*Proceedings of the 2023 Winter Simulation Conference*
*C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.*

# SIMULTANEOUS PERTURBATION-BASED STOCHASTIC APPROXIMATION FOR QUANTILE OPTIMIZATION

Meichen Song
Jiaqiao Hu

Department of Applied Mathematics and Statistics
Stony Brook University
Stony Brook, NY, 11794, USA

Michael C. Fu

Robert H. Smith School of Business
University of Maryland, College Park
College Park, MD, 20742, USA

## ABSTRACT

We study a gradient-based algorithm for solving differentiable quantile optimization problems under a black-box scenario. The algorithm finds improved solutions along the descent direction of the quantile objective function, which is approximated at each step using a simultaneous perturbation technique that involves the difference quotient of the output random variables. Compared to existing quantile optimization methods, our algorithm has a two-timescale stochastic approximation structure and uses only three observations of the output random variable per iteration without requiring knowledge of the underlying system model. We show the local convergence of the algorithm and establish a finite-time bound on the convergence rate of the algorithm. Numerical results are also presented to illustrate the algorithm.

## 1 INTRODUCTION

Consider optimizing the performance of a stochastic system or model, whose output is a random variable $Y(\theta)$ with cumulative distribution function (c.d.f.) $F(\cdot;\theta)$, where $\theta \in \Theta \subset \Re^d$ is the vector of decision variables/parameters and $\Theta$ is the solution space. We allow the presence of both distributional and structural parameters, meaning that elements of $\theta$ may directly appear in the system/model itself and/or as parameters of the input distributions. In conventional stochastic optimization, the emphasis has been on optimizing mean-based performance measures, where the objective functions are formulated in the form of expectations, most commonly $E[Y(\theta)]$. Numerous approaches have been proposed for addressing such problems. We refer the reader to, e.g., Fu (2015) and references therein for a review of these different approaches.

In this work, we focus on the setting where the performance measure is given by the quantile of the output random variable, i.e., an optimization problem of the form:

$$\theta^* \in \arg\min_{\theta \in \Theta} q(\theta;\varphi), \tag{1}$$

where $\varphi \in (0,1)$ is a specified probability level and $q(\theta;\varphi)$ is the $\varphi$-quantile of $Y(\theta)$ defined by

$$q(\theta;\varphi) = \inf\{y : F(y;\theta) \geq \varphi\}.$$

In contrast to traditional mean-based objectives, which measure the average performance of a random system, a quantile function can be used to capture the tail behavior of an output distribution. Therefore, its effective optimization can be especially important in applications involving safety, reliability, or risk factors, e.g., for the purposes of mitigating the effects of extreme events and improving system resilience.

It is well known that solving the problem (1) can be very challenging (Rockafellar 2020), because unlike expectations, a quantile in general does not allow for an unbiased estimation. Prior works on this topic are the mathematical programming approaches presented in Kibzun and Kurbakovskiy (1991), Kibzun

979-8-3503-6966-3/23/$31.00  

et al. (2013), Vasiléva and Kan (2015) and the scenario optimization method of Zamar et al. (2017)—all require knowledge of the output distribution. Other methods include the recursive algorithm presented in Kim and Powell (2011) for heavy-tailed distributions; the quasi-gradient (QG) method proposed in Kibzun and Matveev (2012) for convex objective functions; the stochastic Nelder-Mead method of Chang and Lu (2014); and the Bayesian method used in, e.g., Wang et al. (2021), Sabater et al. (2021), which directly uses simulation samples in a surrogate model to approximate the response surface of an unknown quantile function. More recent developments are the multi-timescale stochastic approximation (SA) procedures introduced in Hu et al. (2022), Jiang et al. (2022), and Hu et al. (2023). The former two procedures incorporate results from quantile sensitivity estimation and can be viewed as SA methods with direct gradient estimation (through techniques such as perturbation analysis or the likelihood ratio method), in which knowledge of the system model is needed in deriving the estimators.

We assume that the quantile objective function is differentiable so that gradient-based search becomes the optimization method of choice. This smoothness condition can be expected in many engineering applications such as power system optimization (Conejo et al. 2010), queueing network optimization (Fu and Hill 1997), and traffic simulation (Spall and Chin 1997). However, central to the context of problem (1) is that neither the underlying system model nor the distribution $F$ defining $q(\theta;\varphi)$ is known, and the question is how to effectively approximate the quantile gradient by using only observations of $Y(\theta)$. Such a "black-box" setting precludes the adoption of the above-mentioned direct gradient techniques. The two pieces of work that are most closely related to ours are the QG algorithm (Kibzun and Matveev 2012) and the algorithms of Hu et al. (2023). The QG algorithm directly estimates quantile gradients via a "standard" finite-difference (FD) method using the difference quotient of the quantile estimator based on order statistics. Similar to the work of Hu et al. (2023), our proposed algorithm is, on the other hand, motivated by the analytical form of the quantile gradient. However, instead of estimating the true quantile gradient as in Hu et al. (2023), we make use of the observation that the descent direction of the objective function $q(\theta;\varphi)$ in (1) is in the negative direction of the the output c.d.f. gradient, which allows its FD approximation via a simultaneous perturbation (SP)-type estimator along the lines of Spall (1992). The resultant approximation procedure, when coupled with another recursive procedure replacing order statistics for quantile estimation, leads to a two-timescale SA algorithm we call quantile optimization via two-timescale simultaneous perturbation (QO-TSP) for optimizing (1). QO-TSP eliminates the estimation bias and noise by averaging all historical data collected over the iterations, and thus only requires three observations of the output random variable at each iteration, regardless of problem dimension. Under mild regularity conditions, we present the general local convergence property of QO-TSP. Then, for the class of strongly convex objective functions, we further establish, through a fixed-point argument introduced in Hu et al. (2023), a finite-time bound on the algorithm performance, suggesting that the optimal convergence rate of QO-TSP, when stated in terms of the mean absolute error (MAEs) of the solutions produced, is upper bounded by $O(k^{-2/7})$, where $k$ is the number of algorithm iterations.

The rest of this paper is organized as follows. In Section 2, we motivate the proposed SP gradient estimator and describe the QO-TSP algorithm. In Section 3, we present the convergence and convergence rate results of the algorithm, followed by preliminary computational experiments in Section 4 to illustrate its performance. Finally, Section 5 concludes the paper.

## 2 A SP-BASED QUANTILE OPTIMIZATION ALGORITHM

When the quantile objective function $q(\theta;\varphi)$ is differentiable, the following simple gradient-based local search algorithm can be applied for solving (1):

$$\theta_{k+1} = \theta_k - \alpha_k \widehat{\nabla} q_k(\theta_k;\varphi), \tag{2}$$

where $\alpha_k > 0$ is the step-size and $\widehat{\nabla} q_k(\cdot;\varphi)$ represents an estimator of the quantile gradient, which is the key element in defining the algorithm. In the black-box setting where only the observations of $Y(\theta)$ are available, the construction of the gradient estimator is often carried out using FD methods. In particular, a straightforward symmetric FD version of the estimator can be written as follows:

$$\widehat{\nabla}_i q_k(\theta;\varphi) = \frac{\hat{q}(\theta + c_k e_i;\varphi) - \hat{q}(\theta - c_k e_i;\varphi)}{2c_k},\ i = 1,\ldots,d, \tag{3}$$

where $\widehat{\nabla}_i$ denotes the $i$th component of the gradient estimator, $e_i$ denotes the unit vector in the $i$th direction, and $c_k > 0$ is the perturbation size satisfying $c_k \to 0$ as $k \to \infty$. The implementation of (3) requires another estimation procedure to approximate the true quantiles at the two perturbed parameter vectors $\theta \pm ce_i$. However, different from mean estimation, one challenge is that the biases involved in estimating these quantiles cannot be easily eliminated in the same way that estimation noises are averaged out over the iterations. Note that even a relatively small bias effect in the $\hat{q}$ estimates could be significantly amplified by a small perturbation size $c_k$ in the denominator, which means that the estimation procedure would need to have a strong enough tracking ability to ensure that the resultant estimates $\hat{q}$ could closely follow the true quantiles as $c_k \to 0$. For example, when order statistics are used to estimate the quantiles, i.e., the approach taken by Kibzun and Matveev (2012), it has been shown that the sample size required at each iteration will at least be growing at the rate of $k^2$ in order to make the gradient descent procedure (2) convergent.

Assume that for each $\theta \in \Theta$, the c.d.f. $F(\cdot;\theta)$ is differentiable with a probability density function (p.d.f.) $f(\cdot;\theta)$ so that the quantile $q(\theta;\varphi)$ can be equivalently written as $q(\theta;\varphi) = F^{-1}(\varphi;\theta)$. We now describe a computationally appealing approach based on the following result, which is obtained by differentiating both sides of the identity $F(q(\theta;\varphi);\theta) = \varphi$, yielding an analytical form for the quantile gradient; see, e.g., Fu et al. (2009):

$$\nabla_\theta q(\theta;\varphi) = -\frac{\nabla_\theta F(y;\theta)|_{y=q(\theta;\varphi)}}{f(q(\theta;\varphi);\theta)}. \tag{4}$$

A simple observation is that the density in the denominator of (4) is always non-negative, so that the quantile gradient shares the same direction as $-\nabla_\theta F(y;\theta)|_{y=q(\theta;\varphi)}$. This allows the descent direction of $q(\theta;\varphi)$ at a given $\theta$ to be approximated by constructing an estimator for $\nabla_\theta F(y;\theta)$ and then replacing $y$ by an estimate of $q(\theta;\varphi)$.

In particular, for a fixed $\theta \in \Theta$, the quantile $q(\theta;\varphi)$ can be estimated in a recursive way by using the procedure introduced in Hu et al. (2022):

$$q_{k+1} = q_k + \gamma_k(\varphi - I\{Y(\theta) \le q_k\}), \tag{5}$$

where $q_k$ is the current estimate of $q(\theta;\varphi)$ at step $k$, $\gamma_k > 0$ is the step-size, and $I\{\cdot\}$ denotes the indicator function. The procedure can be justified as follows. Note that given $q_k$, the conditional expectation $E[I\{Y(\theta) \le q_k\}|q_k] = F(q_k;\theta)$. Thus, the differences $I\{Y(\theta) \le q_k\} - F(q_k;\theta)$, $k = 0,1,\ldots$ basically act as martingale difference noise terms that cancel out each other over the iterations. As a result, recursion (5) can be viewed as an SA procedure for approximating the solution to the stochastic root-finding problem $F(q;\theta) = \varphi$ and hence converges to its unique solution $q(\theta;\varphi)$ as $k \to \infty$.

On the other hand, because $E[I\{Y(\theta) \le y\}] = F(y;\theta)$, a symmetric FD estimator analogous to (3) can be used for estimating the c.d.f. gradient:

$$\widehat{\nabla}_i F(y;\theta) = \frac{I\{Y(\theta + c_k e_i) \le y\} - I\{Y(\theta - c_k e_i) \le y\}}{2c_k},\ i = 1,\ldots,d. \tag{6}$$

Since the perturbation is carried out separately for each element of $\theta$, each iteration of (6) requires $2d$ samples of the output random variable. For high-dimensional problems, this could be computationally demanding. Therefore, we instead consider an SP variant of (6) by following the idea introduced in Spall (1992), in which all components of the parameter vector $\theta$ are simultaneously varied in random directions. Specifically, the symmetric SP estimator for the c.d.f. gradient is given by

$$\widetilde{\nabla} F(y;\theta) = \frac{I\{Y(\theta + c_k\Delta_k) \le y\} - I\{Y(\theta - c_k\Delta_k) \le y\}}{2c_k\Delta_k}, \tag{7}$$

where $\Delta_k = (\Delta_{k,1}, \ldots, \Delta_{k,d})^T$ is a zero-mean random direction with i.i.d. components, and the division by the vector $\Delta_k$ is component-wise. To gain some insight into this procedure, we consider (7) in its deterministic form because just like in equation (5), the estimation noise will be averaged out over the course of the iterations. Note that conditional on $\Delta_k$, the expectations of the two indicator terms in (7) are given by $F(y;\theta \pm c_k\Delta_k)$. Applying a first-order Taylor series expansion to these two functions around $\theta$ and taking the difference, it can be readily observed that

$$\frac{F(y;\theta + c_k\Delta_k) - F(y;\theta - c_k\Delta_k)}{2c_k\Delta_k} \approx \frac{\nabla_\theta F^T(y;\theta)\Delta_k}{\Delta_k}.$$

Thus, by taking expectations (w.r.t. $\Delta_k$) on both sides and invoking the fact that $E[\Delta_{k,i}/\Delta_{k,j}] = 0$ for all $i \neq j$, we have $E[\widetilde{\nabla} F(y;\theta)|\theta] \approx \nabla_\theta F(y;\theta)$. This shows that the similar gradient estimation effect of the usual FD scheme can be achieved by (7) with only two output measurements, which may result in a significant improvement in algorithm efficiency on high-dimensional problems.

Our gradient descent procedure follows the same structure as (2) but with $-\widetilde{\nabla} F(y;\theta)$ replacing $\widehat{\nabla} q_k(\theta_k;\varphi)$. This, when coupled with the recursive quantile estimation procedure (5), leads to the following algorithm we propose:

**QO-TSP**

*Step 0*: Choose initial estimates $q_0$, $\theta_0$. Specify sequences $\{\alpha_k\}$, $\{\gamma_k\}$, $\{c_k\}$, and a quantile level $\varphi$. Set the iteration counter $k = 0$.

*Step 1*: Repeat the following steps until a stopping condition is satisfied:

$$q_{k+1} = q_k + \gamma_k(\varphi - I\{Y(\theta_k) \leq q_k\}), \tag{8}$$

$$\theta_{k+1} = \Pi_\Theta\left[\theta_k - \alpha_k\left(\frac{-I\{Y(\theta_k + c_k\Delta_k) \leq q_k\} + I\{Y(\theta_k - c_k\Delta_k) \leq q_k\}}{2c_k\Delta_k}\right)\right], \tag{9}$$

$$k = k+1,$$

where in (9), $\Pi_\Theta(\cdot)$ is a projection operator that bring an iterate back onto the solution space $\Theta$ whenever it becomes infeasible. QO-TSP is a two-timescale SA algorithm and its implementation requires the step-size $\alpha_k$ to be chosen very small relative to $\gamma_k$. Roughly speaking, this is because the convergence of the quantile estimates $q_k$ relies on the sequence $\{Y(\theta_k)\}$ being generated under a fixed parameter vector $\theta_k$. Setting $\alpha_k$ small makes the increments in $\theta_k$ appear negligible. As a result, when viewed from the $q_k$ recursion (8), the parameter vector $\theta_k$ would appear to be fixed at a constant value.

## 3 CONVERGENCE PROPERTIES OF QO-TSP

Let $\mathscr{F}_k = \sigma\{q_0, \theta_0, \cdots, q_k, \theta_k\}, k = 0, 1, \cdots$, be the sequence of increasing $\sigma$-fields generated by the sequences of random iterates $\{q_k\}$ and $\{\theta_k\}$. We assume that the feasible region $\Theta$ is a convex, compact set that takes the form $\Theta = \{\theta \in \Re^d : h_j(\theta) \leq 0, j = 1, \cdots, m\}$, where $h_j(\cdot)$, $j = 1, \cdots, m$, are continuously differentiable functions satisfying $\nabla_\theta h_j(\theta) \neq 0$ whenever $h_j(\theta) = 0$; see, e.g., Kushner and Yin (1997). For a given $\theta \in \Theta$, let $\mathscr{S}$ be the support of $Y(\theta)$. We assume that $\mathscr{S}$ is unaffected by the choice of $\theta$. Denote $\|\cdot\|$ as the Euclidean norm. For notational ease, we let $I_k^{\pm} := I\{Y(\theta_k \pm c_k\Delta_k) \leq q_k\}$ and $D_k := (-I_k^+ + I_k^-)/2c_k\Delta_k$.

### 3.1 Strong Local Convergence

Using the projection $\Pi_\Theta(\cdot)$ in (9) has the same effect of using an extra correction term $Z_k$ in the recursion; see, e.g., Kushner and Yin (1997). This allows us to put (9) in the equivalent form

$$\theta_{k+1} = \theta_k - \alpha_k\Big(\frac{-I^+ + I^-}{2c_k\Delta_k}\Big) + \alpha_k Z_k, \tag{10}$$

where $\alpha_k Z_k$ is the vector with the shortest Euclidean length required to bring $\theta_k - \alpha_k D_k$ back to $\Theta$ whenever it lies outside of $\Theta$. In our setting, because $\Theta$ is a convex set, $Z_k \in -C(\theta_{k+1})$, where $C(\theta)$ is the normal cone to $\Theta$ at $\theta$ given by $C(\theta) := \{\nu \in \Re^d : \nu^T(\tilde{\theta} - \theta) \leq 0, \forall \tilde{\theta} \in \Theta\}$.

Our main convergence result is established based on the following regularity conditions:

***A1.*** *(i). Let $\mathscr{N}(\Theta)$ be an open neighborhood of $\Theta$. The c.d.f. $F(\cdot;\cdot)$ is thrice continuously differentiable in both arguments with its partial derivatives uniformly bounded on $\mathscr{S} \times \mathscr{N}(\Theta)$*

*(ii). There is a constant $C_f > 0$ such that $f(y;\theta) \leq C_f$ for all $y \in \mathscr{S}$, $\theta \in \Theta$.*

***A2.*** *The random perturbations $\{\Delta_k\}$ are i.i.d. and are independent of $\mathscr{F}_k$. The components of each $\{\Delta_k\}$ are i.i.d., follow the symmetric Bernoulli distribution $P(\Delta_{k,i} = 1) = P(\Delta_{k,i} = -1) = 1/2$ for $i = 1, \cdots, d$.*

***A3.*** *The step-size sequences $\{\alpha_k\}$, $\{\gamma_k\}$, and the perturbation size sequence $\{c_k\}$ satisfy the following conditions:*

*(i). $\alpha_k$, $c_k > 0$, $c_k \to 0$, $\sum_{k=0}^{\infty} \alpha_k = \infty$, $\sum_{k=0}^{\infty} \alpha_k^2/c_k^2 < \infty$;*

*(ii). $\gamma_k > 0$, $\sum_{k=0}^{\infty} \gamma_k = \infty$, $\sum_{k=0}^{\infty} \gamma_k^2 < \infty$, $\alpha_k = o(\gamma_k)$.*

Because the analysis of symmetric SP gradient approximation schemes such as (7) is frequently carried out using a third-order Taylor series expansion, the c.d.f. $F$ is taken to be three-times continuously differentiable; see, e.g., Spall (1992) for a similar condition in the mean-based setting. The Bernoulli random direction in A2 is commonly used when implementing SP-based gradient estimators. A3 is also standard in the SA literature; see, e.g., Spall (1992), Kushner and Yin (1997), Borkar (2008).

We have the following convergence result for QO-TSP, indicating that the asymptotic behavior of the algorithm is primarily governed by the limiting behavior of the solutions to an underlying ordinary differential equation (ODE). If the equilibria of the ODE are isolated and all lie in the interior of $\Theta$, then the algorithm will converge to a first-order stationary point of (1). In particular, when the objective function is strictly convex, the algorithm converges globally to the optimal solution.

**Theorem 1** Let Assumptions A1-A3 hold. Then w.p.1 the sequence $\{\theta_k\}$ generated by QO-TSP converges to some limit set of the projected ODE

$$\dot{\theta}(t) = \nabla_\theta F(y;\theta)|_{\theta=\theta(t), y=q(\theta(t);\varphi)} + z(t), \ t \geq 0, \tag{11}$$

where $z(t) \in -C(\theta(t))$ is the minimum force needed to keep the trajectory $\theta(t)$ in $\Theta$. If, in addition, the objective function $q(\theta;\varphi)$ is strictly convex on $\Theta$, then $\{\theta_k\}$ converges to the unique optimal solution $\theta^*$ to (1) w.p.1.

We briefly comment on the proof of Theorem 1. The rigourous analysis is lengthy but essentially follows the same ODE approach laid out in Hu et al. (2022). First, we show that the quantile estimates $\{q_k\}$ generated by (8) remains bounded w.p.1 and that the gradient estimation bias in (7) is of order $O(c_k^2)$, where the latter is established using a third-order Taylor series expansion as opposed to a first-order expansion as alluded to in Section 2. Then, the joint behavior of the recursions (8) and (9) is examined along the timescale defined by $\{\gamma_k\}$. In particular, the almost sure boundedness of $\{q_k\}$ and $\{\theta_k\}$ (due to the projection) allows us to construct path-wise continuous-time interpolations of $\{q_k, \theta_k\}$ and characterize their long-run behavior using a set of coupled limiting ODEs. This leads us to the conclusion that $q_k \to q(\theta_k;\varphi)$ as $k \to \infty$ w.p.1. This result, together with the $O(c_k^2)$ order of the gradient estimation bias, further suggests that (9) can be written in the form of a generalized Robbins–Monro SA algorithm containing the true c.d.f. gradient, two bias terms (due to the approximation errors of $q_k$ and the SP gradient estimation), a martingale difference noise term, and an additional projection term $Z_k$. Hence, the desired convergence result follows by directly applying Theorem 5.2.3 in Kushner and Yin (1997). Due to space limitation, we omit the detailed derivation/proof steps and instead refer the reader to Hu et al. (2022) for a comprehensive development of the approach.

### 3.2 Rate of Convergence

The convergence rate study is carried out under more stringent conditions. Specifically, we consider the case where $q(\theta;\varphi)$ is twice continuously differentiable on $\Theta$ and is strictly convex with its unique optimal solution $\theta^*$ lying in the interior of $\Theta$. Let the step- and perturbation-sizes to be of the standard forms $\alpha_k = a/k^\alpha$, $\gamma_k = r/k^\gamma$ and $c_k = c/k^\tau$ for $\alpha,\gamma,\tau \in (0,1)$ and $a,r,c > 0$. In addition to the conditions imposed in Section 3.1, we also make the following assumptions:

***B1.*** *For almost all $(q_k,\theta_k)$ pairs, there exists a constant $\varepsilon > 0$ such that $f(y;\theta_k) \geq \varepsilon$ $\forall y$ in the line segment between $q_k$ and $q(\theta_k;\varphi)$.*

***B2.*** *Denote $\lambda(\theta)$ as the smallest eigenvalue of the Hessian matrix $H(\theta) := \nabla^2_\theta q(\theta;\varphi)$. There exists a constant $\rho > 0$ such that $\lambda(\theta) > \rho$ for all $\theta$ in the line segment connecting $\theta_k$ and $\theta^*$.*

B1 essentially requires the output density to be bounded away from zero, and its suitability has been discussed in Hu et al. (2022). B2, in a sense, is stronger than the strictly convex condition stated in the second part of Theorem 1 and is satisfied when $q(\theta;\varphi)$ is strongly convex on $\Theta$, a condition commonly assumed in the literature for analyzing the convergence rates of gradient descent algorithms; see, e.g., Ghadimi and Lan (2012).

The next result characterizes the order of the approximation errors of the sequence of quantile estimates $\{q_k\}$. The proof of this result is based on a fixed-point approach introduced in Hu et al. (2023). The general idea is to exploit the recursive relationship between the MSEs of successive quantile estimates and then devise a sequence of iteration-dependent contraction mappings to quantify the estimation errors accrued during the search. For a full development of the approach, the reader is referred to Hu et al. (2023).

**Lemma 1** If A1-A3 and B1 hold, then the sequence $\{q_k\}$ generated by QO-TSP satisfies

$$\sqrt{E[(q_k - q(\theta_k;\varphi))^2]} = O\Big(\frac{\alpha_k}{c_k\gamma_k}\Big) + O(\gamma_k^{\frac{1}{2}}). \tag{12}$$

Lemma 1 reveals that the convergence rate of $\{q_k\}$ is not only determined by the step-size $\gamma_k$ used in (8), but also by the varying speed of underlying parameter vector $\theta_k$. Note that the ratio $\alpha_k/c_k$ in the first term on the right-hand-side of (12) essentially reflects the amount of increment in $\theta_k$ at each step (see (9)). Thus, to allow proper tracking of the true quantiles $q(\theta_k;\varphi)$ as the underlying $\theta_k$ changes over time, the step-size $\gamma_k$ should be taken large relative to the changes in $\theta_k$ so that $\alpha_k/c_k\gamma_k \to 0$ as $k \to \infty$. However, using too large a value of $\gamma_k$ would cause (12) to be dominated by the $O(\gamma_k^{1/2})$ term, slowing down the convergence. Suppose the underlying $\theta_k$ is fixed, i.e., $\alpha_k = 0$, then the first term on the right-hand-side of (12) vanishes, and the mean absolute errors (MAEs) of $\{q_k\}$ converge at the rate $O(\gamma_k^{1/2})$.

Using the result of Lemma 1, the following main convergence rate result is then obtained through a similar fixed-point argument:

**Theorem 2** If A1-A3 and B1-B2 hold, then the sequence $\{\theta_k\}$ generated by QO-TSP satisfies

$$E[\|\theta_k - \theta^*\|] = O\Big(\frac{\alpha_k}{c_k\gamma_k}\Big) + O(\gamma_k^{\frac{1}{2}}) + O(c_k^2) + O\Big(\frac{\alpha_k^{\frac{1}{2}}}{c_k}\Big). \tag{13}$$

Clearly, the convergence rate of the algorithm is dominated by the order of the slowest component on the right-hand-side of (13). Thus, the best decay rates of the algorithm parameters can be found by solving the optimization problem

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

where the first constraint above is due to A3(i) and the second constraint is imposed to ensure that A3(ii) is satisfied. This in turn yields $\alpha \approx 1$, $\gamma = 4/7$, and $\tau = 1/7$, indicating that the best converge rate of the algorithm, in terms of the MAEs of the sequence $\{\theta_k\}$ generated, is upper bounded by $O(k^{-2/7})$.

## 4 NUMERICAL EXAMPLES

To illustrate the algorithm, we consider some computational experiments on the following test functions:

1. $Y(\theta) = 5\exp\left((\theta_1-1)^2+(\theta_2-2)^2\right)X + \exp\left((\theta_1-1)(\theta_2-2)\right)$, where $\Theta = [0,2]\times[1,3]$;
2. $Y(\theta) = [\sum_{i=1}^{d}(\theta_i - i)^2 + 1]X$, where $\theta_i \in [i-1, i+1]$ and $d = 10$;
3. $Y(\theta) = \frac{\sum_{i=1}^{d}(\theta_i^3 - 5\theta_i^2 + \theta_i)}{d} + X$, where $\Theta = [1,5]^d$ and $d = 20$;
4. $Y(\theta) = \frac{\sum_{i=1}^{d}(\theta_i - i/2)^2}{d}X + \frac{\sum_{i=1}^{d}(\theta_i - i)\theta_i}{d}$, where $\theta_i \in [\frac{i}{2}-1, \frac{i}{2}+1]$ and $d = 50$,

where in each case, $X$ can be viewed as the input random variable. Note that these functions are artificially chosen to have explicit forms so that their exact optimal solutions can be obtained analytically for comparison to the numerical solutions. However, such information is not exploited in QO-TSP.

The dimensions of these test functions vary from 2 to 50. Problem 1 only contains two decision variables. Problem 2 scales the input random variable by stretching it along the horizontal direction. As the problem dimension increases, the output distribution may become very flat, making its high level quantiles difficult to estimate. Problem 3, on the other hand, shifts the input distribution without changing its shape. However, under the optimal parameter values, the output distribution quantiles are very distant from the origin, so predicting their values could also be challenging. In problem 4, the input random variable is both shifted and scaled, but to a moderate degree. For each test problem, we consider two types of noise distribution: $X \sim Cauchy(0,1)$ and $X \sim Normal(0,1)$, and two quantile levels: $\varphi = 0.6$, $\varphi = 0.95$. There are 16 test scenarios in total. Note that under the Cauchy noise, neither the mean nor the variance of the output distribution exists.

The performance of QO-TSP is compared with that of the QG algorithm of Kibzun and Matveev (2012). As described in Section 2, the QG algorithm uses the symmetric FD estimator (3) for approximating the quantile gradient, where each $\hat{q}$ in the numerator is obtained as the $\lceil n_k\varphi \rceil$th order statistic of an output sample of size $n_k$. Thus, one gradient estimate requires $2dn_k$ output observations. In our experiments, we take $n_k = \lceil k^{2.003} \rceil$. The perturbation size and the step-size in the descent recursion (2) are taken to be $1/k^{0.501}$ and $1/k$, respectively. This choice is the minimal required for the convergence of QG; see, Theorem 8 of Kibzun and Matveev (2012). In addition to QG, we have also implemented a common random number (CRN) version of QO-TSP, denoted as QO-CRN. This is conducted in the same spirit as in Kleinman et al. (1999), in which the two output random variables $Y(\theta_k \pm c_k\Delta_k)$ in (9) are generated using the same stream of random numbers with the hope to reduce the variance of the gradient estimator.

In both QO-TSP and QO-CRN, the decay rates of the step- and perturbation-sizes are determined based on the results of Section 3.2, i.e., $\alpha = 0.99$ (note that $\alpha \in (0,1)$), $\gamma = 4/7$, and $\tau = 1/7$. We take these parameters of the forms $\alpha_k = \kappa_1(2R)^\alpha/(k+R)^\alpha$, $c_k = \kappa_2(2R)^\tau/(k+R)^\tau$, and $\gamma_k = R/k^\gamma$, which resemble those suggested in, e.g. Spall (2003), where $R$ is a hyperparameter that represents $m$ percent of the maximum number of iterations allowed and $\kappa_1$, $\kappa_2$ are used to ensure non-negligible $\alpha_k$ and $c_k$ values during the first $R$ iterations. We have experimented with different values of $\kappa_1$, $\kappa_2$ and $m$ by running the algorithm on the test functions with reduced numbers of iterations. Our numerical results are based on the values listed in Table 1, which yield reasonable performance in each test scenario.

We set the total number of output evaluations to $3\times10^4$ for problem 1, and $3\times10^5$ for the rest of the test functions. For each of the respective 16 test scenarios, we perform 40 independent replication runs of all three algorithms, where each run is initialized by randomly selecting $\theta_0$ from $\Theta$. Tables 2-3 show the means and standard errors (over 40 runs) of the true quantile values at the final solutions found by the three comparison algorithms. In each column of the tables, the result that is closest to the true optimal

Table 1: Choices of $\kappa_1$, $\kappa_2$, and $m$ in QO-TSP and QO-CRN.

| Problem | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\kappa_1$ | 0.05 | 0.05 | 1 | 0.05 |
| $\kappa_2$ | 0.5 | 0.5 | 2 | 0.5 |
| $m$ | 0.1 | 0.1 | 0.1 | 1 |

value is shown in bold. The convergence behavior of the algorithms can be visualized in Figures 1 and 2, which plot the true quantile values at the current estimated solutions as functions of the numbers of output evaluations consumed. Test results indicate reasonable performance of QO-TSP. In particular, the algorithm outperforms QG within the allowed budget in the majority of the test cases. Also, in many cases, the use of CRNs may drastically reduce the variance of the gradient estimator, leading to significantly improved convergence behavior. We observe from the tables that with the exception of the 2-dimensional case, QO-CRN not only yields solutions that are closest to the optimal values but also shows smaller standard errors than those of QO-TSP and QG.

Table 2: Performance of QO-TSP, QO-CRN, and QG (Cauchy distribution case).

| Problem | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| | | $\varphi = 0.6$ | | |
| Optimum | 2.62 | 0.32 | -14.91 | -214.63 |
| QO-TSP | 2.74 (1.93e-2) | 0.34 (1.17e-3) | -14.85 (2.22e-3) | -214.60 (1.02e-3) |
| QO-CRN | 2.71 (1.39e-2) | **0.33 (2.47e-4)** | **-14.88 (1.78e-4)** | **-214.62 (3.68e-5)** |
| QG | **2.66 (5.81e-3)** | 0.59 (2.26e-2) | -13.44 (7.56e-2) | -213.96 (9.03e-3) |
| | | $\varphi = 0.95$ | | |
| Optimum | 32.57 | 6.31 | -8.92 | -214.63 |
| QO-TSP | 33.80 (1.94e-1) | 6.38 (7.47e-3) | -8.87 (2.40e-3) | -214.57 (1.91e-3) |
| QO-CRN | **33.81 (2.80e-1)** | **6.33 (4.72e-3)** | **-8.89 (1.71e-4)** | **-214.62 (1.27e-5)** |
| QG | 49.13 (3.62) | 19.59 (1.36) | -4.27 (1.58e-1) | -211.51 (5.16e-2) |

Table 3: Performance of QO-TSP, QO-CRN, and QG (Normal distribution case).

| Problem | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| | | $\varphi = 0.6$ | | |
| Optimum | 2.27 | 0.25 | -14.98 | -214.63 |
| QO-TSP | 2.34 (8.69e-3) | 0.26 (6.43e-4) | -14.76 (6.13e-3) | -214.62 (3.13e-4) |
| QO-CRN | 2.33 (1.59e-2) | **0.25 (1.16e-4)** | **-14.77 (6.86e-3)** | **-214.62 (1.54e-4)** |
| QG | **2.28 (2.26e-3)** | 0.49 (1.64e-2) | -14.13 (5.14e-2) | -214.12 (1.07e-2) |
| | | $\varphi = 0.95$ | | |
| Optimum | 9.22 | 1.64 | -13.59 | -214.63 |
| QO-TSP | 9.25 (4.90e-3) | 1.65 (3.07e-4) | -13.56 (8.92e-4) | -214.61 (2.88e-4) |
| QO-CRN | **9.24 (5.61e-3)** | **1.64 (2.57e-7)** | **-13.57 (7.48e-5)** | **-214.62 (1.88e-5)** |
| QG | 9.25 (4.94e-3) | 1.65 (5.86e-4) | -12.66 (6.48e-2) | -213.59 (1.60e-2) |

## 5 CONCLUSIONS

In this paper, we have proposed a new FD-based gradient descent algorithm for solving differentiable quantile optimization problems under a black-box scenario. The key observation is that the quantile gradient shares the same direction as the negative output c.d.f. gradient, which can be effectively approximated using only the system/model output information through an SP-style estimator. The gradient estimator, when combined

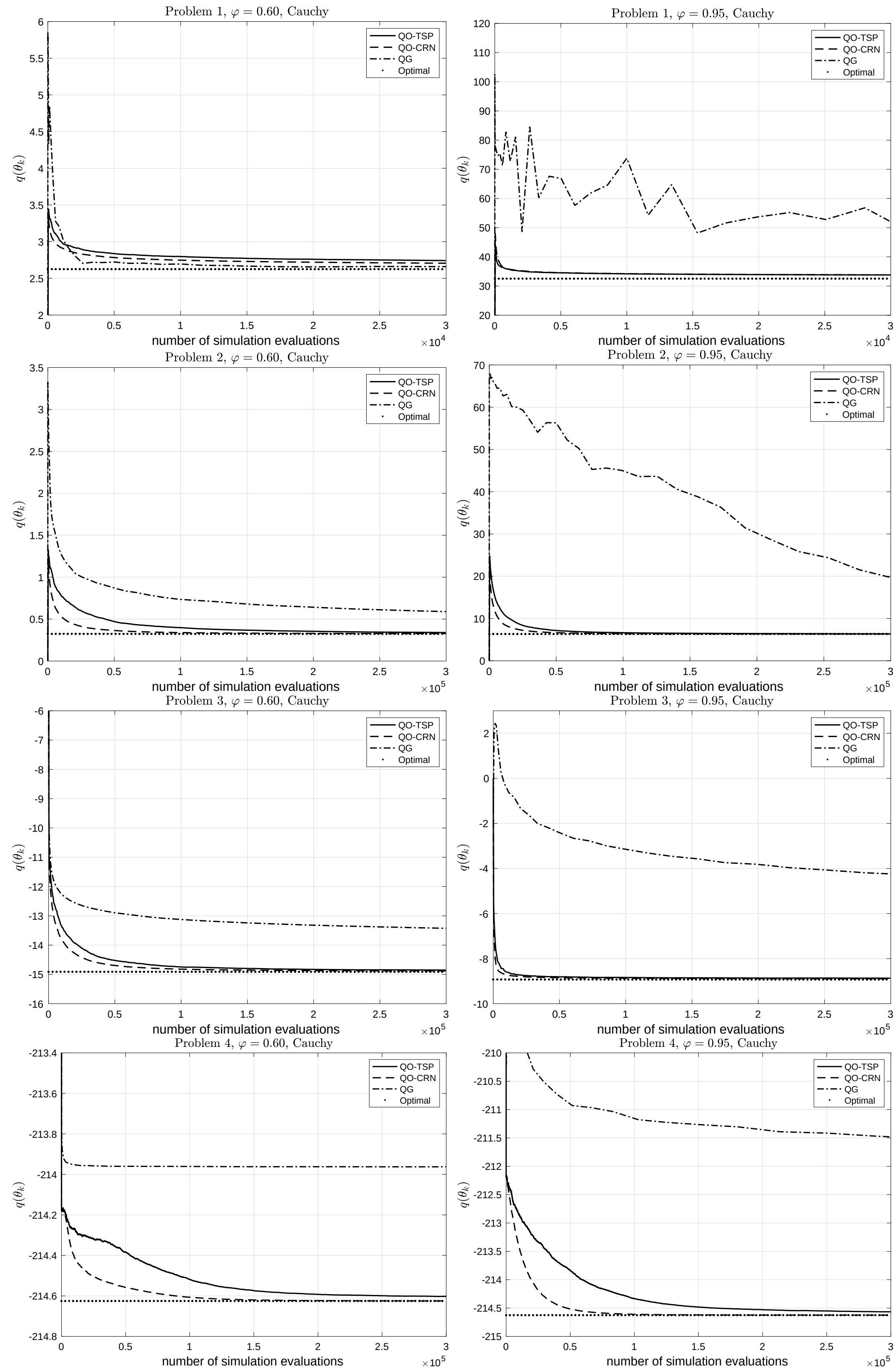


Figure 1: Performance of QO-TSP, QO-CRN and QG under Cauchy noise.

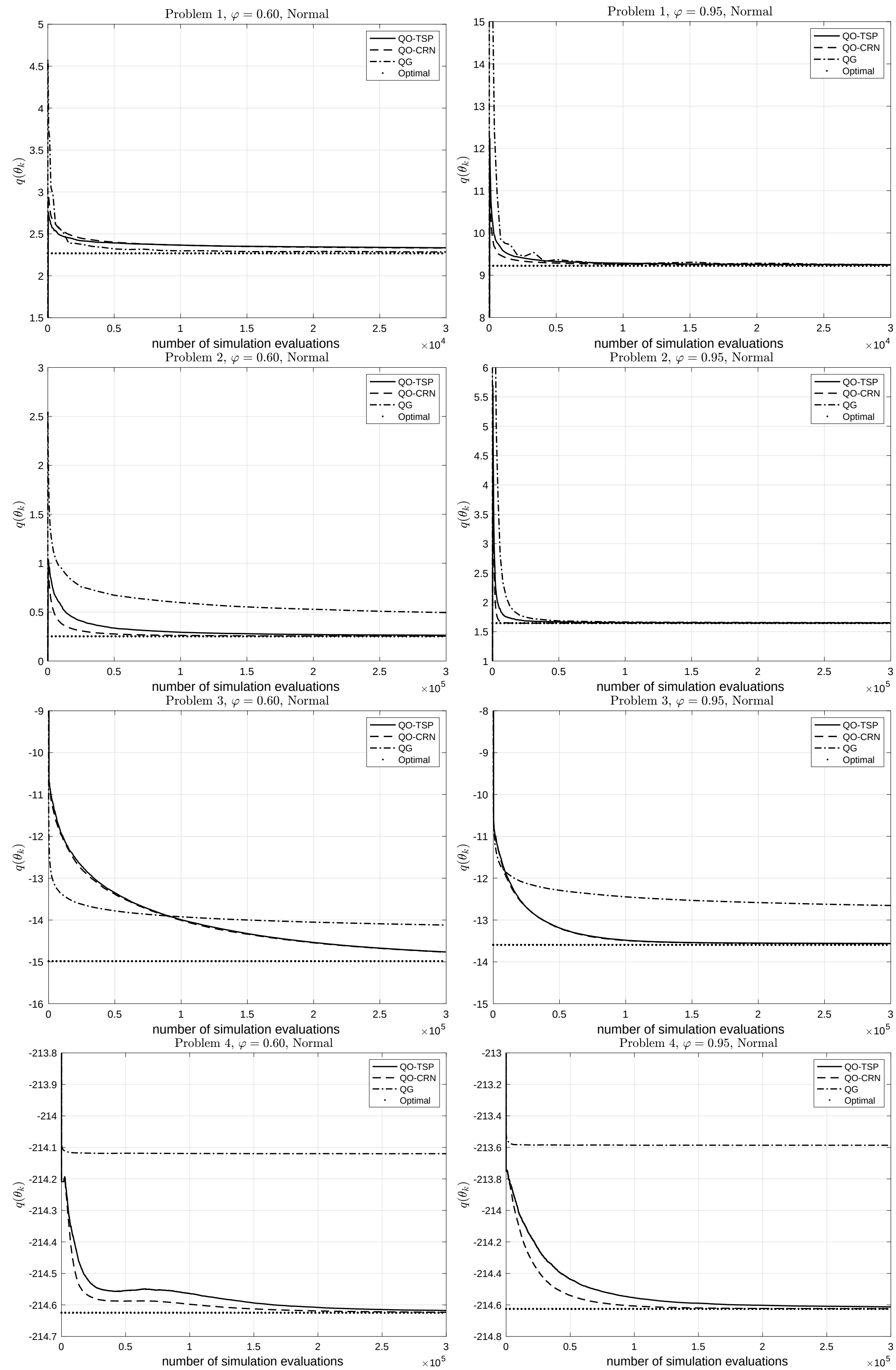


Figure 2: Performance of QO-TSP, QO-CRN and QG under Normal noise.

with a separate recursive quantile estimation procedure, gives rise to a two-timescale SA algorithm that requires only three observations of the output random variable at each iteration. We have shown the general local convergence property of the algorithm and provided a finite-time performance bound to characterize its rate of convergence. Preliminary numerical results suggest that the algorithm could lead to improved performance over an earlier method that relies on traditional FD and order statistics in quantile gradient estimation. However, one practical issue is the choice of algorithm parameters. In particular, because the c.d.f. gradient only provides the descent direction of the quantile function, not its magnitude, our experiments indicate that a good choice of the step-size parameter appears to be highly problem-dependent. How to determine these parameters is an important issue that remains to be investigated.

## ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under Grants CMMI-2027527, IIS-2123684, and by Air Force Office of Scientific Research under Grant FA95502010211.

## REFERENCES

Borkar, V. S. 2008. *Stochastic Approximation: A Dynamical Systems Viewpoint*. UK: Cambridge University Press.

Chang, K.-H., and H.-K. Lu. 2014. "Solving Quantile-based Stochastic Optimization Problems with Modified Stochastic Nelder-Mead Simplex Method". In *Proceedings of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 374–379. IEEE.

Conejo, A. J., M. Carrión, J. M. Morales et al. 2010. *Decision Making Under Uncertainty in Electricity Markets*, Volume 1. Springer.

Fu, M. C. 2015. *Handbook of Simulation Optimization*. New York, NY: Springer Science+Business Media.

Fu, M. C., and S. D. Hill. 1997. "Optimization of Discrete Event Systems via Simultaneous Perturbation Stochastic Approximation". *IIE Transactions* 29(3):233–243.

Fu, M. C., L. J. Hong, and J.-Q. Hu. 2009. "Conditional Monte Carlo Estimation of Quantile Sensitivities". *Management Science* 55(12):2019–2027.

Ghadimi, S., and G. Lan. 2012. "Optimal Stochastic Approximation Algorithms for Strongly Convex Stochastic Composite Optimization I: A Generic Algorithmic Framework". *SIAM Journal on Optimization* 22(4):1469–1492.

Hu, J., Y. Peng, G. Zhang, and Q. Zhang. 2022. "A Stochastic Approximation Method for Simulation-Based Quantile Optimization". *INFORMS Journal on Computing* 34(6):2889–2907.

Hu, J., M. Song, and M. C. Fu. 2023. "Quantile Optimization via Multiple Timescale Local Search for Black-box Functions". *arXiv preprint arXiv:2308.07607*.

Jiang, J., J. Hu, and Y. Peng. 2022. "Quantile-based Policy Optimization for Reinforcement Learning". In *Proceedings of the 2022 Winter Simulation Conference*, edited by B. Feng, G. Pderielli, Y. Peng, S. Shashaani, E. Song, C. G. Corlu, L. H. Lee, and P. Lendermann, 2712–2723. Piscataway, NJ, USA: IEEE Press.

Kibzun, A. I., and V. Kurbakovskiy. 1991. "Guaranteeing Approach to Solving Quantile Optimization Problems". *Annals of Operations Research* 30:81–93.

Kibzun, A. I., and E. Matveev. 2012. "Optimization of the Quantile Criterion for the Convex Loss Function by a Stochastic Quasi-Gradient Algorithm". *Annals of Operations Research* 200(1):183–198.

Kibzun, A. I., A. V. Naumov, and V. I. Norkin. 2013. "On Reducing a Quantile Optimization Problem with Discrete Distribution to a Mixed Integer Programming Problem". *Automation and Remote Control* 74(6):951–967.

Kim, J. H., and W. B. Powell. 2011. "Quantile Optimization for Heavy-Tailed Distributions Using Asymmetric Signum Functions". *Working Paper*:https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.299.9913&rep=rep1&type=pdf, accessed 15th August 2023.

Kleinman, N. L., J. C. Spall, and D. Q. Naiman. 1999. "Simulation-based Optimization with Stochastic Approximation Using Common Random Numbers". *Management Science* 45(11):1463–1611.

Kushner, H. J., and G. G. Yin. 1997. *Stochastic Approximation and Recursive Algorithms and Applications*. New York, NY: Springer.

Rockafellar, R. T. 2020. "Risk and Utility in the Duality Framework of Convex Analysis". In *From Analysis to Visualization*, edited by D. H. Bailey, N. S. Borwein, R. P. Brent, R. S. Burachik, J.-A. H. Osborn, B. Sims, and Q. J. Zhu, 21–42: Springer International Publishing.

Sabater, C., O. Le Maître, P. M. Congedo, and S. Görtz. 2021. "A Bayesian Approach for Quantile Optimization Problems with High-Dimensional Uncertainty Sources". *Computer Methods in Applied Mechanics and Engineering* 376:113632.

Spall, J. C. 1992. "Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation". *IEEE Transactions on Automatic Control* 37(3):332–341.

Spall, J. C. 2003. *Introduction to Stochastic Search and Optimization*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Spall, J. C., and D. C. Chin. 1997. "Traffic-Responsive Signal Timing for System-Wide Traffic Control". *Transportation Research Part C: Emerging Technologies* 5:153–163.

Vasiléva, S. N., and Y. S. Kan. 2015. "A Method for Solving Quantile Optimization Problems with a Bilinear Loss Function". *Automation and Remote Control* 76(9):1582–1597.

Wang, S., S. H. Ng, and W. B. Haskell. 2021. "A Multilevel Simulation Optimization Approach for Quantile Functions". *INFORMS Journal on Computing* 34(1):569–585.

Zamar, D. S., B. Gopaluni, S. Sokhansanj, and N. K. Newlands. 2017. "A Quantile-Based Scenario Analysis Approach to Biomass Supply Chain Optimization under Uncertainty". *Computers & Chemical Engineering* 97(2):114–123.

## AUTHOR BIOGRAPHIES

**MEICHEN SONG** is a Ph.D. student in the Department of Applied Mathematics and Statistics at the State University of New York, Stony Brook. She received the B.S. degree in Mathematics from Southeast University, Nanjing, China. Her research interests are simulation optimization, quantile optimization, and stochastic approximation. Her e-mail address is meichen.song@stonybrook.edu.

**JIAQIAO HU** is an Associate Professor in the Department of Applied Mathematics and Statistics at the State University of New York, Stony Brook. He received the B.S. degree in automation from Shanghai Jiao Tong University, the M.S. degree in applied mathematics from the University of Maryland, Baltimore County, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park. His research interests include Markov decision processes, applied probability, and simulation optimization. He is currently Department Editor for IISE Transactions and Associate Editor for Operations Research. His e-mail address is jiaqiao.hu.1@stonybrook.edu.

**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, at the University of Maryland, College Park. His research interests include stochastic gradient estimation, simulation optimization, and applied probability. He served as WSC2011 Program Chair and received the INFORMS Simulation Society's Distinguished Service Award in 2018. He is a Fellow of INFORMS and IEEE. His e-mail addresses is mfu@umd.edu.