IISE Transactions

ISSN: 2472-5854 (Print) 2472-5862 (Online) Journal homepage: www.tandfonline.com/journals/uiie21

Simulation optimization of conditional value-at-risk

Jiaqiao Hu, Meichen Song, Michael C. Fu & Yijie Peng

To cite this article: Jiaqiao Hu, Meichen Song, Michael C. Fu & Yijie Peng (2025) Simulation optimization of conditional value-at-risk, IISE Transactions, 57:10, 1167-1181, DOI: 10.1080/24725854.2024.2429714

To link to this article: https://doi.org/10.1080/24725854.2024.2429714

IISE TRANSACTIONS
2025, VOL. 57, NO. 10, 1167–1181
https://doi.org/10.1080/24725854.2024.2429714

Simulation optimization of conditional value-at-risk

Jiaqiao Hu[a], Meichen Song[a], Michael C. Fu[b], and Yijie Peng[c]

[a]Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY, USA; [b]H. Smith School of Business & Institute for Systems Research, University of Maryland, College Park, MD, USA; [c]Guanghua School of Management, Peking University, Beijing, China

ABSTRACT

Conditional value-at-risk (CVaR) is a well-established tool for measuring risk. In this article, we consider solving CVaR optimization problems within a general simulation context. We derive an analytical expression for CVaR gradient and propose a simultaneous perturbation-type gradient estimator. This naturally results in a two-time-scale stochastic gradient algorithm for differentiable CVaR optimization. The algorithm is easily implementable and uses only three simulation evaluations at each iteration without requiring knowledge of the simulation model. We prove the almost sure local convergence of the algorithm and show that for the class of strongly convex problems, the mean absolute error of the sequence of solutions produced by the algorithm diminishes at a rate that is bounded from above by $O(k^{-2/7})$, where k is the number of iterations. Simulation experiments are also carried out to illustrate and evaluate the performance of the algorithm.

ARTICLE HISTORY
Received 7 March 2024
Accepted 3 November 2024

KEYWORDS
Simulation optimization; conditional value at risk; stochastic approximation

1. Introduction

Conditional Value-at-Risk (CVaR) (see Rockafellar and Uryasev, 2000) has become a standard technique in the financial industry for assessing and managing risk. CVaR is closely related to Value-at-Risk (VaR), which defines the quantile value quantifying the potential loss on a risky asset or portfolio at a given confidence level. CVaR, on the other hand, takes into account the risk in the tail of a return distribution and measures the average extreme losses beyond the threshold specified by VaR. In recent years, the application of CVaR has been extended to domains beyond finance. For instance, Ahmed *et al.* (2007); Gotohand and Takano (2007); Katariya *et al.* (2014) and Qiu *et al.* (2014) have examined the use of CVaR as the optimization objective to incorporate risk aversion when making inventory management decisions. In Goh and Meng (2009); Felfel *et al.* (2018); and Sawik (2019), CVaR is incorporated into optimization models to avoid excessively large costs associated with ordering, transportation, shortage, and disruption in supply chain design and management. Many other practical applications of CVaR include transportation/traffic control, network design, and renewable energy; see the recent survey article by Filippi *et al.* (2020) and references therein.

Despite its widespread applications, previous studies have rarely examined CVaR optimization from a computational point of view (Filippi *et al.*, 2020), with the majority of the existing literature focusing on applying the stochastic programming framework introduced in Rockafellar and Uryasev (2000). Specifically, let Y be a random variable and $\varphi \in (0,1)$ be a given probability level, a fundamental result given in Rockafellar and Uryasev (2000) is that CVaR can be expressed as the optimal value of a minimization problem, i.e.,

$$\mathrm{CVaR}_{\varphi}(Y) = \min_{z} \left[z + \frac{1}{1-\varphi} E[\max\{Y - z, 0\}] \right]. \tag{1}$$

This representation is convenient from an optimization standpoint, because it allows one to bypass VaR estimation and instead leverage mean-based techniques such as sample average approximation (Rockafellar and Uryasev, 2000; Wang and Ahmed, 2008) and stochastic programming (Kunz-Bay and Mayer, 2006; Schultz and Tedemann, 2006; Huang and Subramanian, 2012; Noyan, 2012) for CVaR optimization. Nevertheless, the application of the framework relies critically on structural properties (e.g., linearity or convexity) of the assumed system model, as well as knowledge of the distribution function (Sarykalin *et al.*, 2008; Filippi *et al.*, 2020). In particular, one limitation of such methods, as pointed out in Tamar *et al.* (2015), is that they cannot be suitably applied in the presence of distributional parameters, i.e., when decision variables appear as the parameters of the input distributions and affect the system/model outcomes through changing the input distributions, a scenario that frequently arises in many engineering problems such as queueing network optimization, resource allocation, and reinforcement learning.

In this article, we consider CVaR optimization under a general simulation context where there is minimal knowledge of the underlying model generating the output, and decision variables may affect the simulation model both directly and via the input distributions. Our setting assumes that the objective function is differentiable, and a primary

CONTACT Jiaqiao Hu jiaqiao.hu.1@stonybrook.edu
Supplemental data for this article can be accessed online at https://doi.org/10.1080/24725854.2024.2429714.

issue is to construct a CVaR gradient estimator that can be effectively integrated within an optimization procedure. In the simulation literature, CVaR gradient estimation is often referred to as CVaR sensitivity analysis, and has been a topic of active research over the past two decades. The early work of Scaillet (2004) considers the class of linear loss functions, for which a kernel-based CVaR sensitivity estimator is developed and analyzed in terms of its asymptotic properties. Hong and Liu (2009) later show that for general Lipschitz continuous loss functions, the CVaR sensitivity can be expressed in the form of a conditional expectation, leading to an Infinitesimal Perturbation Analysis (IPA)-type of estimator with desired properties such as consistency and asymptotic normality. When no structural parameter is contained in the loss function, Tamar *et al.* (2015) also derive an asymptotically unbiased estimator based on the Likelihood Ratio (LR) method and incorporate it into stochastic gradient descent for CVaR optimization. More recently, Glynn *et al.* (2021) propose a sensitivity estimator for general distortion risk measures that obeys a central limit theorem. The estimator uses the Generalized Likelihood Ratio (GLR) method (Peng *et al.*, 2018) to estimate distribution sensitivities, which allows the sample path to be discontinuous. It should be noted that sensitivity analysis in general is a broad area that goes far beyond just CVaR derivative estimation. Due to the vast amount of work in this area, we do not attempt to provide a description of all developments, but rather refer the interested reader to, e.g., Fu (2015) and references therein.

Although the aforementioned sensitivity estimators are valuable tools in providing CVaR gradient information at a fixed parameter value, their constructions rely on the VaR/quantile estimators, which are inherently biased for any finite sample size. Thus, a straightforward implementation of these estimators in a gradient search method would require a simulation sample size that grows with the number of algorithm iterations. In addition to the issue of bias, the selection and application of these estimation techniques could also be highly problem-dependent. For example, the LR approach of Tamar *et al.* (2015) cannot handle problems with structural parameters, whereas IPA- and GLR-based estimators rely on knowledge of the simulation model, and thus may not be applicable when the sample path derivatives of a simulation model are either unavailable or difficult to obtain.

To overcome these limitations, we draw upon ideas from recent results on quantile optimization (Hu *et al.*, 2022) and develop a two-time-scale stochastic gradient algorithm in which the CVaR gradient is approximated through a novel application of the well-known Simultaneous Perturbation (SP) method (Spall, 1992). The key observation is that the true gradient can be expressed in the form of an integral of the output distribution sensitivities. We show that this leads to a strikingly simple SP-style gradient estimator that is not only easy to implement (requiring only three simulation evaluations at each step), but can also be applied in the presence of both distributional and structural parameters. The resultant optimization algorithm we call SP-based CVaR optimization (SPCO) then consists of two coupled Stochastic Approximation (SA) recursions running at different time scales. One computes quantile estimates, and the other searches for improved solutions along the estimated descent directions of CVaR. As compared with existing approaches relying on increasing the simulation sample size to reduce estimation errors (Tamar *et al.*, 2015), SPCO simultaneously eliminates the estimation bias and noise (in both quantile and gradient estimates) by averaging all simulation data collected over the iterations, allowing evaluation (quantile estimation) and search to be conducted in a coherent manner.

We analyze the bias and variance of the proposed CVaR gradient estimator, and prove the probability one convergence of SPCO using an Ordinary Differential Equation (ODE) method from the multi-time-scale SA literature. Then, for the class of strongly convex objective functions, we further characterize the convergence rate of the algorithm by following the fixed-point argument recently introduced in Hu *et al.* (2024). Our result indicates that an optimal bound on the convergence rate of SPCO, when expressed in terms of the Mean Absolute Errors (MAEs) of the solutions produced, is of order $O(k^{-2/7})$, where k is the number of algorithm iterations. We note that although the convergence of two-time-scale SA algorithms is relatively well understood, the research is sparse on analyzing their rates of convergence. The first such result is due to Konda and Tsitsiklis (2004), who establish a central limit theorem for two-time-scale SA with linear structures. An extension of the result to the nonlinear case is presented in Mokkadem and Pelletier (2006); however, their analysis assumes that the gradient estimation errors have bounded moments and that the covariance of the estimation noise converges to a constant matrix, two critical conditions that are not satisfied by a finite-difference-based gradient estimator like ours.

In the rest of this article we start by describing the CVaR optimization problem under a simulation scenario in Section 2. In Section 3, we introduce the proposed gradient estimator and the SPCO algorithm. In Section 4, we prove the convergence of the algorithm and investigate its rate of convergence. Some numerical studies and comparison results are reported in Section 5. Finally, Section 6 concludes the article with a brief summary of the contributions.

2. Problem setting

Let Y be an output random variable from a simulation model $h(\boldsymbol{X}, \boldsymbol{\theta})$, where $\boldsymbol{X} \in \Re^l$ is a random vector generated from an input probability distribution, and $\boldsymbol{\theta} = (\theta_1, ..., \theta_d) \in \Re^d$ is a design parameter vector whose components may appear in both the input distribution and the simulation model h itself. We let $F(\cdot; \boldsymbol{\theta})$ be the cumulative distribution function (c.d.f.) of Y with an (almost everywhere) continuous probability density function (p.d.f.) $f(\cdot; \boldsymbol{\theta})$. Let $\mathcal{S}(\boldsymbol{\theta}) \subseteq \Re$ be the (closed) support of $f(\cdot; \boldsymbol{\theta})$ and assume that it can be expressed as the union of a finite number of disjoint intervals, i.e., $\mathcal{S}(\boldsymbol{\theta}) = \cup_{i=1}^{K} [a_i(\boldsymbol{\theta}), b_i(\boldsymbol{\theta})]$ with $b_i(\boldsymbol{\theta}) < a_{i+1}(\boldsymbol{\theta})$, $i = 1, ..., K-1$, where it is possible that $a_1(\boldsymbol{\theta}) = -\infty$ and/or $b_K(\boldsymbol{\theta}) = \infty$.

For a given probability level $\varphi \in (0,1)$ and a parameter vector $\boldsymbol{\theta}$, we assume that the φ-quantile of the output random variable Y, $q(\boldsymbol{\theta};\varphi)$, lies in the open interval $(a_\ell(\boldsymbol{\theta}), b_\ell(\boldsymbol{\theta}))$ for some $\ell \in \{1,...,K\}$ so that it can be uniquely written as $q(\boldsymbol{\theta};\varphi) = F^{-1}(\varphi;\boldsymbol{\theta})$. The (right-tail) CVaR at the same level φ, denoted by $\phi_Y(\boldsymbol{\theta};\varphi)$, measures the average value of Y in excess of $q(\boldsymbol{\theta};\varphi)$ and can be defined as

$$\phi_Y(\boldsymbol{\theta};\varphi) := \frac{1}{1-\varphi}\int_{q(\boldsymbol{\theta};\varphi)}^{\infty} yf(y;\boldsymbol{\theta})dy, \tag{2}$$

provided that the integral above is well-defined. The existence of CVaR is typically justified when the output distribution is light-tailed, whereas for certain types of heavy-tailed distributions, a truncation to a finite interval is often necessary to ensure the finiteness of $\phi_Y(\boldsymbol{\theta};\varphi)$. Our goal is to find an optimal parameter vector $\boldsymbol{\theta}^*$ that minimizes $\phi_Y(\boldsymbol{\theta};\varphi)$, i.e.,

$$\boldsymbol{\theta}^* = \operatorname*{argmin}_{\boldsymbol{\theta}\in\Theta} \phi_Y(\boldsymbol{\theta};\varphi), \tag{3}$$

where the feasible region Θ is a compact, convex, full-dimensional subset of $\Re^d$. We assume that Θ can be described by inequality constraints $h_j(\boldsymbol{\theta}) \le 0, j = 1,...,m$, where each $h_j(\cdot)$ is a continuously differentiable function satisfying $\nabla_{\boldsymbol{\theta}} h_j(\boldsymbol{\theta}) \neq 0$ whenever $h_j(\boldsymbol{\theta}) = 0$ (Kushner and Yin, 1997). Many commonly encountered constraint sets such as hyper-balls, hyper-rectangles, and general convex polytopes can all be written in such a form.

Note that, as is common in CVaR optimization, we have formulated (3) as a minimization problem. In this setting, the representation (1) implies that (3) is equivalent to solving a stochastic programming problem over an augmented parameter space (Rockafellar and Uryasev, 2000), i.e.,

$$\min_{\boldsymbol{\theta}\in\Theta,\, z\in\Re}\left[z + \frac{1}{1-\varphi}E[\max\{Y-z,0\}]\right], \tag{4}$$

which can be tackled using tools from traditional mean-based optimization. There are instances, however, when it is preferable to adopt a risk-seeking stance (Xia *et al.*, 2023), in which case the objective becomes CVaR maximization. For those problems, considering their stochastic programming equivalences may no longer be beneficial, because (4) would transform into a challenging max-min problem with very few tractable solutions. Our approach, in contrast, is based on a general gradient descent/ascent idea, which is unaffected by this change in problem formulation, equally accommodating both minimization and maximization objectives.

3. Simultaneous perturbation-based CVaR gradient estimator

It is well-known (Glynn *et al.*, 2021) that the CVaR $\phi_Y(\boldsymbol{\theta};\varphi)$ defined in (2) belongs to a general class of distortion risk measures that take the form

$$\rho_g(Y) = \int_0^\infty g(\bar{F}(y;\boldsymbol{\theta}))dy - \int_0^\infty \left[1 - g(\bar{F}(-y;\boldsymbol{\theta}))\right]dy,$$

where $g(\cdot):[0,1]\to[0,1]$ is called the distortion function and $\bar{F}(y;\boldsymbol{\theta}) := 1 - F(y;\boldsymbol{\theta})$ is the complementary c.d.f. Thus, if we assume the smoothness of $F(y;\boldsymbol{\theta})$ and suppose that derivative and integral can be interchanged, then using the chain rule, the gradient of $\rho_g(Y)$ can be easily obtained as

$$\nabla_{\boldsymbol{\theta}}\rho_g(Y) = -\int_{-\infty}^{\infty} \frac{dg(u)}{du}\Big|_{u=1-F(y;\boldsymbol{\theta})} \nabla_{\boldsymbol{\theta}} F(y;\boldsymbol{\theta})dy.$$

In the case of CVaR, the distortion function specializes to $g(u) = u/(1-\varphi)$ if $0 \le u \le 1-\varphi$, and $g(u) = 1$ if $1-\varphi \le u \le 1$. This immediately suggests the following form of the CVaR gradient:

$$\nabla_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta};\varphi) = -\frac{1}{1-\varphi}\int_{q(\boldsymbol{\theta};\varphi)}^{\infty} \nabla_{\boldsymbol{\theta}} F(y;\boldsymbol{\theta})dy.$$

We now collect the assumptions needed to formalize this observation.

Assumption A1: *Let $\mathcal{N}(\Theta)$ be an open neighborhood of Θ.*

(i) *The support boundary points $a_i : \mathcal{N}(\Theta) \to \Re, i = 2,...K$ and $b_i : \mathcal{N}(\Theta) \to \Re, i = 1,...,K-1$ are differentiable functions of $\boldsymbol{\theta}$. The endpoint $b_K : \mathcal{N}(\Theta) \to \Re$ is either a differentiable function of $\boldsymbol{\theta}$ or $b_K(\boldsymbol{\theta}) = \infty$ for all $\boldsymbol{\theta} \in \mathcal{N}(\Theta)$.*

(ii) *The output c.d.f. $F(\cdot;\cdot)$ is twice continuously differentiable in both arguments on $\{(y,\boldsymbol{\theta}) : y \in int(\mathcal{S}(\boldsymbol{\theta})), \boldsymbol{\theta} \in \mathcal{N}(\Theta)\}$, where $int(\mathcal{S}(\boldsymbol{\theta}))$ is the interior of $\mathcal{S}(\boldsymbol{\theta})$.*

(iii) *$\lim_{y\to\infty} y\nabla_{\boldsymbol{\theta}} F(y;\boldsymbol{\theta}) = 0$ for all $\boldsymbol{\theta} \in \Theta$.*

(iv) *For all $q < \infty$, the integrals $\int_q^r yf(y;\boldsymbol{\theta})dy$, $\int_q^r y\nabla_{\boldsymbol{\theta}} f(y;\boldsymbol{\theta})dy$, $\int_q^r \nabla_{\boldsymbol{\theta}} f(y;\boldsymbol{\theta})dy$, and $\int_q^r \nabla_{\boldsymbol{\theta}} F(y;\boldsymbol{\theta})dy$ converge uniformly on $\mathcal{N}(\Theta)$ as $r \to \infty$.*

Because Θ is compact, the differentiability of $b_K(\boldsymbol{\theta})$ in condition A1(i) implies that $b_K(\boldsymbol{\theta})$ is bounded for all $\boldsymbol{\theta} \in \Theta$, which covers output distributions supported on bounded intervals, e.g., uniform, Beta, and their mixture distributions. On the other hand, the $b_K(\boldsymbol{\theta}) = \infty$ case is satisfied by many distributions supported on (semi-)infinite intervals, such as Cauchy, exponential, and normal distributions. A1(ii) requires the output distribution to be sufficiently smooth. A1(iii) is generally satisfied by light-tailed distributions, e.g., those belonging to the exponential family, as well as general distributions with bounded support. A1(iv) holds automatically for distributions supported on bounded intervals, and can be verified for many frequently encountered light-tailed distributions. A general sufficient condition for A1(iv) to hold is the existence of integrable functions $g_1(y)$, $g_2(y)$, $g_3(y)$, and $g_4(y)$ on $[q,\infty)$ that dominate the integrands in the sense that $|yf(y;\boldsymbol{\theta})| \le g_1(y)$, $||y\nabla_{\boldsymbol{\theta}} f(y;\boldsymbol{\theta})|| \le g_2(y)$, $||\nabla_{\boldsymbol{\theta}} f(y;\boldsymbol{\theta})|| \le g_3(y)$, and $||\nabla_{\boldsymbol{\theta}} F(y;\boldsymbol{\theta})|| \le g_4(y)$ for all $\boldsymbol{\theta} \in \Theta$.

Lemma 3.1. *Suppose that A1 holds. Then $\phi_Y(\boldsymbol{\theta};\varphi)$ is differentiable with*

$$\nabla_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta};\varphi) = -\frac{1}{1-\varphi}\int_{q(\boldsymbol{\theta};\varphi)}^{\infty} \nabla_{\boldsymbol{\theta}} F(y;\boldsymbol{\theta})dy. \tag{5}$$

Proof. See Section A of the Online Appendix for a proof. ■

Because both the output distribution $F(\cdot;\boldsymbol{\theta})$ and the quantile $q(\boldsymbol{\theta};\varphi)$ are unknown, the CVaR gradient needs to be estimated from simulation data. An important observation from (5) is that the gradient operator only appears inside the integral, and the integral itself can thus be viewed as a continuous function of the integration limit $q(\boldsymbol{\theta};\varphi)$. This suggests that the CVaR gradient can be estimated by approximating an integral of the form $\int_q^\infty \nabla_{\boldsymbol{\theta}} F(y;\boldsymbol{\theta})dy$ and then replacing q by an estimate of $q(\boldsymbol{\theta};\varphi)$.

In particular, let $\boldsymbol{\theta}\in\Theta$ be fixed and q_k be an estimate of $q(\boldsymbol{\theta};\varphi)$, the quantile estimation can be carried out separately using the recursive procedure proposed in Hu *et al.* (2022):

$$q_{k+1} = q_k + \gamma_k(\varphi - I\{Y_k \le q_k\}), \tag{6}$$

where γ_k is the step size, $I\{\cdot\}$ is the indicator function, and $Y_k \sim F(\cdot;\boldsymbol{\theta})$ is an output random variable distributed according to $F(\cdot;\boldsymbol{\theta})$. Note that given q_k, $I\{Y_k \le q_k\}$ is an unbiased estimator for $F(q_k;\boldsymbol{\theta})$. Consequently, if the simulation noises $I\{Y_k \le q_k\} - F(q_k;\boldsymbol{\theta})$ can be averaged out over the iterations, then the sequence $\{q_k\}$ generated by (6) should approach the unique solution $q(\boldsymbol{\theta};\varphi)$ to the root-finding problem $F(y;\boldsymbol{\theta}) = \varphi$.

Now, given a quantile estimate q_k, consider the following simple simultaneous perturbation procedure for estimating $\nabla_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta};\varphi)$ based on simulation output information:

$$-\frac{1}{1-\varphi}\frac{-\max\{Y_k^+, q_k\} + \max\{Y_k^-, q_k\}}{2c_k\Delta_k}, \tag{7}$$

where $c_k > 0$ is the perturbation size, $\Delta_k = (\Delta_{k,1},\ldots,\Delta_{k,d})^T$ is a zero-mean random direction with i.i.d. bounded inverse second moments, $Y_k^\pm \sim F(\cdot;\boldsymbol{\theta}\pm c_k\Delta_k)$ are the output random variables generated under the perturbed parameters $\boldsymbol{\theta}\pm c_k\Delta_k$, and the division by a vector is element-wise. We justify this estimation procedure through the help of the following result.

Lemma 3.2. *Let $\boldsymbol{\theta}\in\Theta$ and q_k be fixed and suppose that A1(iv) holds. Then we have* $E\big[\max\{Y_k^\pm, q_k\}|q_k,\boldsymbol{\Delta}_k\big] = q_k + \int_{q_k}^\infty [\bar{F}(z;\boldsymbol{\theta}\pm c_k\boldsymbol{\Delta}_k)]dz$.

Proof. See Section B of the Online Appendix for a proof. ■

In view of Lemma 3.2, the expectation of the estimator (7) (conditional on q_k) can be written as

$$\begin{aligned}
&E\left[-\frac{1}{1-\varphi}\frac{-\max\{Y_k^+, q_k\} + \max\{Y_k^-, q_k\}}{2c_k\boldsymbol{\Delta}_k}\middle| q_k\right]\\
&= E\left[E\left[-\frac{1}{1-\varphi}\frac{-\max\{Y_k^+, q_k\} + \max\{Y_k^-, q_k\}}{2c_k\boldsymbol{\Delta}_k}\middle| q_k,\boldsymbol{\Delta}_k\right]\middle| q_k\right]\\
&= -\frac{1}{1-\varphi}E\left[\int_{q_k}^\infty \frac{F(z;\boldsymbol{\theta}+c_k\boldsymbol{\Delta}_k) - F(z;\boldsymbol{\theta}-c_k\boldsymbol{\Delta}_k)}{2c_k\boldsymbol{\Delta}_k}dz\middle| q_k\right].
\end{aligned} \tag{8}$$

We observe that the difference quotient in (8) is precisely the SP estimator for approximating the gradient $\nabla_{\boldsymbol{\theta}}F(z;\boldsymbol{\theta})$. In particular, if we suppose that the expectation (w.r.t. Δ_k) can be taken inside the integral, then using Taylor's theorem along with the fact $E[\Delta_{k,i}/\Delta_{k,j}] = 0$ for all $i\neq j$ and ignoring the resulting estimation bias, it can be readily seen that (8) reduces to $-\frac{1}{1-\varphi}\int_{q_k}^\infty \nabla_{\boldsymbol{\theta}}F(z;\boldsymbol{\theta})dz$, which differs from (5) only in the integration limit. As a result, when q_k approaches $q(\boldsymbol{\theta};\varphi)$, it is reasonable to expect that the CVaR gradient could be closely approximated by (7).

The above discussion suggests a natural gradient-based method for solving (3), leading to our proposed SPCO algorithm. Specifically, let $\boldsymbol{\theta}_k$ be a current estimate of $\boldsymbol{\theta}^*$, the algorithm uses two coupled recursions to search for an improved $\boldsymbol{\theta}_{k+1}$ along the descent direction of (7) while simultaneously computing a new quantile estimate as follows:

$$q_{k+1} = q_k + \gamma_k(\varphi - I\{Y_k \le q_k\}), \tag{9}$$

$$\boldsymbol{\theta}_{k+1} = \Pi_\Theta\left[\boldsymbol{\theta}_k - \beta_k\left(-\frac{1}{1-\varphi}\frac{-\max\{Y_k^+, q_k\} + \max\{Y_k^-, q_k\}}{2c_k\Delta_k}\right)\right], \tag{10}$$

where β_k and γ_k are step-size parameters, $Y_k \sim F(\cdot;\boldsymbol{\theta}_k)$, $Y_k^\pm \sim F(\cdot;\boldsymbol{\theta}_k\pm c_k\Delta_k)$, and $\Pi_\Theta(\cdot)$ stands for a projection operator that forces the iterates generated by (10) to stay within the feasible region Θ. Since SPCO is a two-time-scale SA method, its effective implementation requires the step-size β_k to be chosen very small relative to γ_k. Intuitively, this is because the convergence of the quantile estimates q_k relies on the sequence $\{Y_k\}$ being generated under a fixed parameter vector $\boldsymbol{\theta}_k$. Thus, setting β_k small has the effect of making the increments in $\boldsymbol{\theta}_k$ become progressively negligible when viewed from the q_k recursion (9), and this in turn allows proper tracking of the true quantile value $q(\boldsymbol{\theta}_k,\varphi)$ as the underlying parameter $\boldsymbol{\theta}_k$ slowly varies over time.

4. Asymptotic properties of SPCO

In this section, we establish the general local convergence behavior of SPCO and characterize its convergence rate in terms of the MAEs of $\{\boldsymbol{\theta}_k\}$. Let $(\Omega,\mathcal{F},P)$ be the probability space induced by the algorithm, where Ω is the collection of all sample trajectories generated by the algorithm, $\mathcal{F}$ is a σ-field of subsets of Ω, and P is a probability measure on $\mathcal{F}$. We also define $\mathcal{F}_k = \sigma\{q_0,\boldsymbol{\theta}_0,\ldots,q_k,\boldsymbol{\theta}_k\}$, $k = 0,1,\ldots$ as the sequence of increasing σ-fields generated by the collection of all random iterates obtained up to iteration k. Throughout the analysis, we let $\|\cdot\|$ be the usual Euclidean norm and denote

$$\boldsymbol{D}_k = -\frac{1}{1-\varphi}\frac{-\max\{Y_k^+, q_k\} + \max\{Y_k^-, q_k\}}{2c_k\Delta_k}$$

for notational convenience.

4.1. Strong convergence

The projection operation $\Pi_\Theta(\cdot)$ in (10) essentially serves as a correction step that sends an iterate back onto the region Θ whenever it becomes infeasible (Kushner and Yin, 1997). Thus, we can replace the operation with an extra correction term and put (10) in the following equivalent form:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \beta_k\left(-\frac{1}{1-\varphi}\frac{-\max\{Y_k^+, q_k\} + \max\{Y_k^-, q_k\}}{2c_k\Delta_k}\right) + \beta_k \boldsymbol{Z}_k, \tag{11}$$

where $\beta_k \boldsymbol{Z}_k := \boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k + \beta_k \boldsymbol{D}_k$ is the real vector with the smallest Euclidean norm needed to maintain the feasibility of $\boldsymbol{\theta}_{k+1}$. More precisely, because Θ is a convex set, the correction term $\boldsymbol{Z}_k$ lies in the convex cone generated by the inward normals to the surface of Θ at the point $\boldsymbol{\theta}_{k+1}$, that is, $\boldsymbol{Z}_k \in -C(\boldsymbol{\theta}_{k+1})$, where $C(\boldsymbol{\theta})$ is the normal cone to Θ at $\boldsymbol{\theta}$,

$$C(\boldsymbol{\theta}) := \{\boldsymbol{v} \in \Re^d : \boldsymbol{v}^T(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}) \le 0, \forall \tilde{\boldsymbol{\theta}} \in \Theta\}. \tag{12}$$

Clearly, from this definition, we have that $C(\boldsymbol{\theta}) = \{0\}$ for all $\boldsymbol{\theta}$ that lie in the interior of Θ.

Our convergence analysis is based on a standard ODE argument from multi-time-scale SA literature (Kushner and Yin, 1997; Borkar, 2009; Hu *et al.*, 2022) and proceeds in two major steps. First, we show that the sequence of quantile estimates $\{q_k\}$ remains bounded w.p.1. This, together with the boundedness of $\{\boldsymbol{\theta}_k\}$ (due to the projection), allows us to construct continuous-time interpolations of the iterates $\{q_k, \boldsymbol{\theta}_k\}_{k=0}^{\infty}$ and subsequently use a set of two coupled ODEs to characterize their long-run behavior. This part of the analysis is aimed at establishing the tracking ability of the q_k iterates, in the sense that $q_k \to q(\boldsymbol{\theta}_k; \varphi)$ as $k \to \infty$. Then using this result, we can write (11) in the form of a generalized SA algorithm in terms of the true CVaR gradient, two bias terms (due to the approximation error of q_k and SP estimation), a simulation noise term, and the additional projection term Z_k. Consequently, by directly applying existing results from single-time-scale SA, we are able to conclude the strong convergence of $\{\boldsymbol{\theta}_k\}$ to the limiting solution of the projected ODE

$$\dot{\boldsymbol{\theta}}(t) = -\nabla_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta}; \varphi)|_{\boldsymbol{\theta}=\boldsymbol{\theta}(t)} + \boldsymbol{z}(t), t \ge 0, \tag{13}$$

where $\boldsymbol{z}(t) \in -C(\boldsymbol{\theta}(t))$ is the minimum force needed to keep the trajectory $\boldsymbol{\theta}(t)$ in Θ.

Our analysis is conducted under the following assumptions:

Assumptions:

A2: *Let $\mathcal{N}(\Theta)$ be an open neighborhood of Θ.*

(i) *$F(\cdot;\cdot)$ is three times continuously differentiable in both arguments on $\{(y, \boldsymbol{\theta}) : y \in int(\mathcal{S}(\boldsymbol{\theta})), \boldsymbol{\theta} \in \mathcal{N}(\Theta)\}$.*

(ii) *There exists $C_1, C_2 > 0$ such that $\int_{-\infty}^{\infty} y^2 f(y; \boldsymbol{\theta})dy \le C_1$ and $\int_{-\infty}^{\infty} [\nabla^3_{\boldsymbol{\theta}} F(y; \boldsymbol{\theta})]_{i,j,k} dy \le C_2$ for all $\boldsymbol{\theta} \in \mathcal{N}(\Theta)$, where $[\nabla^3_{\boldsymbol{\theta}} F(y; \boldsymbol{\theta})]_{i,j,k}$ denotes the (i,j,k)th element of $\nabla^3_{\boldsymbol{\theta}} F(y; \boldsymbol{\theta})$,*

(iii) *There exists a constant $C_f > 0$ such that $f(y; \boldsymbol{\theta}) \le C_f$ for all $y \in \mathcal{S}$, $\boldsymbol{\theta} \in \Theta$.*

A3: *The random perturbations $\{\Delta_k\}$ are i.i.d., independent of $\mathcal{F}_k$. The components of each Δ_k are mutually independent following the symmetric Bernoulli distribution $P(\Delta_{k,i} = 1) = P(\Delta_{k,i} = -1) = 1/2$ for all $i = 1, ..., d$.*

A4: *The sequences $\{\beta_k\}$, $\{\gamma_k\}$, and $\{c_k\}$ satisfy*

(i) $\gamma_k > 0$, $\sum_{k=0}^{\infty} \gamma_k = \infty$, $\sum_{k=0}^{\infty} \gamma_k^2 < \infty$.

(ii) β_k, $c_k > 0$, $c_k \to 0$, $\sum_{k=0}^{\infty} \beta_k = \infty$, $\sum_{k=0}^{\infty} \beta_k^2/c_k^2 < \infty$, $\beta_k = o(\gamma_k)$.

Condition A2(i) is a strengthened version of A1(ii) and assumes the output distribution to be three times continuously differentiable. The condition $\int_{-\infty}^{\infty} y^2 f(y; \boldsymbol{\theta})dy \le C_1$ in A2(ii) ensures the output random variable has a bounded second-order moment under all $\boldsymbol{\theta}$, whereas the condition $\int_{-\infty}^{\infty} [\nabla^3_{\boldsymbol{\theta}} F(y; \boldsymbol{\theta})]_{i,j,k} dy \le C_2$ is needed to bound the approximation error of the gradient estimator (7). Due to A2(i), both conditions in A2(ii) hold trivially for distributions supported on finite intervals. A2(iii) requires the output density to be uniformly bounded, which is generally satisfied in many practical situations. Both A3 and A4 are regularity conditions on the algorithm input parameters. In particular, the Bernoulli random direction in A3 is a common choice when implementing SP estimators. The conditions on step- and perturbation-sizes in A4 are also standard in the SA literature (Spall, 1992; Kushner and Yin, 1997; Borkar, 2009; Zhang and Hu, 2019).

Lemma 4.1. *Let Assumptions A2(i), A2(iii), and A4(i) hold. Then the sequence $\{q_k\}$ generated by recursion (9) satisfies: (i) $\sup_k |q_k| < \infty$ w.p.1; (ii) $\sup_k E[q_k^2] < \infty$.*

Proof. See Section C of the Online Appendix. ■

The next result quantifies the estimation error of the SP estimator (7) and shows that the bias of the estimator is of order $O(c_k^2)$, whose proof is given in Section D of the Online Appendix.

Lemma 4.2. *Define*

$$\boldsymbol{b}_k := E\left[-\frac{1}{1-\varphi}\frac{-\max\{Y_k^+, q_k\} + \max\{Y_k^-, q_k\}}{2c_k\Delta_k}|\mathcal{F}_k\right] + \frac{1}{1-\varphi}\int_{q_k}^{\infty} \nabla_{\boldsymbol{\theta}} F(z; \boldsymbol{\theta}_k)dz.$$

Suppose that A1(iv), A2, and A3 hold. Then we have (i) $\boldsymbol{b}_k = O(c_k^2)$ w.p.1; (ii) $E[||\boldsymbol{b}_k||] = O(c_k^2)$.

In Lemma 4.3 below, we characterize the asymptotic behavior of $(\boldsymbol{\theta}_k, q_k)$ along the time scale defined by $\{\gamma_k\}$ and establish the convergence of the quantile estimates q_k to the true quantiles $q(\boldsymbol{\theta}_k; \varphi)$ as $k \to \infty$.

Lemma 4.3. *If Assumptions A1-A4 hold, then $|q_k - q(\boldsymbol{\theta}_k; \varphi)| \to 0$ as $k \to \infty$ w.p.1.*

Proof. See Section E of the Online Appendix for a proof. ■

The main convergence result for SPCO is stated in the following theorem.

Theorem 4.4. *If Assumptions A1-A4 hold, then the sequence $\{\boldsymbol{\theta}_k\}$ generated by SPCO converges to some limit set of the projected ODE (13) w.p.1. In addition, if $\phi_Y(\boldsymbol{\theta}; \varphi)$ is strictly*

convex on Θ, then $\{\boldsymbol{\theta}_k\}$ converges to the unique optimal solution $\boldsymbol{\theta}^*$ to the CVaR optimization problem (3) w.p.1.

Proof. Denote by

$$\tilde{\boldsymbol{b}}_k := -\frac{1}{1-\varphi}\int_{q_k}^{q(\boldsymbol{\theta}_k;\varphi)} \nabla_{\boldsymbol{\theta}} F(z;\boldsymbol{\theta}_k)dz,$$

$$\boldsymbol{\xi}_k := \boldsymbol{D}_k - E[\boldsymbol{D}_k|\mathcal{F}_k]$$

and recall that

$$\nabla_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta};\varphi)|_{\boldsymbol{\theta}=\boldsymbol{\theta}_k} = -\frac{1}{1-\varphi}\int_{q(\boldsymbol{\theta}_k;\varphi)}^{\infty} \nabla_{\boldsymbol{\theta}} F(z;\boldsymbol{\theta}_k)dz.$$

Then, under the notation of Lemma 4.3, recursion (11) can be written as

$$\begin{aligned}\boldsymbol{\theta}_{k+1} &= \boldsymbol{\theta}_k - \beta_k\left(-\frac{1}{1-\varphi}\int_{q(\boldsymbol{\theta}_k;\varphi)}^{\infty} \nabla_{\boldsymbol{\theta}} F(z;\boldsymbol{\theta}_k)dz + \tilde{\boldsymbol{b}}_k + \boldsymbol{b}_k + \boldsymbol{\xi}_k\right) + \beta_k \boldsymbol{Z}_k \\ &= \boldsymbol{\theta}_k - \beta_k\left(\nabla_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta};\varphi)|_{\boldsymbol{\theta}=\boldsymbol{\theta}_k} + \tilde{\boldsymbol{b}}_k + \boldsymbol{b}_k + \boldsymbol{\xi}_k\right) + \beta_k \boldsymbol{Z}_k.\end{aligned}$$

Since $\nabla_{\boldsymbol{\theta}} F(z;\boldsymbol{\theta})$ is continuous for all $\boldsymbol{\theta}\in\Theta$ (A1(ii)) and $|q_k - q(\boldsymbol{\theta}_k;\varphi)| \to 0$ w.p.1. (Lemma 4.3), we have $||\tilde{\boldsymbol{b}}_k|| \to 0$ as $k\to\infty$ w.p.1. In addition, from Lemma 4.2, $||\boldsymbol{b}_k|| \to 0$ w.p.1. Thus, it can be seen that all assumptions except A2.1 in Theorem 5.2.3 of Kushner and Yin (1997) are satisfied. We remark that A2.1 in Kushner and Yin (1997) is stated for expositional ease and is in fact stronger than what is needed. The assumption could instead be relaxed to the weaker condition $\sum_{k=0}^{\infty}\beta_k^2 E[||\boldsymbol{\xi}_k||^2] < \infty$, which holds true in our setting (see the proof of Lemma 4.3). It is then easy to check that all steps in the convergence proof given in Theorem 5.2.3 of Kushner and Yin (1997) still go through without requiring any modification. Hence, the convergence of $\{\boldsymbol{\theta}_k\}$ is a direct consequence of Theorem 5.2.3 in Kushner and Yin (1997). The second part of the result amounts to showing that when the objective function is strictly convex, the optimal solution $\boldsymbol{\theta}^*$ to (3) is a unique globally asymptotically stable equilibrium of ODE (13). The proof is identical to that of Corollary 1 in Hu *et al.* (2022), and we omit the details. ■

4.2. Rate of convergence

To analyze the convergence rate of the algorithm, we consider the case where the optimization problem (3) has a strictly convex objective function, with its optimal solution $\boldsymbol{\theta}^*$ lying in the interior of the constraint set Θ. Clearly, by the first-order condition for optimality, $\nabla_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta};\varphi)|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} = 0$. In addition, we have from Theorem 4.4 that $\boldsymbol{\theta}_k \to \boldsymbol{\theta}^*$ as $k\to\infty$ w.p.1.

To fix ideas, we take the step- and perturbation-sizes to be of the standard forms $\beta_k = b/k^{\beta}$, $\gamma_k = r/k^{\gamma}$, and $c_k = c/k^{\tau}$ for constants $b, r, c > 0$ and $\beta, \gamma, \tau \in (0,1)$. Let $\phi_Y(\boldsymbol{\theta};\varphi)$ be twice continuously differentiable with Hessian $H(\boldsymbol{\theta}) := \nabla^2_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta};\varphi)$. We impose the following additional assumptions on the output density and the Hessian matrix:

Assumptions:

B1: *For almost all $(q_k, \boldsymbol{\theta}_k)$ pairs, there exist constants $\epsilon > 0$ and $C_F > 0$ such that (i) $f(y;\boldsymbol{\theta}_k) \geq \epsilon$ and (ii) $||\nabla_{\boldsymbol{\theta}} F(y;\boldsymbol{\theta}_k)|| \leq C_F$ for all y on the line segment between q_k and $q(\boldsymbol{\theta}_k;\varphi)$.*

B2: *Let $\lambda(\boldsymbol{\theta})$ be the smallest eigenvalue of $H(\boldsymbol{\theta})$. There exists a constant $\varrho > 0$ such that $\lambda(\boldsymbol{\theta}) \geq \varrho$ for all $\boldsymbol{\theta}$ that lie on the line segment between $\boldsymbol{\theta}_k$ and $\boldsymbol{\theta}^*$.*

B1 essentially requires $f(\cdot;\boldsymbol{\theta}_k)$ and $\nabla_{\boldsymbol{\theta}} F(\cdot;\boldsymbol{\theta}_k)$ to be uniformly bounded on the line segment connecting q_k and $q(\boldsymbol{\theta}_k;\varphi)$. This is not a very restrictive condition in practice, and its suitability has been discussed in Hu *et al.* (2022), where it is shown that (9) can be replaced by a projected recursion that projects quantile estimates onto any closed interval H containing the true quantiles $q(\boldsymbol{\theta};\varphi)$ for all $\boldsymbol{\theta}\in\Theta$. This projection will not have an influence on the convergence behavior of $\{q_k\}$ and ensures $\{q_k\}$ to stay uniformly bounded along all sample trajectories. Consequently, B1 is guaranteed to hold when $f(\cdot;\cdot)$ and $\nabla_{\boldsymbol{\theta}} F(\cdot;\boldsymbol{\theta})$ are continuous on the compact region $H\times\Theta$. B2 is obviously satisfied when $\lambda(\boldsymbol{\theta})$ is bounded from below uniformly for all $\boldsymbol{\theta}\in\Theta$, which is the case for CVaR objectives that are strongly convex; see, e.g., Hu and Fu (2024). From this perspective, B2 can be viewed as a strengthened version of the strict convexity condition adopted in Theorem 4.4.

Lemma 4.5 below establishes the convergence rate of the quantile estimates $\{q_k\}$ in terms of their mean squared errors.

Lemma 4.5. *Let Assumptions A1-A4 and B1 hold. The sequence $\{q_k\}$ generated by SPCO satisfies*

$$\sqrt{E\left[(q_k - q(\boldsymbol{\theta}_k;\varphi))^2\right]} = O\left(\frac{\beta_k}{\gamma_k c_k}\right) + O\left(\gamma_k^{\frac{1}{2}}\right).$$

Proof. Let $\zeta_k = q_k - q(\boldsymbol{\theta}_k;\varphi)$ and write (9) in terms of ζ_k as

$$\zeta_{k+1} = \zeta_k + \gamma_k(\varphi - I\{Y_k \leq q_k\}) + q(\boldsymbol{\theta}_k;\varphi) - q(\boldsymbol{\theta}_{k+1};\varphi).$$

We square both sides of the equation to obtain

$$\begin{aligned}\zeta_{k+1}^2 = \zeta_k^2 &+ \gamma_k^2(\varphi - I\{Y_k \leq q_k\})^2 + \left(q(\boldsymbol{\theta}_k;\varphi) - q(\boldsymbol{\theta}_{k+1};\varphi)\right)^2 \\ &+ 2\zeta_k\gamma_k(\varphi - I\{Y_k \leq q_k\}) \\ &+ 2\zeta_k\left(q(\boldsymbol{\theta}_k;\varphi) - q(\boldsymbol{\theta}_{k+1};\varphi)\right) \\ &+ 2\gamma_k(\varphi - I\{Y_k \leq q_k\})\left(q(\boldsymbol{\theta}_k;\varphi) - q(\boldsymbol{\theta}_{k+1};\varphi)\right).\end{aligned} \tag{14}$$

Under A1(ii), $q(\boldsymbol{\theta};\varphi)$ is continuously differentiable on Θ and therefore is Lipschitz continuous. Let L_q be its associated Lipschitz constant. Then we have from (11) and the fact that $||\boldsymbol{Z}_k|| \leq ||\boldsymbol{D}_k||$ (Hu *et al.*, 2022, lemma 5)

$$\begin{aligned}|q(\boldsymbol{\theta}_k;\varphi) - q(\boldsymbol{\theta}_{k+1};\varphi)| &\leq L_q||\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k+1}|| \\ &\leq \beta_k L_q(||\boldsymbol{D}_k|| + ||\boldsymbol{Z}_k||) \leq 2\beta_k L_q||\boldsymbol{D}_k||,\end{aligned} \tag{15}$$

Substituting the above into (14), we get

$$\begin{aligned}\zeta_{k+1}^2 \le \zeta_k^2 + \gamma_k^2 + 4\beta_k^2 L_q^2 ||\boldsymbol{D}_k||^2 + 2\zeta_k\gamma_k(\varphi - I\{Y_k \le q_k\}) \\ +4\beta_k L_q|\zeta_k|||\boldsymbol{D}_k|| \\ + 2\gamma_k(\varphi - I\{Y_k \le q_k\})\big(q(\boldsymbol{\theta}_k;\varphi) - q(\boldsymbol{\theta}_{k+1};\varphi)\big).\end{aligned}$$

Because ζ_k is $\mathcal{F}_k$-measurable and Y_k is conditionally independent of $Y_k^{\pm}$ given $\mathcal{F}_k$, the conditional expectation $E[\zeta_{k+1}^2|\mathcal{F}_k]$ can be bounded as follows:

$$\begin{aligned}E\Big[\zeta_{k+1}^2|\mathcal{F}_k\Big] &\le \zeta_k^2 + \gamma_k^2 + 4\beta_k^2 L_q^2 E\big[||\boldsymbol{D}_k||^2|\mathcal{F}_k\big] + 2\zeta_k\gamma_k(F(q(\boldsymbol{\theta}_k;\varphi);\boldsymbol{\theta}_k) - F(q_k;\boldsymbol{\theta}_k)) \\ &\quad + 4\beta_k L_q|\zeta_k|E[||\boldsymbol{D}_k|||\mathcal{F}_k] + 2\gamma_k(F(q(\boldsymbol{\theta}_k;\varphi);\boldsymbol{\theta}_k) - F(q_k;\boldsymbol{\theta}_k))E\big[q(\boldsymbol{\theta}_k;\varphi) - q(\boldsymbol{\theta}_{k+1};\varphi)|\mathcal{F}_k\big] \\ &= \zeta_k^2 + \gamma_k^2 + 4\beta_k^2 L_q^2 E\big[||\boldsymbol{D}_k||^2|\mathcal{F}_k\big] + 2\zeta_k\gamma_k f(\bar{q}_k;\boldsymbol{\theta}_k)(q(\boldsymbol{\theta}_k;\varphi) - q_k) \\ &\quad + 4\beta_k L_q|\zeta_k|E[||\boldsymbol{D}_k|||\mathcal{F}_k] + 2\gamma_k f(\bar{q}_k;\boldsymbol{\theta}_k)(q(\boldsymbol{\theta}_k;\varphi) - q_k)E\big[q(\boldsymbol{\theta}_k;\varphi) - q(\boldsymbol{\theta}_{k+1};\varphi)|\mathcal{F}_k\big] \\ &= (1 - 2\gamma_k f(\bar{q}_k;\boldsymbol{\theta}_k))\zeta_k^2 + \gamma_k^2 + 4\beta_k^2 L_q^2 E\big[||\boldsymbol{D}_k||^2|\mathcal{F}_k\big] + 4\beta_k L_q|\zeta_k|E[||\boldsymbol{D}_k|||\mathcal{F}_k] \\ &\quad + 2\gamma_k f(\bar{q}_k;\boldsymbol{\theta}_k)|\zeta_k|E\big[|q(\boldsymbol{\theta}_k;\varphi) - q(\boldsymbol{\theta}_{k+1};\varphi)||\mathcal{F}_k\big] \\ &\le (1 - 2\gamma_k\epsilon)\zeta_k^2 + \gamma_k^2 + 4\beta_k^2 L_q^2 E\big[||\boldsymbol{D}_k||^2|\mathcal{F}_k\big] + 4\beta_k L_q(1 + \gamma_k C_f)|\zeta_k|E[||\boldsymbol{D}_k|||\mathcal{F}_k],\end{aligned}$$

where in the second step we have used a first-order Taylor series expansion of $F(q_k;\boldsymbol{\theta}_k)$ around $q(\boldsymbol{\theta}_k;\varphi)$ and $\bar{q}_k$ is on the line segment between q_k and $q(\boldsymbol{\theta}_k;\varphi)$. The last inequality follows from B1, $f(\bar{q}_k;\boldsymbol{\theta}_k) \le C_f$ (condition A2(iii)), and (15).

Without loss of generality, we take ϵ to be sufficiently small so that $2\gamma_k\epsilon < 1$ for all k. Then, by unconditioning on $\mathcal{F}_k$, it follows that

$$\begin{aligned}E\Big[\zeta_{k+1}^2\Big] &\le (1 - 2\gamma_k\epsilon)E\big[\zeta_k^2\big] + \gamma_k^2 + 4\beta_k^2 L_q^2 E\big[||\boldsymbol{D}_k||^2\big] \\ &\quad + 4\beta_k L_q(1 + \gamma_k C_f)E[|\zeta_k|E[||\boldsymbol{D}_k|||\mathcal{F}_k]] \\ &\le (1 - 2\gamma_k\epsilon)E\big[\zeta_k^2\big] + \gamma_k^2 + 4\beta_k^2 L_q^2 E\big[||\boldsymbol{D}_k||^2\big] \\ &\quad + 4\beta_k L_q(1 + \gamma_k C_f)\sqrt{E\big[\zeta_k^2\big]E\big[||\boldsymbol{D}_k||^2\big]} \\ &\le \left(\sqrt{1 - 2\gamma_k\epsilon}\sqrt{E\big[\zeta_k^2\big]} + \frac{2\beta_k L_q(1 + \gamma_k C_f)\sqrt{E\big[||\boldsymbol{D}_k||^2\big]}}{\sqrt{1 - 2\gamma_k\epsilon}}\right)^2 + \gamma_k^2 \\ &\le ((1 - \gamma_k\epsilon)\sqrt{E\big[\zeta_k^2\big]} + \mathcal{M}_k)^2 + \gamma_k^2,\end{aligned}$$

where the second inequality is due to the Cauchy–Schwarz inequality, the last step follows from $\sqrt{1-x} \le 1 - x/2$ for $x \in [0,1]$, and we have defined

$$\mathcal{M}_k := \frac{2\beta_k L_q(1 + \gamma_k C_f)\sqrt{E\big[||\boldsymbol{D}_k||^2\big]}}{\sqrt{1 - 2\gamma_k\epsilon}}.$$

From (A-11) in the Appendix, we have $E[||\boldsymbol{D}_k||^2] = O(1/c_k^2)$. Thus, there exist a constant $\mathcal{C} > 0$ and an integer $\mathcal{N} > 0$ such that $\mathcal{M}_k \le \mathcal{C}\beta_k/c_k$ for all $k \ge \mathcal{N}$.

Now consider a sequence of mappings

$$\mathcal{T}_k(x) := \sqrt{((1 - \gamma_k\epsilon)x + \mathcal{C}\beta_k/c_k)^2 + \gamma_k^2} \text{ for } x \ge 0,\ k \ge \mathcal{N}.$$

Let $\{x_k\}$ be the sequence of real numbers generated by $x_{k+1} = \mathcal{T}_k(x_k)$. Through a simple inductive argument, it can be shown that if $x_{\mathcal{N}} = \sqrt{E[\zeta_{\mathcal{N}}^2]}$, then $\sqrt{E[\zeta_k^2]} \le x_k$ for all $k = \mathcal{N}, \mathcal{N}+1, \ldots$. On the other hand, because $d\sqrt{y^2 + \gamma_k^2}/dy = y/\sqrt{y^2 + \gamma_k^2} < 1$ for all $y > 0$, the mapping $\mathcal{T}_k(\cdot)$ satisfies that $|\mathcal{T}_k(x) - \mathcal{T}_k(y)| \le (1 - \gamma_k\epsilon)|x - y|$ for all $x, y > 0$, and hence, is a contraction. The unique fixed point x_k^* of $\mathcal{T}_k$ can be obtained by solving the quadratic equation $\mathcal{T}_k(x_k^*) = x_k^*$ and its order can be derived as follows:

$$\begin{aligned}x_k^* &= \frac{\mathcal{C}(1 - \gamma_k\epsilon)\beta_k/c_k + \sqrt{\mathcal{C}^2\beta_k^2/c_k^2 + \gamma_k^3\epsilon(2 - \gamma_k\epsilon)}}{\gamma_k\epsilon(2 - \gamma_k\epsilon)} \\ &= O\left(\frac{\beta_k}{c_k\gamma_k}\right) + O\left(\gamma_k^{\frac{1}{2}}\right).\end{aligned}$$

Finally, by the contraction property of $\mathcal{T}_k(\cdot)$, the same argument as in Hu *et al.* (2024) can be used to show that $|x_k - x_k^*| = o(x_k^*)$. This in turn implies

$$\sqrt{E\big[\zeta_k^2\big]} \le x_k \le x_k^* + o(x_k^*) = O(\beta_k/(\gamma_k c_k)) + O(\gamma_k^{1/2}).$$

■

The main result of this section is presented in the following theorem, which provides a characterization of the convergence rate of the MAE $E[||\boldsymbol{\theta}_k - \boldsymbol{\theta}^*||]$.

Theorem 4.6. *If Assumptions A1-A4 and B1-B2 hold, then the sequence $\{\boldsymbol{\theta}_k\}$ generated by SPCO satisfies*

$$E[||\boldsymbol{\theta}_k - \boldsymbol{\theta}^*||] = O\left(\frac{\beta_k}{c_k\gamma_k}\right) + O\left(\gamma_k^{\frac{1}{2}}\right) + O(c_k^2) + O\left(\frac{\beta_k^{\frac{1}{2}}}{c_k}\right). \tag{16}$$

Proof. Define $\boldsymbol{\psi}_k := \boldsymbol{\theta}_k - \boldsymbol{\theta}^*$. Recall that

$$\boldsymbol{D}_k = -\frac{1}{1-\varphi}\frac{-\max\{Y_k^+, q_k\} + \max\{Y_k^-, q_k\}}{2c_k\Delta_k}$$

and

$$E[\boldsymbol{D}_k|\mathcal{F}_k] = \boldsymbol{b}_k - \frac{1}{1-\varphi}\int_{q_k}^{\infty} \nabla_{\boldsymbol{\theta}} F(z;\boldsymbol{\theta}_k)dz := \boldsymbol{b}_k + \boldsymbol{G}(q_k;\boldsymbol{\theta}_k),$$

where $\boldsymbol{b}_k$ is defined as in Lemma 4.2. From (11), we have

$$\begin{aligned} ||\boldsymbol{\psi}_{k+1}||^2 &= ||\boldsymbol{\psi}_k||^2 - 2\beta_k \boldsymbol{\psi}_k{}^T(\boldsymbol{D}_k - \boldsymbol{Z}_k) + \beta_k^2||\boldsymbol{D}_k - \boldsymbol{Z}_k||^2 \\ &\le ||\boldsymbol{\psi}_k||^2 - 2\beta_k \boldsymbol{\psi}_k{}^T\boldsymbol{D}_k + 2\beta_k \boldsymbol{\psi}_k{}^T\boldsymbol{Z}_k + 4\beta_k^2||\boldsymbol{D}_k||^2, \end{aligned} \tag{17}$$

where the inequality follows from $||\boldsymbol{Z}_k - \boldsymbol{D}_k|| \le ||\boldsymbol{Z}_k|| + ||\boldsymbol{D}_k|| \le 2||\boldsymbol{D}_k||$. Because $\boldsymbol{Z}_k \in -C(\boldsymbol{\theta}_{k+1})$, we have by (12) that $-\boldsymbol{Z}_k^T(\boldsymbol{\theta}^* - \boldsymbol{\theta}_{k+1}) \le 0$, implying $\boldsymbol{Z}_k^T(\boldsymbol{\theta}_k - \boldsymbol{\theta}^*) \le \boldsymbol{Z}_k^T(\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k+1})$. This allows us to bound the $2\beta_k\boldsymbol{\psi}_k{}^T\boldsymbol{Z}_k$ term in Lemma 4.2 as follows:

$$\begin{aligned} &2\beta_k \boldsymbol{\psi}_k{}^T\boldsymbol{Z}_k = 2\beta_k(\boldsymbol{\theta}_k - \boldsymbol{\theta}^*)^T\boldsymbol{Z}_k \le -2\beta_k(\boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k)^T\boldsymbol{Z}_k \\ &= 2\beta_k^2\boldsymbol{D}_k^T\boldsymbol{Z}_k - 2\beta_k^2\boldsymbol{Z}_k{}^T\boldsymbol{Z}_k \le 2\beta_k^2\boldsymbol{D}_k^T\boldsymbol{Z}_k \le 2\beta_k^2||\boldsymbol{D}_k||^2. \end{aligned}$$

Substituting the above into (17) and then taking conditional expectations at both sides, we have

$$\begin{aligned} E\Big[||\boldsymbol{\psi}_{k+1}||^2|\mathcal{F}_k\Big] &\le ||\boldsymbol{\psi}_k||^2 - 2\beta_k\boldsymbol{\psi}_k{}^T E[\boldsymbol{D}_k|\mathcal{F}_k] \\ &\quad + 6\beta_k^2 E\big[||\boldsymbol{D}_k||^2|\mathcal{F}_k\big] \\ &= ||\boldsymbol{\psi}_k||^2 - 2\beta_k\boldsymbol{\psi}_k{}^T[\boldsymbol{G}(q(\boldsymbol{\theta}_k;\varphi),\boldsymbol{\theta}_k) + \boldsymbol{G}(q_k,\boldsymbol{\theta}_k) \\ &\quad - \boldsymbol{G}(q(\boldsymbol{\theta}_k;\varphi),\boldsymbol{\theta}_k) + \boldsymbol{b}_k] \\ &\quad + 6\beta_k^2 E\big[||\boldsymbol{D}_k||^2|\mathcal{F}_k\big]. \end{aligned}$$

Note that because $\nabla_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta};\varphi)|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} = 0$, through a Taylor series expansion around $\boldsymbol{\theta}^*$, we can write

$$\begin{aligned} \boldsymbol{G}(q(\boldsymbol{\theta}_k;\varphi);\boldsymbol{\theta}_k) &= \nabla_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta};\varphi)|_{\boldsymbol{\theta}=\boldsymbol{\theta}_k} \\ &= \nabla_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta};\varphi)|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} + H(\bar{\boldsymbol{\theta}})(\boldsymbol{\theta}_k - \boldsymbol{\theta}^*) \\ &= H(\bar{\boldsymbol{\theta}})\boldsymbol{\psi}_k, \end{aligned}$$

where $\bar{\boldsymbol{\theta}}$ is on the line segment between $\boldsymbol{\theta}_k$ and $\boldsymbol{\theta}^*$. Moreover, by B1(ii) and the definition of $\boldsymbol{G}(q,\boldsymbol{\theta}_k)$, we have

$$\begin{aligned} ||\boldsymbol{G}(q_k;\boldsymbol{\theta}_k) - \boldsymbol{G}(q(\boldsymbol{\theta}_k;\varphi);\boldsymbol{\theta}_k)|| &\le \frac{C_F}{1-\varphi}|q_k - q(\boldsymbol{\theta}_k;\varphi)| \\ &= L_G|\zeta_k|, \end{aligned}$$

where $L_G := C_F/(1-\varphi)$ and recall that $\zeta_k = q_k - q(\boldsymbol{\theta}_k;\varphi)$. Consequently, by the Cauchy–Schwarz inequality,

$$\begin{aligned} E\Big[||\boldsymbol{\psi}_{k+1}||^2|\mathcal{F}_k\Big] &\le ||\boldsymbol{\psi}_k||^2 - 2\beta_k\boldsymbol{\psi}_k{}^T[H(\bar{\boldsymbol{\theta}})\boldsymbol{\psi}_k + \boldsymbol{G}(q_k,\boldsymbol{\theta}_k) \\ &\quad - \boldsymbol{G}(q(\boldsymbol{\theta}_k;\varphi),\boldsymbol{\theta}_k) + \boldsymbol{b}_k] \\ &\quad + 6\beta_k^2 E\big[||\boldsymbol{D}_k||^2|\mathcal{F}_k\big] \\ &\le (1 - 2\beta_k\varrho)||\boldsymbol{\psi}_k||^2 + 2\beta_k||\boldsymbol{\psi}_k||(L_G|\zeta_k| \\ &\quad + ||\boldsymbol{b}_k||) + 6\beta_k^2 E\big[||\boldsymbol{D}_k||^2|\mathcal{F}_k\big], \end{aligned}$$

where in the last step we have used B2 and the Rayleigh-Ritz inequality (Rugh, 1996), i.e., $\boldsymbol{\psi}_k^T H(\bar{\boldsymbol{\theta}})\boldsymbol{\psi}_k \ge \lambda(\bar{\boldsymbol{\theta}})||\boldsymbol{\psi}_k||^2 \ge \varrho||\boldsymbol{\psi}_k||^2$. Since $\beta_k \to 0$, there exists an integer $N > 0$ such that $0 < \beta_k\varrho \le 1/2$ for all $k \ge N$. Then, taking expectations at both sides, we have that for all $k \ge N$

$$\begin{aligned} E\Big[||\boldsymbol{\psi}_{k+1}||^2\Big] &\le (1 - 2\beta_k\varrho)E\big[||\boldsymbol{\psi}_k||^2\big] \\ &\quad + 2\beta_k\Big[L_G\sqrt{E\big[|\zeta_k|^2\big]} + \sqrt{E\big[||\boldsymbol{b}_k||^2\big]}\Big]\sqrt{E\big[||\boldsymbol{\psi}_k||^2\big]} \\ &\quad + 6\beta_k^2 E\big[||\boldsymbol{D}_k||^2\big] \\ &\le \big((1 - \beta_k\varrho)\sqrt{E\big[||\boldsymbol{\psi}_k||^2\big]} + \mathcal{B}_k\big)^2 + 6\beta_k^2 E\big[||\boldsymbol{D}_k||^2\big], \end{aligned}$$

where $\mathcal{B}_k := \beta_k\Big(L_G\sqrt{E[|\zeta_k|^2]} + \sqrt{E[||\boldsymbol{b}_k||^2]}\Big)/\sqrt{1 - 2\beta_k\varrho}$.

From the proof of Lemma 4.2(ii), it can be easily observed that $\sqrt{E[||\boldsymbol{b}_k||^2]} = O(c_k^2)$. Therefore, we have by Lemma 4.5, $\mathcal{B}_k = O(\beta_k^2/(\gamma_k c_k)) + O(\beta_k\gamma_k^{1/2}) + O(\beta_k c_k^2)$. Also note from (A-11) in the Appendix that $E[||\boldsymbol{D}_k||^2] = O(1/c_k^2)$. Let $N' \ge N$ be such that $\mathcal{B}_k \le \mathcal{C}_1(\beta_k^2/(\gamma_k c_k) + \beta_k\gamma_k^{1/2} + \beta_k c_k^2)$ and $E[||\boldsymbol{D}_k||^2] \le \mathcal{C}_2/c_k^2$ for all $k \ge N'$ and constants $\mathcal{C}_1, \mathcal{C}_2 > 0$. Next, we proceed by constructing a mapping $\mathcal{H}_k(y) = \sqrt{((1-\beta_k\varrho)y + \bar{\mathcal{B}}_k)^2 + 6\mathcal{C}_2\beta_k^2/c_k^2}$, where $\bar{\mathcal{B}}_k := \mathcal{C}_1(\beta_k^2/(\gamma_k c_k) + \beta_k\gamma_k^{1/2} + \beta_k c_k^2)$, and consider the sequence of iterates $\{y_k\}$ generated by $y_{k+1} = \mathcal{H}_k(y_k)$ with $y_{N'} = \sqrt{E[|\psi_{N'}||^2]}$. As in the proof of Lemma 4.5, it is easy to see that $\mathcal{H}_k$ is a contraction mapping with a unique fixed point y_k^* satisfying $y_k^* = \mathcal{H}_k(y_k^*)$. In addition, using induction and applying the same fixed-point argument as in the proof of Lemma 4 in Hu *et al.* (2024), we have that $\sqrt{E[||\boldsymbol{\psi}_k||^2]} \le y_k \le y_k^* + o(y_k^*)$. The order of y_k^* can be derived by inspecting the root of the fixed-point equation $y_k^* = \mathcal{H}_k(y_k^*)$, i.e.,

$$\begin{aligned} y_k^* &= \frac{(1-\beta_k\varrho)\bar{\mathcal{B}}_k + \sqrt{\bar{\mathcal{B}}_k{}^2 + 6\varrho\mathcal{C}_2(2-\beta_k\varrho)\beta_k^3/c_k^2}}{\beta_k\varrho(2-\beta_k\varrho)} \\ &= O\left(\frac{\bar{\mathcal{B}}_k}{\beta_k}\right) + O\left(\frac{\beta_k^{\frac{1}{2}}}{c_k}\right) \\ &= O\left(\frac{\beta_k}{\gamma_k c_k}\right) + O\left(\gamma_k^{\frac{1}{2}}\right) + O(c_k^2) + O\left(\frac{\beta_k^{\frac{1}{2}}}{c_k}\right). \end{aligned}$$

Hence, the desired result is proved by noting that $E[||\boldsymbol{\theta}_k - \boldsymbol{\theta}^*||] \le \sqrt{E[||\boldsymbol{\psi}_k||^2]}$. ■

From Theorem 4.6, the convergence rate of the MAE is dominated by the order of the slowest component on the right-hand side of (16). Thus, given the specific forms of β_k, γ_k, and c_k, by taking into account A4, the choice of algorithm parameters that yields the tightest convergence rate bound can be determined by solving the following optimization problem:

$$\max_{\beta,\gamma,\tau} \quad \min\{\beta-\gamma-\tau, \frac{\gamma}{2}, 2\tau, \frac{\beta}{2}-\tau\}$$
$$s.t. \begin{cases} \frac{1}{2} < \tau + \frac{1}{2} < \beta, \\ \frac{1}{2} < \gamma < \beta < 1. \end{cases}$$

The problem can be readily solved to yield $\beta \approx 1$, $\gamma = 4/7$, and $\tau = 1/7$. Under such a choice, we find that the optimal bound on the convergence rate of $E[||\boldsymbol{\theta}_k - \boldsymbol{\theta}^*||]$ is of order $O(k^{-2/7})$. An empirical illustration of this convergence rate result is provided in Section F of the Online Appendix.

5. Numerical experiments

To illustrate the algorithm, we perform some computational experiments on two sets of examples. In Section 5.1, we apply the algorithm to minimize the CVaRs of four artificially constructed output random variables under various input distributions, whereas in Section 5.2, we consider a risk-averse inventory problem, where the goal is to maximize the left-tail CVaR of the seller's revenue.

5.1. Artificial test functions

The following test functions are used in our experiments:

1. $h_1(X,\boldsymbol{\theta}) = \sum_{i=1}^d (\theta_i - i)\theta_i + X$, where $\theta_i \in [0,i]$ for $i = 1,...,d$ and $X \sim$ Cauchy(0,1).
2. $h_2(X,\boldsymbol{\theta}) = \frac{-10\exp\left(-0.2\left(\sum_{i=1}^d \theta_i^2\right)/d\right)+11}{1-\log(1-\varphi)}X$, where $\boldsymbol{\Theta} = [-2,2]^d$, and $X \sim \exp(1)$.
3. $h_3(X,\boldsymbol{\theta}) = [\sum_{i=1}^d (\theta_i - i)^2 + 1]X + \prod_{i=1}^d (\theta_i - i)$, where $\theta_i \in [i-0.5, i+0.5]$, $i = 1,...,d$ and $X \sim$ Normal(0,1).
4. $h_4(X,\boldsymbol{\theta}) = \left(\frac{\sum_{i=1}^d \theta_i^2}{d} + 0.1\right)\exp(1+X)$, where $\boldsymbol{\Theta} = [-1,1]^d$ and $X \sim$ Normal(0,1).

In function h_1, the Cauchy noise is additive, so the output distribution is also Cauchy and does not have a defined mean. Consequently, we truncate the output random variable $Y(\boldsymbol{\theta})$ to the interval $[-\sum_{i=1}^d i^2/4 - 10^4, 10^4]$ to ensure the existence of CVaR. Function h_2 scales the input random variable X by a factor that varies with the underlying parameter vector $\boldsymbol{\theta}$. In function h_3, the input random variable is both shifted and scaled. For function h_4, the output distribution is log-normal with location parameter $\log(\sum_{i=1}^d \theta_i^2/d + 0.1) + 1$, which is a heavy-tailed distribution but has a well-defined mean. For each test problem, we consider two choices of problem dimension: $d = 20$, $d = 50$ and two different quantile levels: $\varphi = 0.95$, $\varphi = 0.99$, resulting in a total of 16 test cases.

In each of the four test functions, the output Y is monotonic (element-wise) with respect to the input random variable X. Thus, it is natural to adopt the Common Random Numbers (CRN) method for reducing the variance of the gradient estimator in (10). This leads to a CRN version of the algorithm, SPCO-CRN, in which the two output random variables $Y_k^{\pm}$ in (10) are generated at each iteration using the same stream of random numbers. We also consider a version of SPCO that uses the traditional (symmetric) finite difference method in estimating CVaR gradients. The algorithm, we call finite-difference-based CVaR optimization (FDCO), requires $2d+1$ output evaluations at each step and only differs from SPCO in that the perturbation in the parameter vector $\boldsymbol{\theta}_k$ is carried out in (10) in an element-wise (as opposed to simultaneous) manner.

In addition to the two variants of SPCO, we have implemented two other optimization approaches based on direct gradient estimation: the Likelihood-Ratio CVaR optimization (LRCO) of Tamar *et al.* (2015) and a gradient descent procedure using the pathwise CVaR gradient estimator proposed in Hong and Liu (2009) (PWCO). LRCO is a single-timescale SA algorithm that takes the recursive form $\tilde{\boldsymbol{\theta}}_{k+1} = \Pi_\Theta(\tilde{\boldsymbol{\theta}}_k - \rho_k \tilde{\boldsymbol{D}}_k)$. Here, $\rho_k > 0$ is the step-size and $\tilde{\boldsymbol{D}}_k$ is the LR estimator of the CVaR gradient with its jth element given by

$$\tilde{D}_{k,j} = \frac{1}{n_k\varphi}\sum_{i=1}^{n_k}\left(\frac{\partial \log f(Y_i;\boldsymbol{\theta})}{\partial\theta_j}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_k}(Y_i - \tilde{q}(\boldsymbol{\theta}_k))I\{Y_i \geq \tilde{q}(\boldsymbol{\theta}_k)\}\right), \tag{18}$$

where $\tilde{q}(\boldsymbol{\theta})$ is the $\lceil n_k\varphi\rceil$th order statistic of an output sample $Y_1, Y_2, ...Y_{n_k} \sim F(\cdot;\boldsymbol{\theta})$. In our experiments, we set $\rho_k = 1/k$ and $n_k = \lceil(\log k)^4\rceil$, as suggested in Tamar *et al.* (2015). It is worth noting that in all our test cases, θ_i's are structural parameters that directly appear in the test functions. Thus, the application of the LR estimator (18) requires the explicit form of the output density $f(\cdot;\boldsymbol{\theta})$. PWCO, on the other hand, relies on the analytical expression for the sample path derivative $\nabla_{\boldsymbol{\theta}} h(X,\boldsymbol{\theta})$. In particular, let $Y_1, Y_2, ...Y_{n_k}$ be n_k i.i.d output evaluations under $\bar{\boldsymbol{\theta}}_k$, and $\boldsymbol{G}_1, \boldsymbol{G}_2, ...\boldsymbol{G}_{n_k}$ be observations of $\nabla_{\boldsymbol{\theta}} h(X,\boldsymbol{\theta})|_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}_k}$. Similar to LRCO, PWCO is also a recursive gradient descent procedure of the form $\bar{\boldsymbol{\theta}}_{k+1} = \Pi_\Theta(\bar{\boldsymbol{\theta}}_k - \bar{\rho}_k\bar{\boldsymbol{D}}_k)$, where the gradient estimator $\bar{\boldsymbol{D}}_k$ is given by

$$\bar{\boldsymbol{D}}_k = \frac{1}{n_k(1-\varphi)}\sum_{i=1}^{n_k}\boldsymbol{G}_i I\{Y_i > \tilde{q}(\bar{\boldsymbol{\theta}}_k)\},$$

which provides an approximation of the closed-form expression of the CVaR gradient $\nabla_{\boldsymbol{\theta}}\phi_Y(\boldsymbol{\theta};\varphi) = \frac{1}{1-\varphi}E\left[\nabla_{\boldsymbol{\theta}} h(X,\boldsymbol{\theta})I\{h(X,\boldsymbol{\theta}) > q(\boldsymbol{\theta};\varphi)\}\right]$ (Hong and Liu, 2009). In our experiments, we set $\bar{\rho}_k = 1/k$ and $n_k = \lceil(\log k)^4\rceil$, the same as in LRCO for a fair comparison.

In the implementation of SPCO, SPCO-CRN, and FDCO the decay rates of the parameters are determined from the result of Section 4.2, i.e., $\beta = 0.99$, $\gamma = 4/7$, and $\tau = 1/7$. Throughout the experiments, we take $\gamma_k = R/k^\gamma$, where R is 10% of the total number of iteration steps in SPCO and SPCO-CRN. As discussed in Hu *et al.* (2022), the large constant R allows quick tracking of the true quantiles as $\boldsymbol{\theta}_k$ changes, especially when the initial estimate q_0 is far away from their true values. The parameters β_k and c_k are chosen to be of the forms $\beta_k = b/(k+10b)^\beta$ and $c_k = c/(k+10c)^\tau$. In general, the value of b should be taken neither too large (to prevent oscillatory behavior in early iterations) nor too small (to prevent sluggish performance in the long run),

whereas the value of c should be chosen carefully to balance the bias variance tradeoff in gradient estimation. The values of the constants b and c used in SPCO, SPCO-CRN, and FDCO are listed in Table 1. These parameters are selected based on trial and error to achieve reasonable performance of the algorithms. We have also tested the performance of SPCO under different values of b and c. The results are reported in Section G of the Online Appendix, indicating some robustness of the algorithm performance with respect to their choices.

Our comparison results are based on equating the amount of computational budget, where the total number of evaluations is set to 3×10^4 for h_1 and h_3, and 3×10^5 for h_2 and h_4. In SPCO, SPCO-CRN, and FDCO, the initial quantile estimate is set to $q_0 = 0$. In all algorithms, $\boldsymbol{\theta}_0$ is uniformly generated from Θ. Each algorithm is independently repeated 40 times. The means and standard errors (over 40 runs) of the true CVaR values at the final solutions found by the five comparison algorithms are given in Tables 2 and 3, where in each case, the result that is closest to the true optimal value is shown in bold. The convergence behaviors of SPCO, SPCO-CRN, LRCO, FDCP, and PWCO are also illustrated in Figures 1 and 2, which plot the true CVaR values at the current estimated solutions as functions of the numbers of simulation evaluations.

Test results indicate good performance of SPCO in comparison with LRCO and PWCO. In particular, the algorithm outperforms LRCO and PWCO on function h_2 and most cases on function h_1 within the allowed budget. In particular, we conjecture that the relatively slow convergence of LRCO and PWCO on h_2 is due to the flat response surface of the objective function near the optimum, so that finding a good solution requires a large number of algorithm iterations. However, the increasing (per-iteration) sample size n_k required by both algorithms results in a much smaller number of algorithm iterations compared with SPCO under the given budget constraint. For h_3, SPCO shows comparable performance to PWCO and outperforms LRCO. For h_1 and h_4, the true quantiles at the optimal $\boldsymbol{\theta}^*$ values are very distant from the initial estimate $q_0 = 0$. Thus, it may take many iterations of an iterative quantile estimation procedure such as (9) to identify the correct quantile range. On the other hand, because LRCO and PWCO estimate quantiles using order statistics, they do not require specification of an initial estimate and are not heavily affected by the location of the true quantile. We see from the figures that LRCO and PWCO typically show a faster initial improvement than SPCO. Test results also show that the use of CRN may drastically reduce the variance of the gradient estimator, leading to significantly improved finite-time convergence behavior. From Tables 2

Table 1. The values of (b, c) in SPCO, SPCO-CRN and FDCO.

Test Fcn.	*1*	*2*	*3*	*4*
SPCO	(0.1, 50)	(5, 15)	(0.0001, 4)	(0.1, 10)
SPCO-CRN	(0.1, 50)	(5, 15)	(0.0001, 4)	(0.1, 100)
FDCO	(1, 500)	(20, 40)	(0.1, 5)	(0.2, 20)

Table 2. Performance on test functions for $d = 20$, based on 40 independent runs (standard error in parentheses).

	$\varphi = 0.95$			
Case	*1*	*2*	*3*	*4*
Optimum	−670.26	1.00	2.06	24.40
SPCO	−670.24 (1.99e-3)	1.02 (7.75e-4)	2.06 (3.84e-6)	25.78 (7.35e-2)
SPCO-CRN	**−670.26 (3.64e-14)**	**1.00 (4.97e-10)**	2.06 (2.31e-7)	24.40 (7.70e-6)
LRCO	−670.23 (2.10e-3)	1.79 (2.83e-2)	3.49 (5.11e-2)	24.40 (7.40e-6)
FDCO	−670.21 (1.60e-2)	1.04 (1.88e-3)	2.15 (3.75e-3)	24.78 (1.83e-2)
PWCO	−670.24 (1.20e-3)	2.16 (2.59e-2)	**2.06 (1.60e-10)**	**24.40 (1.91e-6)**
	$\varphi = 0.99$			
Optimum	−532.32	1.00	2.67	1394.75
SPCO	−532.28 (7.76e-3)	1.05 (2.11e-3)	2.67 (9.58e-8)	1534.35 (6.89)
SPCO-CRN	**-532.32 (1.82e-14)**	**1.00 (2.48e-11)**	2.67 (1.09e-7)	**1394.75 (7.28e-14)**
LRCO	−530.87 (7.69e-2)	2.11 (3.47e-2)	4.77 (8.06e-2)	1394.75 (9.48e-6)
FDCO	−530.45 (1.63)	1.09 (4.03e-3)	2.93 (1.19e-2)	1478.31 (3.75)
PWCO	−530.93 (5.91e-2)	2.64 (4.34e-2)	**2.67 (5.03e-8)**	1394.75 (1.58e-6)

Table 3. Performance on test functions for $d = 50$, based on 40 independent runs (standard error in parentheses).

	$\varphi = 0.95$			
Case	*1*	*2*	*3*	*4*
Optimum	−10683.31	1.00	2.06	24.40
SPCO	−10679.26 (6.02e-1)	1.05 (1.24e-3)	2.06 (8.01e-7)	38.18 (3.31e-1)
SPCO-CRN	−10671.45 (1.19)	**1.00 (3.21e-9)**	2.06 (4.64e-7)	24.75 (1.54e-2)
LRCO	−10682.88 (3.13e-2)	2.57 (2.78e-2)	8.01 (9.14e-2)	24.83 (2.29e-2)
FDCO	−10682.54 (3.41e-1)	1.06 (1.77e-3)	2.69 (2.71e-2)	26.11 (4.83e-2)
PWCO	**−10682.98 (1.14e-2)**	2.76 (3.31e-2)	**2.06 (6.44e-9)**	**24.75 (8.76e-3)**
	$\varphi = 0.99$			
Optimum	−10542.60	1.00	2.67	1394.75
SPCO	−10540.23 (5.40e-1)	1.14 (4.33e-3)	2.67 (3.14e-7)	1901.34 (18.32)
SPCO-CRN	**−10542.60 (2.91e-13)**	**1.00 (1.32e-12)**	2.67 (1.66e-7)	**1394.75 (2.64e-6)**
LRCO	−10520.42 (1.06)	2.78 (3.93e-2)	11.36 (1.65e-1)	1398.98 (4.50e-1)
FDCO	−10542.30 (1.34e-1)	1.13 (4.26e-3)	4.73 (7.90e-2)	1547.65 (4.69)
PWCO	−10519.85 (5.57e-1)	2.97 (2.68e-2)	**2.67 (5.37e-8)**	1398.75 (1.33e-1)

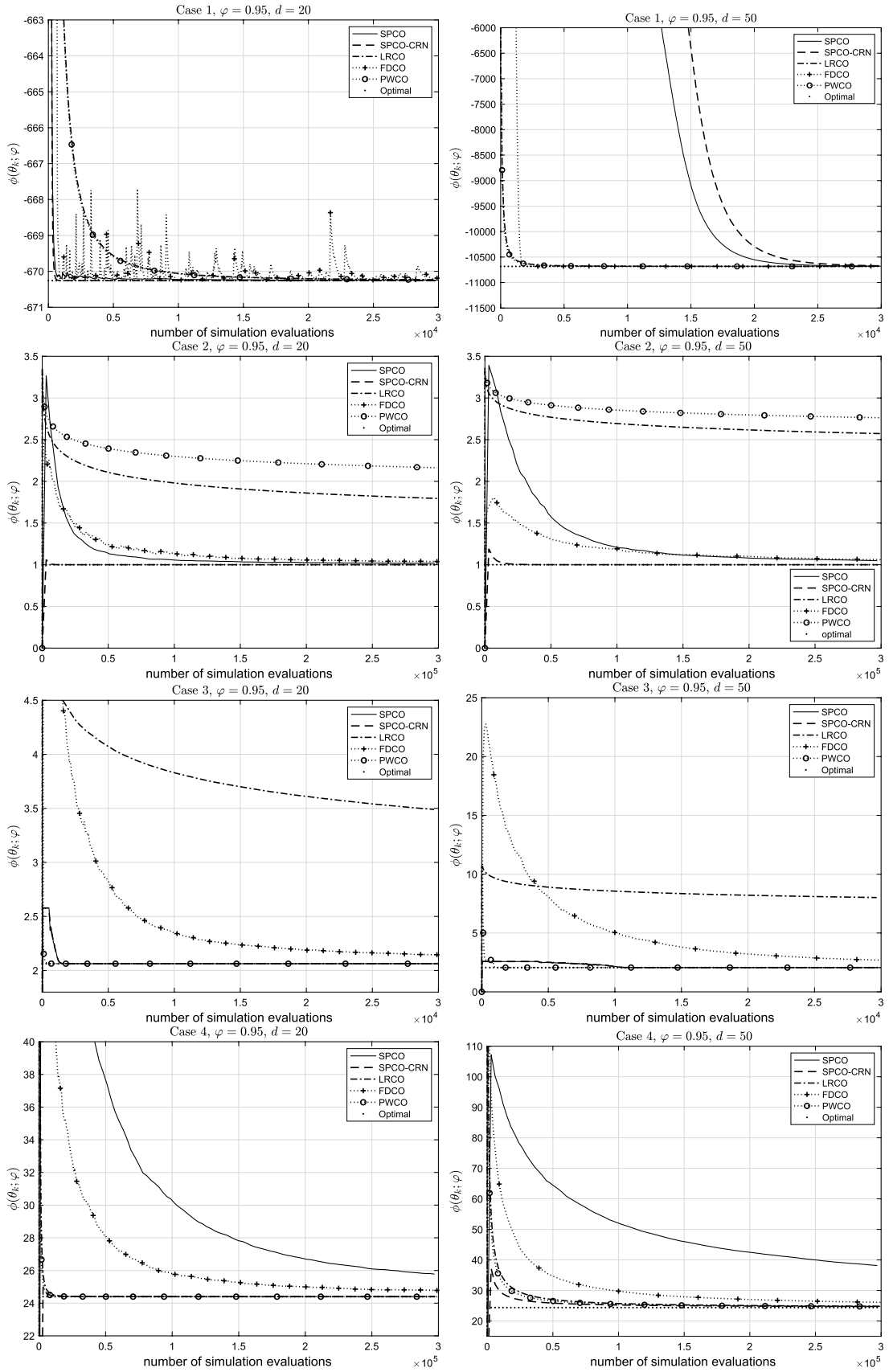

Figure 1. Performance of SPCO, SPCO-CRN, LRCO, FDCO, and PWCO on $h_1 - h_4$ ($\varphi = 0.95$).

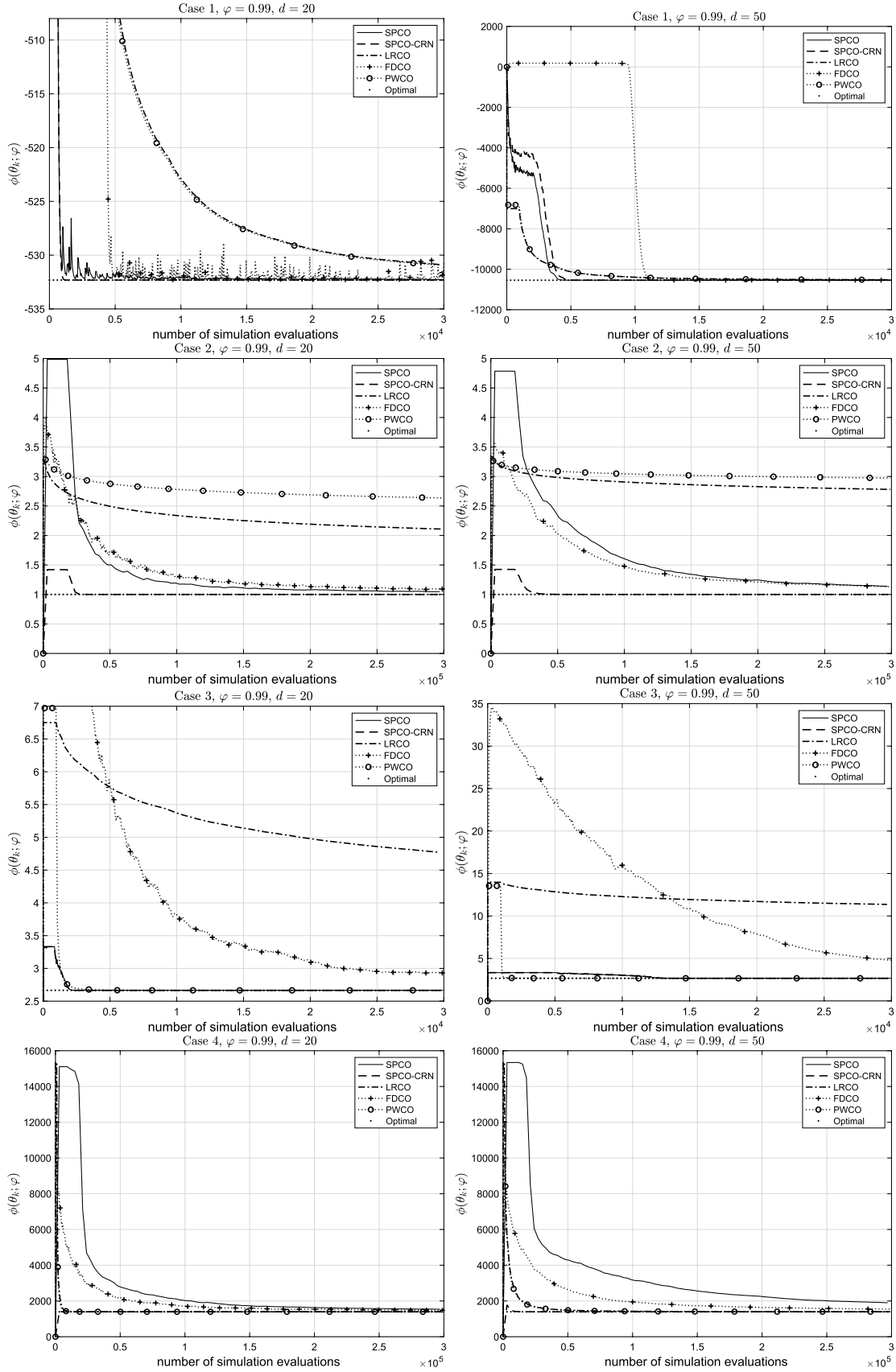

Figure 2. Performance of SPCO, SPCO-CRN, LRCO, FDCO, and PWCO on $h_1 - h_4$ ($\varphi = 0.99$).

and 3, it can be observed that SPCO-CRN yields superior performance in the majority of the 16 test cases and has significantly smaller standard errors than those of SPCO. In the remaining seven cases, the availability of a direct gradient estimator results in PWCO having the best performance of all the methods. Note that we have used a single set of parameter values for PWCO in all test cases, the performance of the algorithm could be further improved through a more careful selection of algorithm parameters tailored to each case. From the comparison results, it is clear that when structural knowledge of the underlying model is available, it should be exploited. On the other hand, because SPCO is model-free and only relies on simulation output samples, it applies more generally in black-box settings.

5.2. An inventory example

We consider a single-period risk-averse inventory problem adapted from Chen *et al.* (2009), where the goal is to determine how many items of a product should be kept in inventory, as well as the selling price of the product in order to maximize the CVaR of the seller's profit. Let θ_1 denote the inventory level at the beginning of a selling season, p the per-unit ordering cost, θ_2 the per-unit selling price, $d(\theta_2, X)$ the price-dependent demand during the selling season, where X signifies the demand risk, which is a random variable with a known distribution and is assumed to be independent of θ_2. At the end of the selling season, the leftover inventory is salvaged at price s. Then the seller's profit can be represented as

$$\begin{aligned} Y(\boldsymbol{\theta}) &= \theta_2 \min\{\theta_1, d(\theta_2, X)\} + s(\theta_1 - d(\theta_2, X))^{\dagger} - p\theta_1 \\ &= (\theta_2 - p)\theta_1 - (\theta_2 - s)(\theta_1 - d(\theta_2, X))^{\dagger}, \end{aligned} \tag{19}$$

Table 4. Optimal solutions found by SPCO on the inventory example, based on 40 independent runs (standard errors in parentheses).

	$\varphi = 0.95$		$\varphi = 0.99$	
	Optimum	*SPCO*	*Optimum*	*SPCO*
θ_1	92.73	93.88 (4.19e-1)	92.94	94.15 (3.59e-1)
θ_2	53.64	53.35 (2.03e-1)	53.53	53.25 (1.70e-1)
G_T	99.74% (2.88e − 4)		99.77% (2.56e − 4)	

where $z^{\dagger} = \max\{z, 0\}$. The objective is to determine the optimal choice of the parameter vector $\boldsymbol{\theta} = (\theta_1, \theta_2)$ that maximizes the left-tail φ-level CVaR of $Y(\boldsymbol{\theta})$ defined by

$$\phi_Y^L(\boldsymbol{\theta}) = \frac{1}{\varphi} \int_{-\infty}^{q(\boldsymbol{\theta};\varphi)} y f(y; \boldsymbol{\theta}) dy.$$

Note that because this is a problem for maximizing the left-tail CVaR, whereas SPCO is presented for minimizing the right-tail CVaR, the following modification of (10) is required in actual implementation of the algorithm:

$$\boldsymbol{\theta}_{k+1} = \Pi_{\Theta}\left[\boldsymbol{\theta}_k + \beta_k \left(-\frac{1}{\varphi} \frac{-\min\{Y_k^+, q_k\} + \min\{Y_k^-, q_k\}}{2c_k \Delta_k}\right)\right].$$

For the numerical experiments, we set $s = 1$, $p = 4$, and use a multiplicative demand model $d(\theta_2, X) = (100 - \theta_2)X$, where X follows a generalized Beta distribution with p.d.f.

$$f_X(x) = \frac{x\left(1 - \frac{x}{2}\right)}{4Beta(2,2)}, \; x \in [0, 2]$$

with *Beta* being the beta function. The feasible region $\boldsymbol{\Theta}$ is taken to be the set characterized by inequality constraints $0 \le \theta_1 \le 100$, $4 \le \theta_2 \le 100$, and $200 \le \theta_1 + 2\theta_2$. In SPCO, we take $\boldsymbol{\theta}_0 = (10, 10)$, $\beta_k = 5/(k + 50)^{0.99}$, $\gamma_k = 10^4/k^{4/7}$, $c_k = 15/(k + 150)^{1/7}$, and consider two CVaR levels: $\varphi = 0.95$ and $\varphi = 0.99$.

For the purpose of comparing with the optimum, we note that an expression for the left-tail CVaR of $Y(\boldsymbol{\theta})$ in (19) can be obtained as

$$\phi_Y^L(\boldsymbol{\theta}) = \frac{1}{\varphi} \int_a^{2lq(\boldsymbol{\theta};\varphi)} \frac{y}{4lBeta(2,2)} \frac{y-a}{l} \left(1 - \frac{y-a}{2l}\right) dy, \tag{20}$$

where $a = -3\theta_1$, $l = (\theta_2 - 1)(100 - \theta_2)$, and $q(\boldsymbol{\theta}; \varphi)$ is given by the solution to the integral equation

$$\int_a^{q(\boldsymbol{\theta};\varphi)} \frac{1}{4lBeta(2,2)} \frac{y-a}{l} \left(1 - \frac{y-a}{2l}\right) dy = \varphi.$$

Problem (20) can be numerically optimized using tools from nonlinear programming. Its (approximate) optimal solutions under the two φ-values are given in Table 4.

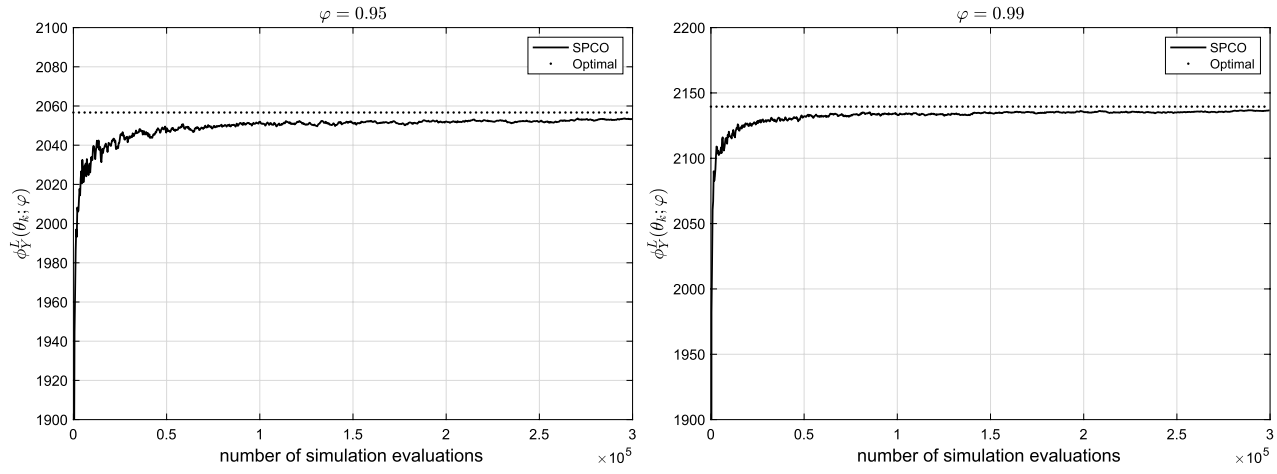

Figure 3. Performance of SPCO on the inventory example.

The performance of SPCO (averaged over 40 runs) is shown in Figure 3 and Table 4. In particular, the graphs plot the objective CVaR values $\phi_Y^L(\boldsymbol{\theta}_k;\varphi)$ as a function of the number of simulation evaluations consumed, whereas the table records the solutions returned by the algorithm upon termination in each of the two respective test cases. To further gauge the performance of the algorithm, we consider a procedure used in Huang *et al.* (2006) for measuring the relative efficacy of an optimization procedure. Specifically, let $\boldsymbol{\theta}_T$ be the solution found by SPCO upon termination, we define a performance measure

$$G_T = \frac{\phi_Y^L(\boldsymbol{\theta}_0;\varphi) - \phi_Y^L(\boldsymbol{\theta}_T;\varphi)}{\phi_Y^L(\boldsymbol{\theta}_0;\varphi) - \phi_Y^L(\boldsymbol{\theta}^*;\varphi)},$$

which signifies the reduction in the optimality gap of a returned solution relative to that of the initial values. The means and standard errors of G_T values are reported in the last row of Table 4. We see that in both test cases, SPCO is able to achieve more than 99% reduction in optimality gap.

6. Conclusions

In this article, we have considered the problem of optimizing CVaR under a general simulation environment. We have derived a closed-form formula for the CVaR gradient and proposed a simple SP-based estimator for approximating its value. The estimator, when combined with a separate procedure for estimating quantiles, results in a two-time-scale SA method for general differentiable CVaR optimization. The algorithm only requires three simulation evaluations at each step and has the advantage of allowing simulation errors to be incrementally eliminated using past simulation data accrued over the iterations. Under appropriate conditions, we have analyzed the asymptotic behavior of the algorithm in terms of its convergence and rate of convergence. The empirical performance of the algorithm has also been illustrated through simulation experiments, indicating its promising performance on a variety of testing scenarios.

Although our discussion of SPCO has primarily focused on a specific SP-style gradient estimator, the two-time-scale structure of the algorithm offers the flexibility to consider many other alternative CVaR gradient procedures. For instance, if the sample path derivatives of a simulation model can be readily obtained, then both IPA and GLR techniques could potentially be applied to lead to new algorithms with direct CVaR gradients. From this perspective, SPCO can be viewed as a framework for differentiable CVaR optimization. Thus, one line of future work will be to examine the use of other suitable gradient procedures within the framework and to investigate the theoretical performance and computational efficiency of the resultant algorithms.

Acknowlegments

The authors thank the editors and two anonymous referees for their helpful comments and suggestions that have led to a substantially improved paper.

Funding

This work was supported in part by the National Science Foundation (NSF) under Grants CMMI-1434419, CMMI-2027527, IIS-2123684, and by the Air Force Office of Scientific Research under Grant FA95502010211. The work of Y. Peng was supported in part by the National Natural Science Foundation of China (NSFC) under Grants 72325007, 72250065, and by the Key Project of Xiangjiang Laboratory (Grant 3XJ02004).

Notes on contributors

Jiaqiao Hu is a professor in the Department of Applied Mathematics and Statistics at the State University of New York, Stony Brook. He received his PhD degree in Electrical and Computer Engineering from the University of Maryland, College Park. His research interests include Markov decision processes, simulation-based optimization, stochastic modeling and analysis, and computational learning theory.

Meichen Song is a PhD student in the Department of Applied Mathematics and Statistics at the State University of New York, Stony Brook. She received a BS degree in Mathematics from Southeast University, Nanjing, China. Her research interests are simulation optimization, quantile optimization, and stochastic approximation.

Michael C. Fu holds the Smith Chair of Management Science in the Robert H. Smith School of Business, with a joint appointment in the Institute for Systems Research and affiliate faculty appointment in the Department of Electrical and Computer Engineering, all at the University of Maryland, College Park. His research interests include Markov decision processes, stochastic gradient estimation, simulation optimization, and applied probability. He is a Fellow of INFORMS and IEEE.

Yijie Peng is currently an associate professor of the Department of Management Science and Information Systems in Guanghua School of Management and is also affiliated with Institute of Artificial Intelligence, both at Peking University (PKU). He received his PhD from the Department of Management Science at Fudan University and his BS degree from the School of Mathematics at Wuhan University. His research interests include stochastic modeling and analysis, simulation optimization, machine learning, and healthcare.

ORCID

Jiaqiao Hu http://orcid.org/0000-0002-9999-672X
Michael C. Fu http://orcid.org/0000-0003-2105-4932

References

Ahmed, C., Cakmak, S.U. and Shapiro, A. (2007) Coherent risk measures in inventory problems. *European Journal of Operational Research*, **182**, 226–238.

Borkar, V.S. (2009) *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge University Press, New York, NY.

Chen, Y., Xu, M. and Zhang, Z.G. (2009) A risk-averse newsvendor model under the CVaR criterion. *Operations Research*, **57**(4), 1040–1044.

Felfel, H., Yahia, W., Ayadi, O. and Masmoudi, F. (2018) Stochastic multi-site supply chain planining in textile and apparel industry under demandand price uncertainties with risk aversion. *Annals of Operations Research*, **271**, 551–574.

Filippi, C., Guastaroba, G. and Speranza, M.G. (2020) Conditional value-at-risk beyond finance: A survey. *International Transactions in Operational Research*, **27**, 1277–1319.

Fu, M.C. (2015) Stochastic gradient estimation, in *Handbook of Simulation Optimization*, pp. 105–147. Springer, New York, NY.

Glynn, P.W., Peng, Y., Fu, M.C. and Hu, J.-Q. (2021) Computing sensitivities for distortion risk measures. *INFORMS Journal on Computing*, **33**(4), 1520–1532.

Goh, M. and Meng, F. (2009) A stochastic model for supply chain risk management using conditional value at risk, in *Managing Supply Chain Risk and Vulnerability*, pp. 141–157. Springer, Berlin, Germany.

Gotohand, J.Y. and Takano, Y. (2007) News vendor solutions via conditional value-at-risk minimization. *European Journal of Operational Research*, **179**, 80–96.

Hong, L.J. and Liu, G. (2009) Simulating sensitivities of conditional value at risk. *Management Science*, **55**(2), 281–293.

Hu, J. and Fu, M.C. (2024) Technical note: On the convergence rate of stochastic approximation for gradient-based stochastic optimization. *Operations Research*, ePub ahead of print March 8,

Hu, J., Peng, Y., Zhang, G. and Zhang, Q. (2022) A stochastic approximation method for simulation-based quantile optimization. *INFORMS Journal on Computing*, **34**(6), 2889–2907.

Hu, J., Song, M. and Fu, M. (2024) Quantile optimization via multiple timescale local search for black-box functions. *Operations Research*, ePub ahead of print March 12,

Huang, D., Allen, T.T., Notz, W.I. and Zeng, N. (2006) Global optimization of stochastic black-box systems via sequential kriging meta-models. *Journal of Global Optimization*, **34**(3), 441–466.

Huang, P. and Subramanian, D. (2012) Iterative estimation maximization for stochastic linear programs with conditional value-at-risk constraints. *Computational Management Science*, **9**, 441–458.

Katariya, A., Cetinkaya, S. and Tekin, E. (2014) On the comparison of risk-neutral and risk-averse newsvendor problems. *Journal of the Operational Research Society*, **65**, 1090–1107.

Konda, V.R. and Tsitsiklis, J.N. (2004) Convergence rate of linear two-time-scale stochastic approximation. *The Annals of Applied Probability*, **14**(2), 796–819.

Künzi-Bay, A. and Mayer, J. (2006) Computational aspects of minimizing conditional value-at-risk. *Computational Management Science*, **3**, 3–27.

Kushner, H.J. and Yin, G.G. (1997) *Stochastic Approximation Algorithms and Applications*, Volume 35 of *Applications of Mathematics*. Springer, New York, NY.

Mokkadem, A. and Pelletier, M. (2006) Convergence rate and averaging of nonlinear two-time-scale stochastic approximation algorithms. *The Annals of Applied Probability*, **16**(3), 1671–1702.

Noyan, N. (2012) Risk-averse two-stage stochastic programming with an application to disaster management. *Computers & Operations Research*, **39**, 541–559.

Peng, Y., Fu, M.C., Hu, J.-Q. and Heidergott, B. (2018) A new unbiased stochastic derivative estimator for discontinuous sample performances with structural parameters. *Operations Research*, **66**(2), 487–499.

Qiu, R., Shang, J. and Huang, X. (2014) Robust inventory decision under distribution uncertainty: A CVaR-based optimization approach. *International Journal of Production Economics*, **153**, 13–23.

Rockafellar, R.T. and Uryasev, S. (2000) Optimization of conditional value-at-risk. *Journal of Risk*, **2**, 21–42.

Rugh, W.J. (1996) *Linear System Theory* (Second ed.), Prentice Hall, Upper Saddle River NJ.

Sarykalin, S., Serraino, G. and Uryasev, S. (2008) Value-at-risk vs. conditional value-at-risk in risk management and optimization, in *Tutorials in Operations Research*, INFORMS, Hanover, Germany, pp. 270–294.

Sawik, T. (2019) Disruption mitigation and recovery in supply chains using portfolio approach. *Omega*, **84**, 232–248.

Scaillet, O. (2004) Nonparametric estimation and sensitivity analysis of expected shortfall. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, **14**(1), 115–129.

Schultz, R. and Tiedemann, S. (2006) Conditional value-at-risk in stochastic programs with mixed-integer recourse. *Mathematical Programming*, **105**, 365–386.

Spall, J.C. (1992) Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, **37**(3), 332–341.

Tamar, A., Glassner, Y. and Mannor, S. (2015) Optimizing the CVaR via sampling, in *Twenty-Ninth AAAI Conference on Artificial Intelligence*. AAAI Press, Washington, DC.

Wang, W. and Ahmed, S. (2008) Sample average approximation of expected value constrained stochastic programs. *Operations Research Letters*, **36**, 515–519.

Xia, L., Yu, Z. and Glynn, P. (2023) On the maximization of long-run reward CVaR for Markov decision processes. *arXiv preprint: arXiv: 2312.01586*.

Zhang, Q. and Hu, J. (2019) Simulation optimization using multi-time-scale adaptive random search. *Asia-Pacific Journal of Operational Research*, **36**(6), 1940014.