

## RESPONSE SURFACE COMPUTATION VIA SIMULATION IN THE PRESENCE OF CONVEXITY

Eunji Lim

Industrial Engineering  
University of Miami  
Coral Gables, FL 33124, USA

### ABSTRACT

We consider the problem of computing a response surface when the underlying function is known to be convex. We introduce a methodology that incorporates the convexity into the function estimator. The proposed response surface estimator is formulated as a finite dimensional quadratic program and exhibits convergence properties as a global approximation to the true function. Numerical results are presented to illustrate the convergence behavior of the proposed estimator and its potential application to simulation optimization.

### 1 INTRODUCTION

In this paper, we consider the problem of computing a response surface for the underlying function  $g_* : \mathbb{R}^d \rightarrow \mathbb{R}$  with the goal of obtaining a global approximation to  $g_*(\cdot)$  where  $g_*(\cdot)$  can only be computed via simulation at each  $x$  in the domain of  $g_*(\cdot)$ . Computation of response surfaces has a long history due to its wide applications in the areas of simulation metamodeling and optimization. Traditional approaches to the computation of response surfaces include fitting a low order polynomial, and using common random numbers or antithetic variates; see [Box and Wilson \(1951\)](#), [Law and Kelton \(2000\)](#), and [Myers and Montgomery \(2002\)](#) for details.

In this paper, we realize that in many stochastic models, the response surfaces of interest  $g_*(\cdot)$  are known to satisfy a certain shape characteristic such as convexity in the underlying parameters even though closed-form formulas for  $g_*(\cdot)$  are not available. For instance, (1) the mean waiting time of a customer in a single server queue is convex in the mean service and interarrival times when the interarrival and service times belong to certain types of classes ([Shanthikumar and Yao 1991](#)); (2) in tandem, cyclic, or tree-like network of queues, the mean sojourn time of a customer is convex in the mean service times at the servers when the interarrival and service times belong to certain types of classes ([Shanthikumar and Yao 1991](#)); (3) in a tandem queueing network with finite intermediate buffer storage spaces and independent, exponential interarrival and service times, the throughput is concave in the buffer storage capacities ([Meester and Shanthikumar 1990](#)); (4) in a multiserver queue with no buffer storage spaces, the fraction of arrivals that are lost is convex in the number of servers ([Wolff and Wang 2002](#)).

Another setting where the convexity of the response surface is naturally assumed is the case where a consumer utility function is estimated through empirical data in the economics context; see [Meyer and Pratt \(1968\)](#) for example. Convexity also arises in estimation of supply and demand functions.

In the presence of convexity, we wish to force the convexity into our response surface estimator to get a better global approximation to  $g(\cdot)$ . The main idea of our proposed response surface estimator is to fit a convex function to a given dataset to minimize the sum of squared errors. When the parameter space is one-dimensional, this approach leads to a well-known “convex regression” estimator in the statistics literature, which can be formulated as a solution to a finite-dimensional quadratic program (QP); see [Hildreth \(1954\)](#), [Hanson and Pledger \(1976\)](#), [Mammen \(1991\)](#), and [Groeneboom, Jongbloed, and Wellner \(2001\)](#) for the formulation, consistency, and convergence rate of the convex regression estimator in one-dimensional parameter spaces.

However, when the parameter space is multi-dimensional, fitting a convex function may appear to involve solving an infinite dimensional minimization problem which is computationally intractable. In this paper, we show that this minimization can, in fact, be formulated as a finite-dimensional QP with  $(d+1)n$  decision variables and  $n(n-1)$  linear constraints where  $n$  is the number of points in the domain. Hence, we extend the idea of fitting a convex function to multi-dimensional cases and propose a multi-dimensional convex regression estimator. We also derive a

desired property of the proposed estimator as a global approximation to  $g_*(\cdot)$ , which is the uniform almost sure (a.s.) convergence over a compact subset of the domain.

Incorporating other shape characteristics such as monotonicity into the function estimator has been studied intensively in the statistics literature under the name of the “isotonic regression” estimator; see [Lim and Glynn \(2006\)](#) for the properties of the isotonic regression estimator in the settings that are suitable for the simulation environment.

This paper is organized as follows. In Section 2, we introduce the mathematical framework for our proposed estimator and Section 3 includes some properties of the proposed estimator as a global approximation to  $g_*(\cdot)$ , which is the uniform a.s. convergence over a compact subset of the domain. In Section 4, we present some numerical results that illustrate the convergence behavior of the proposed estimator and compare the performance of the proposed estimator to that of the standard regression method. We also illustrate the use of the proposed estimator in the context of sensitivity analysis in simulation optimization. Finally, Section 5 includes concluding remarks and future research directions.

## 2 PROBLEM FORMULATION

With the goal of computing a function  $g_* : \mathbb{R}^d \rightarrow \mathbb{R}$ , we observe the simulated value  $Y_i$  of  $g_*$  at  $X_i$  and presume that  $Y_i$  satisfies

$$Y_i = g_*(X_i) + \varepsilon_i$$

for  $i \geq 1$ , where the “noise” satisfies  $E\varepsilon_i = 0$  and  $E\varepsilon_i^2 < \infty$ . To estimate  $g_*(\cdot)$  from the observed pairs  $(X_1, Y_1), \dots, (X_n, Y_n)$ , we propose the estimator  $\hat{g}_n(\cdot)$  which is the minimizer of the sum of squares

$$\sum_{i=1}^n (Y_i - g(X_i))^2 \tag{1}$$

over functions  $g \in \mathcal{C} \triangleq \{g : \mathbb{R}^d \rightarrow \mathbb{R} \text{ such that } g \text{ is convex}\}$ . When the parameter space is one-dimensional ( $d = 1$ ), the minimization problem (1) can be formulated as

$$\begin{aligned} \min \quad & \sum_{i=1}^n (Y_i - g(X_i))^2 \\ \text{s/t} \quad & (Y_2 - Y_1)/(X_2 - X_1) \leq \dots \leq (Y_n - Y_{n-1})/(X_n - X_{n-1}), \end{aligned} \tag{2}$$

provided  $X_1 < \dots < X_n$ . (2) is a QP with  $n$  decision variables and  $n - 2$  linear constraints, so  $\hat{g}_n(\cdot)$  can be computed easily via a quadratic program solver.

However, when the parameter space is multi-dimensional ( $d > 1$ ), the minimization (1) may appear to be computationally intractable since the set of convex functions is infinite-dimensional. But, it turns out that this minimization can be formulated as a finite-dimensional QP as the following result shows ([Lim and Glynn 2010](#)).

**Proposition 1.** *Consider the quadratic program (in the decision variables  $(g_1, \xi_1), \dots, (g_n, \xi_n)$ )*

$$\begin{aligned} \min \quad & \frac{1}{n} \sum_{i=1}^n w(X_i) (Y_i - g_i)^2 \\ \text{s/t} \quad & g_j \geq g_i + \xi_i^T (X_j - X_i), \quad 1 \leq i, j \leq n. \end{aligned} \tag{3}$$

*For  $n \geq d + 1$ , the QP (3) has a minimizer  $(\hat{g}_1, \hat{\xi}_1), \dots, (\hat{g}_n, \hat{\xi}_n)$  and the minimizing values  $\hat{g}_1, \dots, \hat{g}_n$  are unique. Furthermore, any minimizer  $\tilde{g}_n$  of (1) satisfies  $\tilde{g}_n(X_i) = \hat{g}_i$  for  $1 \leq i \leq n$ .*

The main idea behind Proposition 1 is that the constraints of (1) can be written as linear inequalities because a convex function passing through  $(X_1, g_1), \dots, (X_n, g_n)$  exists if and only if there exist  $\xi_1, \dots, \xi_n \in \mathbb{R}^d$  (subgradients) at  $X_1, \dots, X_n$  satisfying

$$g_j \geq g_i + \xi_i^T (X_j - X_i), \quad i, j = 1, \dots, n.$$

Subgradients of a convex function are illustrated in Figure 1 for the case where  $d = 1$  and  $n = 3$ .

While Proposition 1 asserts that  $(\hat{g}_1, \dots, \hat{g}_n)$  is unique, there are many convex functions  $g$  satisfying  $g(X_i) = \hat{g}_i$  for  $1 \leq i \leq n$ . To define our estimator  $\hat{g}_n(x)$  at  $x \neq X_i$ ,  $1 \leq i \leq n$ , we set

$$\hat{g}_n(x) = \sup\{g(x) : g \in \mathcal{C}, g(X_i) = \hat{g}_i, 1 \leq i \leq n\}. \tag{4}$$

Lim

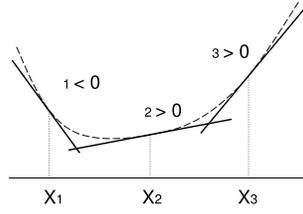


Figure 1: Subgradients of a convex function

The function  $\widehat{g}_n(\cdot)$  is guaranteed to be convex; see, for example, p. 75 of [Avriel \(1976\)](#). Furthermore,  $\widehat{g}_n(\cdot)$  is finite-valued on  $\text{conv}(X_1, \dots, X_n)$  (where  $\text{conv}(A) \triangleq$  convex hull of  $A$ , for  $A \subset \mathbb{R}^n$ ), and infinite elsewhere.

In principle, the maximization that defines (4) again appears to be infinite-dimensional (over  $\mathcal{C}$ ). However, as with Proposition 1, the next result makes clear that  $\widehat{g}_n(x)$  can be computed as the solution of a finite-dimensional linear program ([Lim and Glynn 2010](#)).

**Proposition 2.** For each  $x \in \mathbb{R}^d$ ,  $\widehat{g}_n(x)$  can be computed as the optimal value  $\widehat{y}$  of the LP (in the decision variables  $y, \xi_1, \dots, \xi_n, \tilde{\xi}$ )

$$\begin{aligned} \max \quad & y \\ \text{s/t} \quad & \widehat{g}_j \geq \widehat{g}_i + \xi_i^T (X_j - X_i), \quad 1 \leq i, j \leq n \\ & y \geq \widehat{g}_i + \xi_i^T (x - X_i), \quad 1 \leq i \leq n \\ & \widehat{g}_j \geq y + \tilde{\xi}^T (X_j - x), \quad 1 \leq j \leq n. \end{aligned}$$

The convex function  $\widehat{g}_n(\cdot)$  is our estimator for  $g_*(\cdot)$ .

### 3 CONVEX REGRESSION AS GLOBAL APPROXIMATION

This section emphasizes the role of  $\widehat{g}_n(\cdot)$  as a global approximation to  $g_*(\cdot)$ . In particular, one of the desired properties of the global approximation is the convergence to  $g_*(\cdot)$  uniformly over a subset of  $\mathbb{R}^d$ . It turns out that our proposed estimator  $\widehat{g}_n(\cdot)$  has this property as  $n \rightarrow \infty$  even though the simulated values  $Y_i$ 's do not necessarily converge to the true values. This kind of behavior occurs due to the convexity constraint enforced to  $\widehat{g}_n(\cdot)$ .

In order to analyze the proposed estimator  $\widehat{g}_n(\cdot)$ , we will impose some probabilistic assumptions on the  $(X_i, Y_i)$ 's. In particular, we require that:

- A1.  $g_* : \mathbb{R}^d \rightarrow \mathbb{R}$  is convex.
- A2.  $X, X_1, X_2, \dots$  is a sequence of independent and identically distributed (iid)  $\mathbb{R}^d$ -valued random vectors having a common continuous positive density  $k(\cdot)$ .
- A3. For  $i \geq 1$ ,  $Y_i = g_*(X_i) + \varepsilon_i$ , where the  $\varepsilon_i$ 's satisfy

$$P(\varepsilon_i \in dz_i, 1 \leq i \leq n | X_1, X_2, \dots) = \prod_{i=1}^n F(dz_i | X_i)$$

- for some family  $(F(\cdot|x) : x \in \mathbb{R}^d)$  of cumulative distribution functions.
- A4.  $E(Y_1^2 + \|X_1\|^2 + 1) < \infty$ , thereby implying that

$$\sigma^2(X_1) \triangleq \int_{\mathbb{R}} z^2 F(dz | X_1) < \infty \quad \text{a.s.}$$

- A5. For each  $x \in \mathbb{R}^d$ ,

$$\int_{\mathbb{R}} z F(dz | x) = 0.$$

The following results justify our choice of  $\widehat{g}_n(\cdot)$  as a global approximation to  $g_*(\cdot)$ ; see [Lim and Glynn \(2010\)](#) for details.

**Theorem 1.** Assume A1–A5. Then, for any compact subset  $C$  of  $\mathbb{R}^d$ ,

$$\sup_{x \in C} |\widehat{g}_n(x) - g_*(x)| \rightarrow 0 \quad \text{a.s.} \quad \text{as } n \rightarrow \infty.$$

**Corollary 2.** Assume A1 and A3–A5. Assume that for any subset  $A$  of  $\mathbb{R}^d$ ,  $\liminf_{n \rightarrow \infty} A_n/n > 0$  where  $A_n$  is the number of the  $X_i$ 's in  $A$ . Then, for any compact subset  $C$  of  $\mathbb{R}^d$ ,

$$\sup_{x \in C} |\widehat{g}_n(x) - g_*(x)| \rightarrow 0 \quad \text{a.s.} \quad \text{as } n \rightarrow \infty.$$

#### 4 NUMERICAL RESULTS

In this section, we investigate the performance of our proposed estimator  $\widehat{g}_n(\cdot)$  in the settings of consumer preference models and queueing models. We will compare the performance of the proposed convex regression estimator to that of the standard regression technique. Throughout this section,  $X_i$ 's will be equally spaced over a hyperrectangle in  $\mathbb{R}^d$ . For instance, when  $n = 25$  and  $d = 2$ ,  $X_i$ 's will be distributed as in Figure 2. We recognize that (3) requires solving a QP with  $(d + 1)n$  decision variables and  $n(n - 1)$  constraints. To reduce the computational burden in computing  $\widehat{g}_n(\cdot)$  when  $n$  gets large, we consider another computational strategy. In particular, we consider the least squares estimator  $\widetilde{g}_n$  over all *coordinate-wise* convex functions, which is the minimizer of the sum of squares

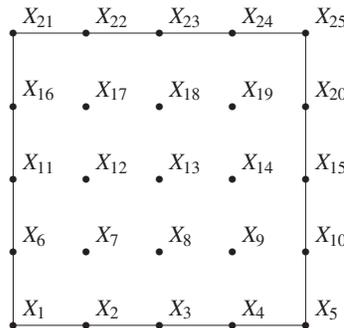
$$\sum_{i=1}^n (Y_i - g(X_i))^2$$

over functions  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  that are “coordinatewise-convex” (i.e.,  $g$  is convex in each coordinate). This simplification allows one to compute  $\widetilde{g}_n$  via a QP with lower complexity especially when the  $X_i$ 's are equally spaced. For instance, when  $n = 25$ ,  $d = 2$ , and the  $X_i$ 's are equally spaced over a rectangle in  $\mathbb{R}^2$  as illustrated in Figure 2,  $\widetilde{g}_n$  is the solution to the following QP (in the decision variables  $g_1, \dots, g_{25}$ ):

$$\begin{aligned} \min_{g_1, \dots, g_{25}} \quad & \sum_{i=1}^{25} (Y_i - g_i)^2 & (5) \\ \text{s/t} \quad & -g_1 + 2g_2 - g_3 \leq 0 \\ & -g_2 + 2g_3 - g_4 \leq 0 \\ & \vdots \\ & -g_1 + 2g_6 - g_{11} \leq 0 \\ & -g_2 + 2g_7 - g_{12} \leq 0 \\ & \vdots \end{aligned}$$

because when  $x_1, \dots, x_n$  are equally spaced points in  $\mathbb{R}$  such that  $x_1 < x_2 < \dots < x_n$ , a convex function passing through  $(x_1, g_1), \dots, (x_n, g_n)$  exists if and only if  $x_2 - x_1 \leq \dots \leq x_n - x_{n-1}$ . Observe that (5) is a QP with  $n$  decision variables and  $O(n)$  constraints, which enjoys less computational burden than (3).

Figure 2:  $X_i$ 's are equally spaced over a rectangle in  $\mathbb{R}^2$  when  $n = 25$ .



To compare the performance of  $\widehat{g}_n$  and  $\widetilde{g}_n$  to that of the standard regression technique, we compute the multiple regression fit to the  $(X_i, Y_i)$ 's. Let  $\widehat{a}_n(0), \dots, \widehat{a}_n(d)$  be the solution to

$$\min_{a_0, a_1, \dots, a_d} \sum_{i=1}^n (Y_i - (a_0 + a_1 X_1^i + \dots + a_d X_d^i))^2,$$

where  $X_i = (X_1^i, \dots, X_d^i)$ . For  $x = (x_1, \dots, x_d)$ , the multiple regression  $\widehat{l}_n(x)$  is computed by  $\widehat{l}_n(x) = \widehat{a}_n(0) + \widehat{a}_n(1)x_1 + \dots + \widehat{a}_n(d)x_d$ . The behavior of  $\widehat{g}_n$  and  $\widetilde{g}_n$  will be compared to that of  $\widehat{l}_n$  across a wide range of  $n$  values. Quadratic programming problems in (3) and (5) were solved using “cvx” which was designed by Grant and Boyd and available at <http://cvxr.com/>.

### 4.1 Estimating Functions Using Convex Regression

#### 4.1.1 Consumer Preference Function

**2-dimensional case:** Suppose we wish to predict consumer preferences by estimating a consumer utility function. We assume there is an underlying utility function  $u : [0.8, 0.86] \times [0.1, 0.16] \rightarrow \mathbb{R}$  which is given by

$$u(x_1, x_2) = 10x_1^{1/2} + 20x_2^{1/2} \tag{6}$$

for a goods basket  $(x_1, x_2) \in [0.8, 0.86] \times [0.1, 0.16]$ . Now suppose we are given some consumer preference data  $(X_1, Y_1), \dots, (X_n, Y_n)$ , where the  $X_i$ 's are equally spaced over  $[0.8, 0.86] \times [0.1, 0.16]$  as in Figure 2, the noisy measurement  $Y_i$  of  $u(X_i)$  is given by  $u(X_i) + \varepsilon_i$ , and  $\varepsilon_i$  follows a uniform distribution between  $-0.1$  and  $0.1$ . Table 1 reports the averages (Mean) and standard deviations (Std) of the  $Y_i$ 's,  $\widehat{l}_n$ ,  $\widehat{g}_n$  and  $\widetilde{g}_n$  at  $X_i = (0.82, 0.12), (0.82, 0.14), (0.84, 0.12)$ , and  $(0.84, 0.14)$  over 100 replications for each value of  $n$ . – indicates that “cvx” could not solve the corresponding QP due to limited computational capacity. While  $\widehat{g}_n$  and  $\widetilde{g}_n$  converge nicely to the true value,  $\widehat{l}_n$  exhibits bias as  $n \rightarrow \infty$  due to the nonlinearity of  $u$  and the  $Y_i$ 's fail to converge due to high variance.

Table 1: Performance of  $Y_i$ 's,  $\widehat{l}_n(\cdot)$ ,  $\widehat{g}_n(\cdot)$  and  $\widetilde{g}_n(\cdot)$  as an estimator for  $u(\cdot)$  in (6)

		$X_i = (0.82, 0.14)$							
		$Y_i$		$\widehat{l}_n(X_i)$		$\widehat{g}_n(X_i)$		$\widetilde{g}_n(X_i)$	
$n$	Mean	Std	Mean	Std	Mean	Std	Mean	Std	
16	16.54	0.06	16.52	0.02	16.56	0.03	16.56	0.03	
49	16.55	0.06	16.52	0.01	16.55	0.02	16.55	0.02	
100	16.54	0.06	16.52	0.01	16.55	0.01	16.55	0.01	
169	16.54	0.05	16.53	0.01	16.55	0.01	16.54	0.01	
256	16.54	0.06	16.53	0.00	16.54	0.01	16.54	0.01	
8281	16.54	0.06	16.53	0.00	–	–	16.54	0.00	
$u(X_i)$		16.54							
		$X_i = (0.84, 0.14)$							
		$Y_i$		$\widehat{l}_n(X_i)$		$\widehat{g}_n(X_i)$		$\widetilde{g}_n(X_i)$	
$n$	Mean	Std	Mean	Std	Mean	Std	Mean	Std	
16	16.65	0.06	16.63	0.02	16.67	0.03	16.66	0.03	
49	16.65	0.06	16.63	0.01	16.66	0.02	16.66	0.01	
100	16.65	0.06	16.64	0.01	16.66	0.01	16.66	0.01	
169	16.66	0.06	16.64	0.01	16.65	0.01	16.65	0.01	
256	16.65	0.06	16.64	0.00	16.65	0.01	16.65	0.01	
8281	16.65	0.06	16.64	0.00	–	–	16.65	0.00	
$u(X_i)$		16.65							

**3-dimensional case:** Suppose the underlying utility function  $u : [0.8, 0.86] \times [0.6, 0.66] \times [0.1, 0.16] \rightarrow \mathbb{R}$  is now given by

$$u(x_1, x_2, x_3) = 5x_1^{1/2} + 10x_2^{1/2} + 15x_3^{1/2} \tag{7}$$

for  $(x_1, x_2, x_3) \in [0.8, 0.86] \times [0.6, 0.66] \times [0.1, 0.16]$ . Consumer preference data  $(X_1, Y_1), \dots, (X_n, Y_n)$  are given, where the  $X_i$ 's are equally spaced over  $[0.8, 0.86] \times [0.6, 0.66] \times [0.1, 0.16]$  as in Figure 2, the noisy measurement  $Y_i$  of  $u(X_i)$  is

given by  $u(X_i) + \varepsilon_i$ , and  $\varepsilon_i$  follows a uniform distribution between  $-0.1$  and  $0.1$ . Table 2 reports the averages (Mean) and standard deviations (Std) of the  $Y_i$ 's,  $\widehat{l}_n$ ,  $\widehat{g}_n$  and  $\widetilde{g}_n$  at  $X_i = (0.82, 0.62, 0.12)$ ,  $(0.82, 0.64, 0.14)$ ,  $(0.84, 0.62, 0.14)$ , and  $(0.84, 0.64, 0.12)$  over 100 replications for each value of  $n$ . “–” indicates that “cvx” could not solve the corresponding QP due to limited computational capacity.

Table 2: Performance of  $Y_i$ 's,  $\widehat{l}_n(\cdot)$ ,  $\widehat{g}_n(\cdot)$ , and  $\widetilde{g}_n(\cdot)$  as an estimator for  $u(\cdot)$  in (7)

$X_i = (0.82, 0.62, 0.12)$								
$n$	$Y_i$		$\widehat{l}_n(X_i)$		$\widehat{g}_n(X_i)$		$\widetilde{g}_n(X_i)$	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
64	17.61	0.06	17.58	0.01	17.62	0.02	17.62	0.02
4096	17.61	0.06	17.59	0.00	–	–	17.60	0.00
$u(X_i)$	17.60							
$X_i = (0.84, 0.62, 0.14)$								
$n$	$Y_i$		$\widehat{l}_n(X_i)$		$\widehat{g}_n(X_i)$		$\widetilde{g}_n(X_i)$	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
64	18.08	0.05	18.05	0.01	18.09	0.02	18.09	0.02
4096	18.06	0.06	18.06	0.00	–	–	18.07	0.00
$u(X_i)$	18.07							

### 4.1.2 Single Server Queue

We consider a single server queue with infinite buffer space where the interarrival times and service times are i.i.d. uniform random variables on  $[x_0 - 0.5, x_0 + 0.5]$  and  $[x_1 - 0.5, x_1 + 0.5]$ , respectively. The interarrival times and service times are independent and the service discipline is first in/first out (FIFO). We wish to compute the expected waiting time  $w_{100}(x_0, x_1)$  of the 100th customer for  $(x_0, x_1) \in [2.92, 2.98] \times [2.85, 2.91]$ . Even though no explicit formula for  $w_{100}(\cdot, \cdot)$  is known, the convexity of  $w_{100}(\cdot, \cdot)$  is established; see p. 137 of [Shanthikumar and Yao \(1991\)](#) for details. To compute  $w_{100}$ , we chose  $n$  points  $X_1, \dots, X_n$  equally spaced over  $[2.92, 2.98] \times [2.85, 2.91]$  as in Figure 2. At each  $X_i = (x_0^i, x_1^i)$ , we simulated the single server queue with the mean interarrival and service times equal to  $x_0^i$  and  $x_1^i$ , respectively, and computed the waiting time  $Y_i$  of the 100th customer. We then obtained the average of 30 independent copies of  $Y_i$ , say  $\bar{Y}_i$ . Using  $(X_1, \bar{Y}_1), \dots, (X_n, \bar{Y}_n)$ , we computed  $\widehat{l}_n$ ,  $\widehat{g}_n$  and  $\widetilde{g}_n$  as our estimator for  $d_{10}$ .

Table 3 reports the averages (Mean) and standard deviations (Std) of the  $\bar{Y}_i$ 's,  $\widehat{l}_n$ ,  $\widehat{g}_n$  and  $\widetilde{g}_n$  at  $X_i = (2.89, 2.94)$  and  $(2.89, 2.96)$  over 100 replications for each value of  $n$ . The value of  $w_{100}(X_i)$  in the last row of Table 3 is obtained from averaging 100,000 independent copies of  $Y_i$  for each  $i$ . “–” indicates that “cvx” could not solve the corresponding QP due to limited computational capacity.

### 4.1.3 Tandem Queue

We consider a system of three single-server stations in tandem where the interarrival times are i.i.d. uniform random variables on  $[x_0 - 0.5, x_0 + 0.5]$  and the service times are i.i.d. uniform random variables on  $[x_i - 0.5, x_i + 0.5]$  for server  $i$ ,  $i = 1, 2, 3$ . The interarrival times and service times are independent and the service discipline is FIFO at all servers. Each server has unlimited buffer space. We wish to compute the expected sojourn time  $s_{600}(x_0, x_1, x_2, x_3)$  of the 600th customer for  $(x_0, x_1, x_2, x_3) \in [2.92, 2.94] \times [2.85, 2.91] \times [2.85, 2.91] \times [2.85, 2.91]$ . The convexity of  $s_{600}$  is known; see p. 141 of [Shanthikumar and Yao \(1991\)](#) for details. To compute  $s_{600}$ , we chose  $n$  points  $X_1, \dots, X_n$  equally spaced over  $[2.92, 2.94] \times [2.85, 2.91] \times [2.85, 2.91] \times [2.85, 2.91]$  as in Figure 2. At each  $X_i = (x_0^i, x_1^i, x_2^i, x_3^i)$ , we simulated the tandem queue with the mean interarrival and service times equal to  $x_0^i, x_1^i, x_2^i$ , and  $x_3^i$ , respectively, and computed the sojourn time  $Y_i$  of the 600th customer. We then obtained the average of 100 independent copies of  $Y_i$ , say  $\bar{Y}_i$ . Using  $(X_1, \bar{Y}_1), \dots, (X_n, \bar{Y}_n)$ , we computed  $\widetilde{g}_n$  as our estimator for  $s_{600}$ . Table 4 reports the averages (Mean) and standard deviations (Std) of the  $Y_i$ 's,  $\widehat{l}_n$ , and  $\widehat{g}_n$  at  $X_i = (2.93, 2.88, 2.88, 2.88)$  over 100 replications for each value of  $n$ . The value of  $s_{600}(2.93, 2.88, 2.88, 2.88)$  in the last row of Table 4 is obtained from averaging 500,000 independent copies of  $Y_i$  at  $X_i = (2.93, 2.88, 2.88, 2.88)$ .

## 4.2 Sensitivity Analysis in Simulation Optimization

**Tandem Queue** We consider a system of two single-server stations in tandem where the external arrivals follow a Poisson process with a rate of 0.25 and the service times are i.i.d. exponential random variables with mean parameter  $x_i$  for server  $i$ ,  $i = 1, 2$ . The interarrival times and service times at all stations are independent and the service discipline

Table 3: Performance of  $Y_i$ 's,  $\widehat{l}_n(\cdot)$ ,  $\widehat{g}_n(\cdot)$ , and  $\widetilde{g}_n(\cdot)$  as an estimator for  $d_{10}(\cdot)$

		$X_i = (2.89, 2.94)$							
		$Y_i$		$\widehat{l}_n(X_i)$		$\widehat{g}_n(X_i)$		$\widetilde{g}_n(X_i)$	
$n$		Mean	Std	Mean	Std	Mean	Std	Mean	Std
16		1.26	0.28	1.41	0.09	1.21	0.29	1.20	0.16
49		1.30	0.25	1.40	0.05	1.25	0.07	1.27	0.08
100		1.32	0.27	1.38	0.04	1.26	0.05	1.27	0.06
169		1.27	0.24	1.37	0.03	1.28	0.04	1.28	0.05
256		1.31	0.25	1.37	0.02	1.28	0.03	1.29	0.04
8281		1.34	0.25	1.35	0.01	–	–	1.30	0.01
$w_{100}(X_i)$		1.30							
		$X_i = (2.89, 2.96)$							
		$Y_i$		$\widehat{l}_n(X_i)$		$\widehat{g}_n(X_i)$		$\widetilde{g}_n(X_i)$	
$n$		Mean	Std	Mean	Std	Mean	Std	Mean	Std
16		0.94	0.22	1.09	0.06	0.89	0.10	0.88	0.10
49		0.96	0.20	1.07	0.04	0.92	0.06	0.93	0.06
100		0.93	0.20	1.06	0.03	0.92	0.04	0.93	0.05
169		0.98	0.21	1.05	0.02	0.93	0.03	0.94	0.03
256		0.94	0.19	1.05	0.02	0.93	0.02	0.94	0.03
8281		0.82	0.19	1.04	0.00	–	–	0.95	0.01
$w_{100}(X_i)$		0.95							

Table 4: Performance of  $Y_i$ 's,  $\widehat{l}_n(\cdot)$ , and  $\widetilde{g}_n(\cdot)$  as an estimator for  $s_{600}(\cdot)$

		$X_i = (2.93, 2.88, 2.88, 2.88)$					
		$Y_i$		$\widehat{l}_c(X_i)$		$\widetilde{g}_n(X_i)$	
$n$		Mean	Std	Mean	Std	Mean	Std
81		13.04	0.30	14.56	0.04	12.99	0.25
625		13.05	0.28	14.15	0.02	13.00	0.11
2401		13.03	0.28	14.01	0.01	13.03	0.06
$s_{600}(X_i)$		13.03					

is FIFO at all stations. Each server has unlimited buffer space. The costs related to this system are the waiting cost which is proportional to the steady-state mean sojourn time of a customer in the system and the operational cost at each server which is proportional to the service rate at each server. Suppose the unit waiting cost at each server is 1 and the unit operational cost for servers 1 and 2 are 1 and 2, respectively. So the total cost per unit time  $c(x_1, x_2)$  is given by

$$\begin{aligned}
 c(x_1, x_2) &= \text{Steady-state mean sojourn time} + 1 \cdot (\text{Service rate at server 1}) + 2 \cdot (\text{Service rate at server 2}) \\
 &= \left( \frac{x_1}{1 - 0.25x_1} + \frac{x_2}{1 - 0.25x_2} \right) + 1 \cdot \frac{1}{x_1} + 2 \cdot \frac{1}{x_2}.
 \end{aligned}$$

Currently, the mean service times at the two server are set be to 0.8 and 2.8, respectively. We now wish to consider five alternatives for the mean service time at server 2, namely  $x_2^{(1)} = 0.3, x_2^{(2)} = 0.8, x_2^{(3)} = 1.3, x_2^{(4)} = 1.8, x_2^{(5)} = 2.3$ . In particular, we wish to estimate the reduction in the total cost per unit time for each alternative.

First, 100 customers are simulated independently at five alternatives,  $x_2^{(i)}, 1 \leq i \leq 5$ , and at  $x_2^{(0)} = 2.8$ , for the service rate in server 2 with fixed interarrival rate of 0.25 and service rate at server 1 of 0.8. Let  $S_i, 0 \leq i \leq 5$ , be the average of the 100 customers' sojourn time, then the total cost at  $(x_1, x_2) = (0.8, x_2^{(i)})$  is computed by

$$Y_i \triangleq S_i + 1/0.8 + 2/x_2^{(i)},$$

for  $0 \leq i \leq 5$ . We now recognize that the mean sojourn time of tandem queueing system with this type of interarrival and service times is convex in the mean interarrival time and in the mean service time at each server; see p. 141 in [Shanthikumar and Yao \(1991\)](#) for details. So  $c(\cdot)$  is convex in the mean service time at server 2. Now our estimator  $\hat{g}_n$ , which is identical to  $\tilde{g}_n$  in one-dimensional case, for  $c(\cdot)$  is computed using  $(x_2^{(0)}, Y_0), \dots, (x_2^{(5)}, Y_5)$ . Table 4 reports the computed reduction in the total cost for each of the five alternatives based on the  $Y_i$ 's and the proposed estimator  $\hat{g}_n(\cdot)$ . The averages (Mean) and standard deviations (Std) are obtained based on 1000 replications. The proposed estimator shows superior performance in terms of both the mean and standard deviation.

Table 5: Performance of  $\hat{g}_n(0.8, x_2^{(0)}) - \hat{g}_n(0.8, \cdot)$  as an estimator for  $c(0.8, x_2^{(0)}) - c(0.8, \cdot)$

	Mean		Std	
	$Y_0 - Y_1$	$\hat{g}_n(0.8, x_2^{(0)}) - \hat{g}_n(0.8, x_2^{(1)})$	$Y_0 - Y_1$	$\hat{g}_n(0.8, x_2^{(0)}) - \hat{g}_n(0.8, x_2^{(1)})$
	2.22	3.17	4.30	3.29
$c(0.8, x_2^{(0)}) - c(0.8, x_2^{(1)})$	3.06			
	Mean		Std	
	$Y_0 - Y_2$	$\hat{g}_n(0.8, x_2^{(0)}) - \hat{g}_n(0.8, x_2^{(2)})$	$Y_0 - Y_2$	$\hat{g}_n(0.8, x_2^{(0)}) - \hat{g}_n(0.8, x_2^{(2)})$
	4.94	5.28	3.68	3.58
$c(0.8, x_2^{(0)}) - c(0.8, x_2^{(2)})$	5.66			

### 5 CONCLUDING REMARKS

As discussed in the numerical studies, the computation of our proposed estimator  $\hat{g}_n(\cdot)$  involves solving a QP with  $(d + 1)n$  decision variables and  $n(n - 1)$  linear constraints, which becomes computationally challenging as  $n$  increases. Consequently, it would be valuable to develop efficient algorithms for computing the convex regression estimator which do not involve solving a QP. In the setting of estimating monotone functions, numerical procedures for computing the isotonic regression estimator have received a great deal of attention and various algorithms have been proposed to compute the isotonic regression estimator in linear time  $O(n)$ ; see [Barlow et al. \(1972\)](#) for details.

Another subject of future research is the verification of the convex property in more general settings of queueing networks. For instance, the throughput of a closed queueing network of Jackson type is shown to be concave in the number of jobs; see [Shanthikumar and Yao \(1988\)](#) for details. So one open question is the identification of sufficient conditions for more general queueing networks having convex performance measures.

### ACKNOWLEDGMENTS

The author thanks Peter W. Glynn for helpful comments and suggestions.

### REFERENCES

Avriel, M. 1976. *Nonlinear programming: Analysis and methods*. Englewood Cliffs, NJ: Prentice-Hall.

Barlow, R. E., D. J. Bartholomew, J. M. Bremner, and H. D. Brunk. 1972. *Statistical inference under order restrictions*. New York: Wiley.

Box, G., and K. Wilson. 1951. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society* 13:1–45.

Groeneboom, P., G. Jongbloed, and J. A. Wellner. 2001. Estimation of a convex function: characterizations and asymptotic theory. *Ann. Statist.* 29:1653–1698.

Hanson, D. L., and G. Pledger. 1976. Consistency in concave regression. *Ann. Statist.* 4:1038–1050.

Hildreth, C. 1954. Point estimates of ordinates of concave functions. *J. Amer. Statist. Assoc.* 49:598–619.

Law, A. M., and W. D. Kelton. 2000. *Simulation modeling and analysis*. New York: McGraw-Hill.

Lim, E., and P. W. Glynn. 2006. Simulation-based response surface computation in the presence of monotonicity. In *Proceedings of the 2006 Winter Simulation Conference*, 264–271.

Lim, E., and P. W. Glynn. 2010. Consistency of multi-dimensional convex regression. *Submitted*.

Mammen, E. 1991. Nonparametric regression under qualitative smoothness assumptions. *Ann. Statist.* 19:741–759.

Meester, L. E., and J. G. Shanthikumar. 1990. Concavity of the throughput of tandem queueing systems with finite buffer storage space. *Adv. Appl. Prob.* 22:764–767.

Meyer, R. F., and J. W. Pratt. 1968. The consistent assessment and fairing of preference functions. *IEEE Transactions on Systems Science and Cybernetics* 4:270–278.

Myers, R. H., and D. C. Montgomery. 2002. *Response surface methodology*. New York: Wiley.

- Shanthikumar, J. G., and D. D. Yao. 1988. Second-order properties of the throughput of a closed queueing network. *Mathematics of Operations Research* 13:524–534.
- Shanthikumar, J. G., and D. D. Yao. 1991. Strong stochastic convexity: closure properties and applications. *J. Appl. Prob.* 28:131–145.
- Wolff, R. W., and C.-L. Wang. 2002. On the convexity of loss probabilities. *J. Appl. Prob.* 39:402–406.

#### **AUTHOR BIOGRAPHY**

**EUNJI LIM** is an Assistant Professor in the Industrial Engineering Department at the University of Miami. She received her Ph.D. in Management Science and Engineering from Stanford University. Her research interest includes statistical inference under shape restrictions and simulation optimization. Her email address is [lim@miami.edu](mailto:lim@miami.edu).