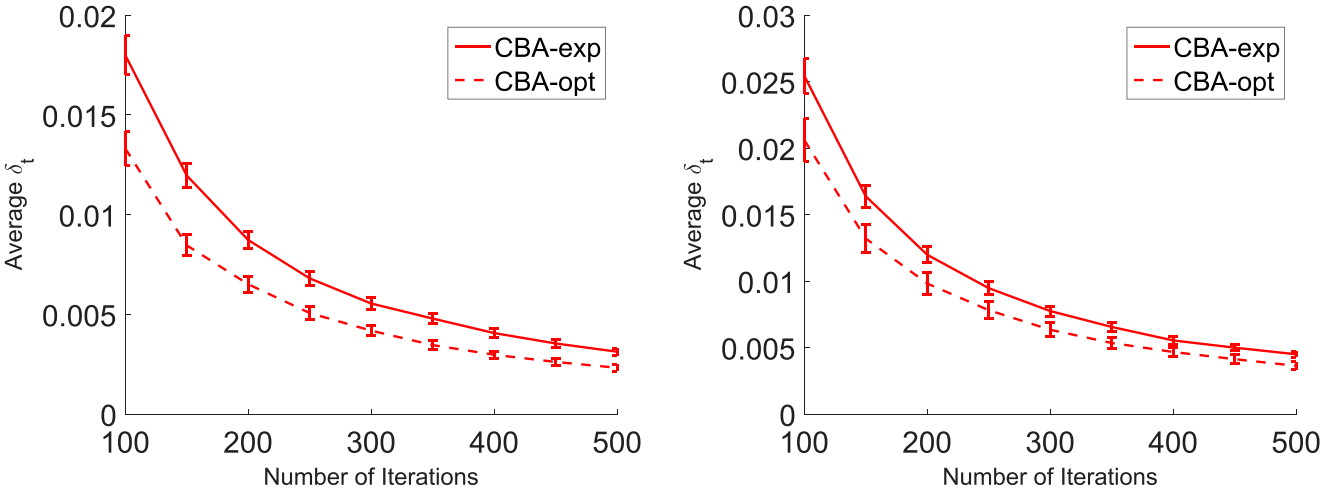
In Figure 2, the x-axis represents the value of $\log_2 \lambda$, and the y-axis represents the relative optimality gap δ_T after 500 iterations evaluated as the average value of 2,000 independent trials (again the bars show the standard error in these trials). The influences of different values of λ in f_- and f_+ are presented for the CBA, CBAstc, and MCBA algorithms. We can see that the value of λ does influence the convergence speed of the algorithms. Particularly, in both panels in Figure 2, the optimality gap after $T = 500$ iterations decreases first as λ increases and starts to increase when λ is large. Moreover, the influence is relatively small when λ is small and is large when λ is large. And the influence is more pronounced for the CBAstc algorithm. In our setting, the best performance is obtained around $\lambda = 2^{-4} = 0.0625$ for all algorithms.

Figure 2. (Color online) The Impact of λ in f_+ and f_- to Different Algorithms, Measured by the Average Relative Optimality Gap After 500 Iterations



Notes. Left: $h_1(x,\xi)$; right: $h_2(x,\xi)$. In both cases, $\xi \sim \mathcal{N}(100,100)$.

Figure 3. (Color online) The Convergence of the Average Relative Optimality Gap in the CBA When Choosing f_+ and f_- to Be the Exponential Distribution and the Optimal Distribution in (8) and (9)



Notes. Left: $\xi \sim \mathcal{U}[50, 150]$; right: $\xi \sim \mathcal{N}(100, 100)$. In both cases, $h(x, \xi) = (x - \xi)^2$.

In Section 3, we provided the optimal choice of f_- and f_+ in (8) and (9). In Figure 3, we present the difference in the performances of the CBA when f_- and f_+ are chosen optimally versus when they are chosen as exponential distributions with $\lambda_- = \lambda_+ = 0.0625$. In this experiment, we choose $h(x, \xi) = h_1(x, \xi)$ and run the CBA for 500 iterations and compute the average relative optimality gap δ_{500} over 2,000 independent trials. The results are plotted in Figure 3.

In Figure 3, we can see that the optimal choice of f_- and f_+ does improve the performance of the CBA, which confirms our analysis in Section 3. However, the improvement is not essential, yet generating samples from the optimal distribution is much more time-consuming. For example, when $\xi \sim \mathcal{N}(100, 100)$, the computational time is less than 0.05 seconds when using exponential distribution but about one second when using the optimal distribution. Therefore, as we discussed in the end of Section 3, one can just choose a simple distribution in practice.

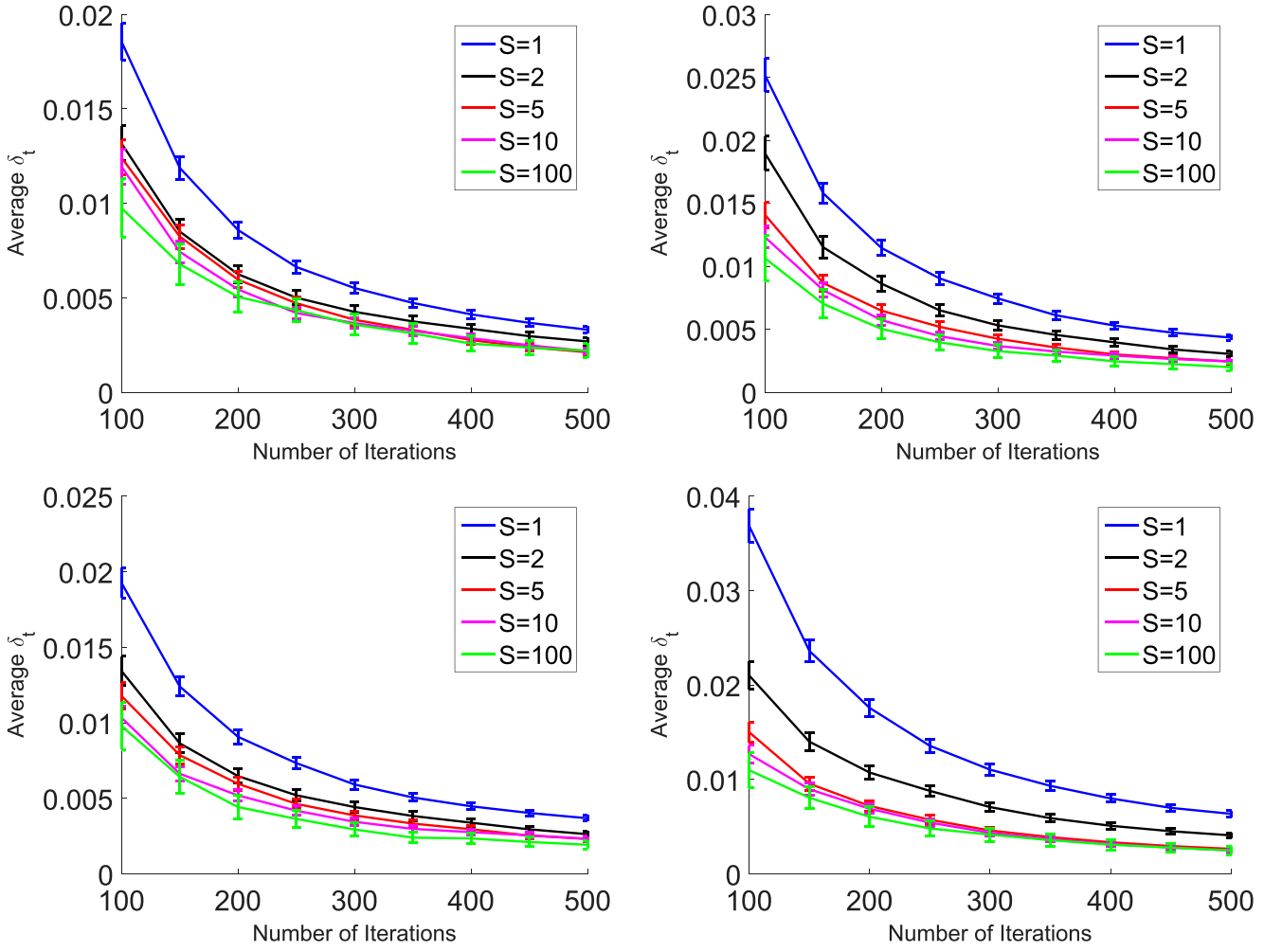
5.3. Minibatch Method with Additional Comparisons

In this section, we numerically test how the performance of the CBA depends on the sample size S in the minibatch technique described in Section 4.3. The instances and the choice of parameters are all identical to Section 5.1. We present the convergence of the CBA with $S = 1, 2, 5, 10, 100$ in Figure 4 and the associated runtimes in Table 2. According to the figures, one additional comparison (i.e., $S = 2$) with z can improve the convergence of Algorithm 1 in all four instances. Although increasing S can still improve the performance further, the effect diminishes quickly. This is because the noise in the stochastic gradient is generated from the sample noise of both ξ and z. The minibatch technique for sampling z does not help reduce the noise because of ξ, which eventually dominates the noise because of z when S is large enough.[11] Although a similar minibatch technique can be applied to ξ, it might not be practical when applying our algorithm in practice. For example, when ξ represents the ideal product of a customer, creating a minibatch for ξ means asking different customers' preferences without updating the solution, which is not consistent with the setting of online optimization in which the solution is updated after the visit of each customer.

5.4. Multidimensional Problems

In this section, we conduct numerical experiments to test the performance of the CBA-QP and MCBA-QP on the multidimensional stochastic quadratic program (10). To generate a testing instance of (10), we first generate a $d \times d$ matrix Q' in which each entry is sampled from an independent and identically distributed standard normal distribution $\mathcal{N}(0, 1)$ and then set $Q = (Q')^\top Q'/d + I_d$ in (10), in which I_d is a $d \times d$ identity matrix. We choose the random variable ξ in (10) with a multivariate normal distribution $\mathcal{N}(100\mathbf{1}_d, 50^2 I_d)$, where $\mathbf{1}_d$ is the all-one vector in $\mathbb{R}^d$. It is easy to show that, with this choice, Assumption 6 holds and we have $\mathbb{E}\exp(\|\xi\|_2) \le \bar{\sigma}$ for some $\bar{\sigma} > 0$ so that we can choose f in Algorithms 3 and 4 to be an exponential distribution in order to guarantee Condition 4 according to Example 8. More specifically, we choose f to be the exponential

Figure 4. (Color online) The Convergence of the Average Relative Optimality Gap in the CBA When Using Different Numbers of Comparisons (S)



Notes. First row: $h_1(x,\xi)$; second row: $h_2(x,\xi)$. First column: $\xi \sim \mathcal{U}[50,150]$; second column: $\xi \sim \mathcal{N}(100,100)$.

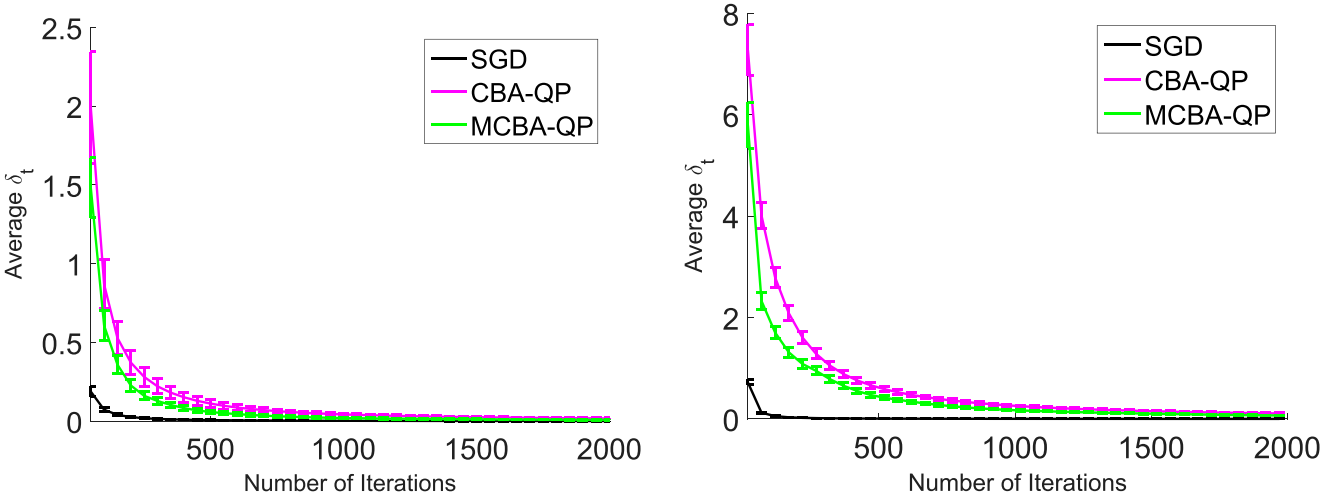
distribution (11) with $\lambda = 2^{-4} = 0.0625$ (same as λ_+ and λ_- in Section 5.1). Finally, we choose the feasible set of (10) to be the box $\mathcal{X} = \prod_1^d [50,150]$.

We compare the performance of the SGD, CBA-QP, and MCBA-QP methods. In the SGD method, each iteration computes $x_{t+1} = \mathrm{Proj}_{\mathcal{X}}(x_t - \eta_t Q(x_t - \xi_t))$, where ξ_t is sampled from the distribution of ξ and $Q(x_t - \xi_t)$ is the gradient of $h(x,\xi)$ with respect to x. As suggested by Proposition 6, we choose $\eta_t = 1/(\mu t + L)$ in the CBA-QP in which μ and L are the smallest and largest eigenvalues of Q, respectively. For the MCBA, we choose $\eta_t^k = \frac{1}{2^{k+1}\mu + L}$ and $T_k = 2^{k+3} + 4$ as suggested by Proposition 7. In the experiments, we apply the same step lengths in the SGD method as in the CBA-QP method.

We randomly generate Q in the aforementioned method, start each algorithm at the same initial point x_1 that is uniformly randomly sampled in the box $\mathcal{X}$, and run each algorithm for $T = 2{,}000$ iterations. We report the average relative optimality gap $\delta_t = \frac{H(\tilde{x}_t) - H(x^*)}{H(x^*)}$ and its standard error over 2,000 independent trails for each t. We run these experiments with the dimension $d = 5$ and $d = 20$. The results are reported in Figure 5.

From the results shown in Figure 5, we see that, even though the CBA-QP and MCBA-QP use much less information than the SGD method, they can eventually find a solution with an accuracy comparable to SGD. As shown in Table 3, the computation time for 2,000 iterations is less than one second in each algorithm, and the time is not much different across different algorithms. (This short runtime is because of the low dimensionality of the problems.) Moreover, in our tests, the MCBA-QP has a faster convergence rate than the CBA-QP, which is consistent with Propositions 6 and 7. However, the CBA-QP and MCBA-QP converge more slowly than SGD. This is because SGD utilizes the information of ξ with the full dimension although the

Figure 5. (Color online) The Convergence of the Average Relative Optimality Gap for Different Algorithms for the Multidimensional Quadratic Program (10)



Note. Left: $d = 5$; right: $d = 20$.

CBA-QP and MCBA-QP can only exploit information along a random direction u. This difference is more significant as the dimension increases, which is confirmed by Figure 5.

6. Concluding Remarks

In this paper, we considered a stochastic optimization problem when neither the underlying uncertain parameters nor the objective value at the sampled point can be observed. Instead, the decision maker can only access the comparative information between the sample point and two chosen decision points in each iteration. We proposed an algorithm that gives unbiased gradient estimates for this problem, which achieves the same asymptotic convergence rate as standard stochastic gradient methods. Numerical experiments demonstrate that our proposed algorithm is efficient.

There is one remark we would like to make. In this paper, we assumed that ξ follows a continuous distribution. However, we only need that, in each iteration, the probability that $\xi_t = x_t$ is zero. This can be guaranteed by only requiring $\mathbb{P}(\xi = \ell) = \mathbb{P}(\xi = u) = 0$, and then, in the CBA, adding a small and decaying random perturbation to $g(x_t, \xi_t, z_t)$ (for example, a uniform distribution on $[-1/2^t, 1/2^t]$ in iteration t). By doing this, one can still use the same analysis and the same performance guarantee holds for the modified algorithm.

There are several future directions for research. First, for multidimensional problems, we only considered convex quadratic problems in this paper. (As mentioned in Section 4.1, we can also generalize our method to separable objective function cases.) It would be of interest to see whether the ideas and techniques can be generalized to more general multidimensional settings. Second, in this paper, we assumed that the distribution of ξ is stationary over time and the comparative information is always reported accurately. However, in practice, the distribution of ξ may change over time or the comparative information may be reported in a noisy fashion. It would be interesting to see whether we could extend our discussions to consider such situations. Finally, our paper only considers a continuous decision setting, that is, we assumed that all the decision variables as well as the test variables can take any continuous values. In many practical situations, the decision variables may only be chosen from a finite set. It is worth further research to see whether a similar idea can be applied in such settings.

Acknowledgments

The authors thank the editor-in-chief, the associated editor, and three anonymous referees for many useful suggestions and feedback.

Endnotes

[1] Throughout the note, when $\ell = -\infty$ ($u = +\infty$), the notation $\ell \le x$ ($x \le u$) is interpreted as $x > -\infty$ ($x < +\infty$).

[2] There is a vast literature on newsvendor/inventory problems with censored demand. For some recent references, we refer the readers to Ding et al. (2002), Bensoussan et al. (2007), and Besbes and Muharremoglu (2013).

[3] There is abundant recent literature that studies the setting in which the seller can observe the full information of ξ for each customer. In particular, it has been shown that, in this case, the seller can obtain asymptotic optimal revenue as the selling horizon grows. For a review of this literature, we refer the readers to den Boer (2015).

[4] If $h(x,\xi)$ is piecewise linear with two pieces, for example, $h(x,\xi) = \mathfrak{h}\cdot(x-\xi)^+ + \mathfrak{b}\cdot(\xi-x)^+$, only comparing x and ξ may be sufficient to compute the stochastic gradient $h'_x(x,\xi)$ (that equals $\mathfrak{h}$ or $-\mathfrak{b}$).

[5] The indivisible product case with $h(x,\xi) = -x\cdot 1(\xi \geq x)$ does not satisfy Assumption 4. In particular, it does not satisfy $\mathbb{E}_\xi(h'_x(x,\xi)) = H'(x)$ (it does satisfy all the other assumptions under mild conditions though). In order to satisfy $\mathbb{E}_\xi(h'_x(x,\xi)) = H'(x)$, it is sufficient that $h(x,\xi)$ is continuous in x, which holds in the divisible product case.

[6] Note that, in Proposition 4, we only focus on the case in which $\mu > 0$ because the MCBA is mainly designed to improve the convergence rate $O(\log T/T)$ of the CBA when the problem is strongly convex. When $\mu = 0$, the convergence rate of the MCBA is still $O(1/\sqrt{T})$.

[7] In the piecewise linear case, the comparative information between x and ξ implies the knowledge of the stochastic gradient at the current sample point, and the gradient does not depend on the choice of f_- and f_+ functions.

[8] Such issues are common in variance reduction problems in simulation. The optimal choice to reduce variance relies on the knowledge of the underlying distribution. See, for example, Asmussen and Glynn (2007).

[9] Note that, in the minibatch method, the multiple samples are drawn simultaneously. It is worth noting that it is also possible to draw multiple samples sequentially, each one depending on the results of all previous ones. However, that mechanism is quite complex. Moreover, the asymptotic performance does not improve because it does not surpass the asymptotic performance when the samples can be directly observed (which is already achieved by our current algorithm with two comparisons). Therefore, we here only choose to present the minibatch approach.

[10] Note that, in Figure 1, the CBA and CBAstc sometimes perform even better than the MCBA despite the worse theoretical guarantee. In fact, this is common in convex optimization literature for such types of (restarting) methods. For example, Chen et al. (2012) proposed a stochastic gradient method called MORDA, which improves ORDA in the same paper in theoretical convergence rate using a similar restarting technique to our MCBA method. However, in the fourth column of table 2 in Chen et al. (2012), the objective value in MORDA is higher than that in ORDA. Similarly, Lan and Ghadimi (2012) developed a stochastic gradient method called multistage AC-SA, which improves AC-SA in theoretical convergence rate but not necessarily in numerical performance (see table 4.3 in Lan and Ghadimi 2012).

[11] In Appendix D in the e-companion, we also present numerical results of convergence with the minibatch method with respect to the runtime of the algorithm. The results show that choosing a small batch size can usually lead to fastest convergence (in terms of time). This is because of the trade-off between the reduced number of iterations required in the minibatch method and the increased computational efforts required in each iteration.

References

Asmussen S, Glynn PW (2007) *Stochastic Simulation: Algorithms and Analysis* (Springer, New York).
Attouch H, Bolte J, Redont P, Soubeyran A (2010) Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Lojasiewicz inequality. *Math. Oper. Res.* 35(2):438–457.
Ban G-Y, Rudin C (2018) The big data newsvendor: Practical insights from machine learning. *Oper. Res.* 67(1):90–108.
Bensoussan A, Çakanyildirim M, Sethi SP (2007) A multiperiod newsvendor problem with partially observed demand. *Math. Oper. Res.* 32(2): 322–344.
Besbes O, Muharremoglu A (2013) On implications of demand censoring in the newsvendor problem. *Management Sci.* 59(6):1407–1424.
Bolte J, Daniilidis A, Lewis A (2007) The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. *SIAM J. Optim.* 17(4):1205–1223.
Bolte J, Sabach S, Teboulle M (2014) Proximal alternating linearized minimization or nonconvex and nonsmooth problems. *Math. Programming* 146(1–2):459–494.
Cachon G, Swinney R (2011) The value of fast fashion: Quick response, enhanced design, and strategic consumer behavior. *Management Sci.* 57(4): 579–595.
Candogan O, Bimpikis K, Ozdaglar A (2012) Optimal pricing in networks with externalities. *Oper. Res.* 60(4):883–905.
Cesa-Bianchi N, Lugosi G (2006) *Prediction, Learning, and Games* (Cambridge University Press, Cambridge, UK).
Chen X, Lin Q, Pena J (2012) Optimal regularized dual averaging methods for stochastic optimization. *Adv. Neural Inform. Processing Systems* 25: 395–403.
Davis D, Drusvyatskiy D (2018) Stochastic subgradient method converges at the rate $O(k^{-1/4})$ on weakly convex functions. Working paper, Cornell University, Ithaca, NY.
den Boer AV (2015) Dynamic pricing and learning: Historical origins, current research, and new directions. *Surveys Oper. Res. Management Sci.* 20(1):1–18.
Ding X, Puterman M, Bisi A (2002) The censored newsvendor and the optimal acquisition of information. *Oper. Res.* 50(3):517–527.
Duchi J, Singer Y (2009) Efficient online and batch learning using forward-backward splitting. *J. Machine Learn. Res.* 10:2873–2898.
Duchi J, Hazan E, Singer Y (2011) Adaptive subgradient methods for online learning and stochastic optimization. *J. Machine Learn. Res.* 12: 2121–2159.
Gupta D, Denton B (2008) Appointment scheduling in healthcare: Challenges and opportunities. *IIE Trans.* 40(9):800–819.
Halman N, Orlin J, Simchi-Levi D (2012) Approximating the nonlinear newsvendor and single-item stochastic lot-sizing problems when data are given by an oracle. *Oper. Res.* 60(2):429–446.
Hazan E, Kale S (2014) Beyond the regret minimization barrier: Optimal algorithms for stochastic strongly-convex optimization. *J. Machine Learn. Res.* 15:2489–2512.
Hu C, Kwok JT, Pan W (2009) Accelerated gradient methods for stochastic optimization and online learning. *Adv. Neural Inform. Processing Systems* 22:781–789.

Karimi H, Nutini J, Schmidt M (2016) Linear convergence of gradient and proximal-gradient methods under the Polyak-Lojasiewicz condition. Frasconi P, Landwehr N, Manco G, Vreeken J, eds. *Joint Eur. Conf. Machine Learn. Knowledge Discovery Databases, Riva del Garda, Italy*, 795–811.

Kiefer J, Wolfowitz J (1952) Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.* 23(3):462–466.

Kolker A (2017) The optimal workforce staffing solutions with random patient demand in healthcare settings. Khosrow-Pour M, ed. *Encyclopedia of Information Science and Technology,* 4th ed. (IGI Global Information Science, Hershey, PA), 3711–3724.

Lan G (2012) An optimal method for stochastic composite optimization. *Math. Programming* 133(1–2):365–397.

Lan G, Ghadimi S (2012) Optimal stochastic approximation algorithms for strongly convex stochastic composite optimization, Part I: A generic algorithmic framework. *SIAM J. Optim.* 22(4):1469–1492.

Lan G, Ghadimi S (2013) Optimal stochastic approximation algorithms for strongly convex stochastic composite optimization, Part II: Shrinking procedures and optimal algorithms. *SIAM J. Optim.* 23(4):2061–2089.

Levi R, Perakis G, Uichanco J (2015) The data-driven newsvendor problem: New bounds and insights. *Oper. Res.* 63(6):1294–1306.

Li Q, Zhou Y, Liang Y, Varshney PK (2017) Convergence analysis of proximal gradient with momentum for nonconvex optimization. Working paper, Syracuse University, Syracuse, NY.

Lin Q, Chen X, Pena J (2014) A sparsity preserving stochastic gradient methods for sparse regression. *Comput. Optim. Applications* 58(2):455–482.

Nemirovski A, Juditsky A, Lan G, Shapiro A (2009) Robust stochastic approximation approach to stochastic programming. *SIAM J. Optim.* 19(4): 1574–1609.

Nemirovski A, Yudin D (1983) *Problem Complexity and Method Efficiency in Optimization* (Wiley, New York).

Noll D (2014) Convergence of non-smooth descent methods using the Kurdyka–Lojasiewicz inequality. *J. Optim. Theory Appl.* 160(2):553–572.

Rakhlin A, Shamir O, Sridharan K (2012) Making gradient descent optimal for strongly convex stochastic optimization. Langford J, Pineau J, eds. *Proc. 29th Internat. Conf. Machine Learn*. (Omnipress, Madison, WI), 449–456.

Robbins H, Monro S (1951) A stochastic approximation method. *Ann. Math. Statist.* 22(3):400–407.

Shalev-Shwartz S (2012) Online learning and online convex optimization. *Foundations Trends Machine Learn.* 4(2):107–194.

Shamir O, Zhang T (2013) Stochastic gradient descent for non-smooth optimization: Convergence results and optimal averaging schemes. *JMLR* 28:71–79.

Shapiro A, Dentcheva D, Ruszczynski A (2014) *Lectures on Stochastic Programming: Modeling and Theory* (SIAM, Philadelphia).

Widder DV (1990) *Advanced Calculus* (Dover Publication Inc., Englewood Cliffs).

Xiao L (2010) Dual averaging methods for regularized stochastic learning and online optimization. *J. Machine Learn. Res.* 11:2543–2596.

Xu Y, Yin W (2017) A globally convergent algorithm for nonconvex optimization based on block coordinate update. *J. Sci. Comput.* 72(2):700–734.

Zinkevich M (2003) Online convex programming and generalized infinitesimal gradient ascent. Fawcett T, Mishra N, eds. *Proc. 20th Internat. Conf. Machine Learn.* (AAAI Press, Menlo Park, CA), 928–936.