Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

HISTORY OF SEEKING BETTER SOLUTIONS, AKA SIMULATION OPTIMIZATION

Michael C. Fu

Robert H. Smith School of Business,
Institute for Systems Research
University of Maryland
College Park, MD 20742, USA

Shane G. Henderson

School of Operations Research
and Information Engineering
Cornell University
Ithaca, NY 14853, USA

ABSTRACT

Simulation optimization — arguably the ultimate aim of most simulation users — has had a long and illustrious history closely tied with the 50 years of the Winter Simulation Conference (WSC). We touch upon the historical developments of the field, highlighting research progress and their interactions with WSC. Specific areas covered include ranking & selection methods, stochastic approximation and gradient estimation, response surface methodology, and sample average approximation. We discuss the interplay between research and practice, including software developments. We conclude with some reflections on the state of the field.

1 INTRODUCTION

The history of seeking better solutions using simulation is far older than the 50 years of the Winter Simulation Conference that is celebrated with this series of WSC50 history track articles, although the term "simulation optimization" itself may not have been in common usage at that beginning. Indeed, seeking better solutions is arguably the whole purpose of simulation, and simulation users have experimented with input parameters to simulations throughout its history. However, the scope of this article will be limited to simulation optimization, by which we mean the optimization of simulation models where there is a clearly defined objective function that is estimated from a simulation model, in contrast to optimization search methods using sampling-based methods, e.g., simulated annealing. Thus, the terms "simulation optimization" and "optimization via simulation" are treated as synonyms throughout.

Simulation optimization is now a vibrant and important field, as evidenced through a number of observations:

- Simulation optimization is a chapter in many standard textbooks or references (Banks et al. 2010; Kleijnen 2008; Law 2015).
- Simulation optimization is a regular track at the Winter Simulation Conference.
- The recent publication of the *Handbook of Simulation Optimization* (Fu 2015a), with the variety of techniques and principles discussed within, demonstrates the field's maturity.
- In a recent survey of commercial simulation software (Swain 2015), 40 out of 55 software products claim to have some level of optimization capability.

Moreover, there have been many high-impact examples of simulation optimization in practice. Two prominent examples are as follows.

- Part of a major supply-chain planning activity at Caterpillar (Rao, Scheller-Wolf, and Tayur 2000) rested on innovations developed through a collaboration with Carnegie-Mellon University. The *Fortune* magazine article (Siekman 2000) summarizing the successful implementation states:

> "Among the techniques the Carnegie-Mellon group used to attack this complex problem was so-called infinitesimal perturbation analysis, for which no complete explanation is possible for the faint-hearted or mathematically disadvantaged... The technique involves first the creation of an algorithm and computer program that provides a solution to a problem but not the optimal answer, partly because it's based on a mix of fact and educated guesses... The team then slightly changes a guess, possibly increasing a number, and lets the computer try again, recalculating the possible paths... If that seems to move things in the right direction... the assumptions are nudged upward a bit more. If not, they are nudged down."

- One of the 2013 finalists for the Franz Edelman Award for Achievement in Operations Research was an inventory management project with Kroger. The abstract of the *Interfaces* article (Zhang et al. 2014) states:

  > "The Kroger Co. is the largest grocery retailer in the United States. It operates 2,422 supermarkets and 1,950 in-store pharmacies. ... The [simulation-optimization] system was implemented in October 2011 in all Kroger pharmacies in the United States, and has reduced out-of-stocks by 1.6 million per year, ensuring greater patient access to medications. It has resulted in an increase in revenue of $80 million per year, a reduction in inventory of more than $120 million, and a reduction in labor cost equivalent to $10 million per year."

  This is a high-dimensional problem, as the authors write (Zhang et al. 2014, p.72): "simulation-optimization for inventory systems, such as (s,S) systems, presents computational challenges when one considers the need to find solutions for more than 2,000 drugs at each of the 1,950 stores." "Our simulation-optimization approach follows a sample-path problem, as Fu and Healy (1997) propose." (Zhang et al. 2014, p.76)

In this article we do not attempt to provide a comprehensive review of simulation optimization; the field is simply too large. Rather, we try to gather the dominant ideas, and explore their origins. In writing this article we contacted prominent researchers and developers in the area to learn more about their experiences. We include many of their thoughts where appropriate.

Since the **Winter Simulation Conference (WSC)** has been a primary driver of developments in simulation optimization from the outset of the conference, we devote a section to describing its various roles. See §2.

In the remainder of this introduction we will briefly discuss each of several lines of inquiry that have been pursued with the goal of identifying improving solutions. Given the somewhat independent development of each of these streams in the research literature, a separate presentation for all but the last two streams (deterministic optimization and random search) is discussed in substantially more detail in later sections.

Stochastic approximation, in contrast to selection methods discussed later, is applied to problems that have a continuous solution space, where the function one is attempting to optimize is usually assumed to be differentiable. The idea is to employ a gradient *search* with a decreasing step size, using (statistical) estimates of gradients in place of true gradients. Stochastic approximation has a very long history, with the initial work completed in the 1950s, and now has an enormously rich literature. Entry points to that literature include Kushner and Yin (2003), Chau and Fu (2015), Ghadimi and Lan (2015). The area remains very active. See §3.

Gradient estimates are needed for many simulation optimization algorithms, including stochastic approximation. Perhaps surprisingly, it is often possible to obtain estimates of gradients from a *single* simulation replication, i.e., without the need for multiple simulations for finite difference estimates, which can become computationally prohibitive in high-dimensional problems. Larry Ho and his colleagues were early pioneers in this field, which blossomed in the 1980s. Surveys of gradient estimation can be found in Fu (2006), Glasserman (2004), Fu (2015b). From a theoretical standpoint, this led to a flourishing research stream on (stochastic) *direct* gradient estimation, meaning methodologies for providing *unbiased* estimates, where finite-difference estimates, aside from requiring multiple simulations, are inherently biased. While perturbation analysis techniques generally refer to a sample path analysis approach, the likelihood ratio (aka score function) method and weak derivatives method focus on the underlying measure from which the system is simulated, analogous to importance sampling in variance reduction techniques. See §4.

Response-surface methodology (RSM) is another technique for optimizing functions over continuous domains. RSM can be subdivided into local and global metamodeling techniques. The essential idea behind local metamodeling methods is to repeatedly fit a local linear or quadratic model to the function being optimized, and step to the optimal solution of the approximating model. In contrast, global metamodeling methods develop an approximation to the function being optimized that applies over its entire domain. Kriging is an important method within this latter class. Entry points to the literature include Barton and Meckesheimer (2006), Kleijnen (2015). See §5.

Sample-average approximation is not an algorithm, but rather is a principle that can be employed to develop algorithms. In its simplest form it involves fixing the sample size used to evaluate a solution, and using common random numbers to synchronize the function evaluations across different solutions. The same idea was discovered under a different name (Healy and Schruben 1991) and later extended to allow the sample size to increase as progress is made (Chen and Schmeiser 2001). Much is known about sample-average approximation, relating as it does to maximum likelihood estimation. Important surveys include Shapiro (2003), Shapiro, Dentcheva, and Ruszczyński (2009), and an accessible introduction can be found in Kim, Pasupathy, and Henderson (2015). See §6.

Selection methods are a collection of techniques for identifying the best from a finite set of solutions. A key assumption in the ranking and selection subclass of selection methods is that the number of solutions is small enough that every solution can be simulated, at least to some degree. Traditionally the solution set has been assumed to be fixed, so that no search, per se, is involved, and so the only decision to be made is the allocation of simulation replications to the feasible solutions. Many ranking and selection methods provide statistical guarantees, akin to the statistical guarantees associated with confidence intervals for a (population) mean derived from a sample. In fact, some authors would restrict the term "ranking and selection" to apply only to methods that provide such guarantees. However, we take a broader view, including concepts and methods that do not necessarily provide strong statistical guarantees, but that do provide at least some kind of asymptotic guarantee. The additional methods that we discuss include ordinal optimization, optimal computing budget allocation (OCBA), knowledge gradient and expected improvement, and random search. Excellent entry points to the literature on these topics, which dates back to the 1950s, include Kim and Nelson (2006), Andradóttir (2006), Ólafsson (2006), Hong, Nelson, and Xu (2015), Chen, Chick, and Lee (2015). See §7.

Deterministic-optimization methods have frequently been adapted for use in simulation-optimization problems. Prominent examples in this sphere include the Nelder-Mead method (Barton and Ivey Jr. 1996), the cross-entropy method (Rubinstein 1999), model-reference adaptive search (Hu, Fu, and Marcus 2005; Hu, Fu, and Marcus 2007), nested partitions (Shi and Ólafsson 2000a; Shi and Ólafsson 2000b), and adaptive search (Zabinsky and Smith 1992; Zabinsky 2003). Due to lack of space, we do not explore these methods in detail, although we do mention some of them in related sections in the sequel.

Random-search methods are algorithms that explicitly include randomness, typically to ensure that a local-search algorithm can escape from local minima, and thus ensure global convergence. The term usually refers to problems with discrete variables, although randomization is present in many continuous-variable optimization algorithms, as well. Since randomization is so ubiquitous in simulation-optimization algorithms, and again due to space limitations, we do not provide a separate section on these methods but instead discuss selected references throughout. Important work includes the stochastic ruler method (Yan and Mukai 1992), random-search method (Andradóttir 1995), stochastic comparisons (Gong, Ho, and Zhai 1999), simulated annealing (Alrefaei and Andradóttir 1999), nested partitions (Shi and Ólafsson 2000b; Pichitlamken and Nelson 2003), COMPASS (Hong and Nelson 2006), Industrial-Strength COMPASS (Xu, Nelson, and Hong 2010), the adaptive hyperbox algorithm (Xu, Nelson, and Hong 2013), R-SPLINE

(Wang, Pasupathy, and Schmeiser. 2013), and Gaussian process-based search (Sun, Hong, and Hu 2014); see Andradóttir (2015) for a recent summary.

Software has been developed for both commercial and research applications. Few of the methods discussed thus far have been adopted in commercial simulation software. Perhaps the difficulty lies in the need to have algorithms that apply to poorly structured problems and that have clean and simple user interfaces. Commercial software developers have instead opted for metaheuristic approaches that, while not offering guarantees, demonstrate admirable generality and applicability. Tabu search has become the primary method adopted in software, although genetic algorithms are also employed (OptTek 2017; ProModel 2017). Non-commercial software, coming out of the research community, is emerging. Perhaps the most prominent example is Industrial-Strength COMPASS (Xu, Nelson, and Hong 2010). A collection of algorithms and test problems can be found at Henderson and Pasupathy (2017). See §8.

In §9, we reflect on our tour of the history of simulation optimization, with a view towards the future of simulation optimization. We conclude the paper in §10 with a timeline of important events, which seems appropriate for a paper related to discrete-*event* simulation!

2 IMPACT OF THE WINTER SIMULATION CONFERENCE

The Winter Simulation Conference (WSC) has played a central role in the development of simulation optimization. The WSC is, in our opinion, the conference with the highest impact on the field, so in this section we highlight some historical evidence for its impact.

Working from the Winter Simulation Conference Proceedings archive (WSC 2017), it is evident that simulation optimization has figured in the WSC at least since the very first proceedings in 1968. Indeed, Thompson (1968) provided a conceptual discussion of simulation optimization. Emery (1969) described an algorithm, closely related to coordinate search, for a specific class of applications. Mihram (1970) explores the possibility of using response-surface methodology for simulation optimization, focusing the discussion on two-variable problems. Lefkowitz and Schriber (1971) wrote FORTRAN software to interface with GPSS to solve a univariate optimization problem.

Although these papers represented the earliest examples of simulation optimization at the WSC, several of our interviewees pointed to the papers Biles (1973), Biles (1974) as the first instance of simulation optimization at the conference. These papers may have had more impact because they provided a general-purpose algorithm that could be easily adapted by others.

Introductory tutorials became a regular track with full papers in 1977. Prior to that, the coverage was more sporadic, focusing more on software and most commonly without a full accompanying paper. The first WSC tutorial addressing simulation optimization was Garzia, Schmidt, and Garzia (1980). This tutorial discussed both mathematical programming for simulations and search algorithms, but apparently did not have a wide influence on the field, as Google Scholar was unable to provide any citation information.

A watershed moment in the WSC's impact on simulation optimization came in 1986, with two WSC papers by Peter Glynn, comprising the first two entries in Table 1. This table provides an historical perspective of WSC impact on simulation optimization in terms of citations. Note that "Stochastic approximation for Monte Carlo optimization" and "Methods for selecting the best system" were also two of the 10 landmark WSC proceedings papers selected at the 40th anniversary of WSC in 2007, also indicating the prominence of simulation optimization at WSC.

In the decade leading up to 1986, WSC showed robust activity in simulation optimization research, as indicated by the titles of papers presented during that period, listed in Table 2.

WSC has served as a welcome forum for new ideas in simulation optimization, continuing such a role more broadly in terms of simulation modeling and analysis methodology over the fifty years of its existence. Many ideas first appeared in briefer WSC proceedings papers before they were more formalized in journal articles. One prominent example, to be described later, is the concept of single-run optimization, first introduced by Suri and Leung (1987) at WSC. Among others, Lee Schruben has often followed this path, with a couple of prominent examples relating to sensitivity analysis and gradient estimation in the form of

Table 1: Highly cited WSC proceedings papers on simulation optimization (Google Scholar, Apr.2017).

Year	# Citations	Title
1986	93	Optimization of Stochastic Systems
1986	147	Stochastic Approximation for Monte Carlo Optimization
1989	119	Optimization of Stochastic Systems via Simulation
1991	95	Methods for Selecting the Best System
1992	144	A Tutorial on Simulation Optimization
1997	378	Simulation Optimization: Methods and Applications
1998	253	A Review of Simulation Optimization Techniques
1999	303	Simulation Optimization Methodologies
2000	286	A Survey of Simulation Optimization Techniques and Procedures
2001	141	Simulation Optimization
2002	201	Simulation Optimization
2003	335	Practical Introduction to Simulation Optimization
2005	389	Simulation Optimization: A Review, New Developments, and Applications

Table 2: WSC proceedings papers on simulation optimization in the late 1970s through early 1980s.

Year	Title
1986	Stochastic Approximation for Monte Carlo Optimization
	Simulation Optimization Using Frequency Domain Methods
	Tutorial on Indifference-Zone Normal Means Ranking and Selection Procedures
1985	An Application of Optimization-by-Simulation to Discrete Variable Systems
	Optimization of Manufacturing System Simulations Using Perturbation Analysis and SENSE
1984	A Simulation Optimization Approach to Optimum Storage and Retrieval Policies in an Automated Warehousing System
	Application of an Optimization Procedure to Steady-State Simulation
	Simulation Optimization for Decision Support in Operating a Robotic Manufacturing System
1984	A Simulation/Optimization-Based Planning and Decision Support System
1978	Superimposing Direct Search Methods for Parameter Optimization onto Dynamic Simulation Models
	Simulation Optimization Using Response Surfaces Based on Spline Approximations
	Use of Both Optimization and Simulation Models to Analyze Complex Systems

frequency domain experimentation. Both ideas ended up playing a prominent role in several PhD theses of subsequently successful leading scholars in the simulation community, including Sheldon Jacobson, Doug Morrice, Paul Sanchez, and Victor (Wai-Kin) Chan.

Another interesting line of research first introduced at WSC by Lee Schruben (Schruben 2000) is to represent the dynamics of a discrete-event system sample path through mathematical programming, and thus try to exploit the computational efficiencies and structural insights from the optimization community. This approach, more fully developed in Chan and Schruben (2003), Chan and Schruben (2008), bears some similarity to the (max,+) algebra of discrete-event systems (Heidergott 2010). Subsequent work demonstrated how IPA estimators could be simultaneously generated from the sample paths (Chan and Schruben 2006).

3 STOCHASTIC APPROXIMATION

Stochastic approximation (SA) traces its roots back to the two seminal papers Robbins and Monro (1951), Kiefer and Wolfowitz (1952). Both papers are rooted in statistics, appearing in prominent statistics journals, but arguably the biggest developments were made initially in the systems and control community, as we shall shortly describe. One measure of the impact that these two papers have had is captured in the number of Google Scholar citations: over 4700 for the former and over 1500 for the latter. In terms of simulation optimization, SA is an iterative algorithm based on gradient search, and these two SA algorithms directly connect to algorithms driven by direct gradients versus those relying on finite-difference gradients. If the simulation model is treated as a black box, then methods based on finite difference gradient estimates are basically the only option available to the simulation user. The usual finite-difference estimates lead to a linear scaling of computational effort as a function of the decision variable, e.g., $2d$ simulations for a d-dimensional symmetric difference gradient estimate. Arguably, one of the most significant advances in black-box gradient-based approaches is simultaneous perturbation stochastic approximation (SPSA) (Spall 1992) – with over 1400 Google Scholar citations, because the number of simulations is reduced to 2 simulations regardless of the dimension. This is accomplished by randomizing in every direction rather than one at a time, hence the moniker simultaneous perturbation. This idea is particularly beneficial in the simulation optimization context, because the simulations are the expensive part of the computation, so that the randomization itself is essentially negligible in terms of computation, whereas in other contexts, the randomizing itself may constitute a non-trivial portion of the total computation, especially in high dimensions. An analogy in the design of experiments would be to reduce a full 2^k-factorial design to a randomized 2-design experiment regardless of the number of input variables in the system.

Two prominent MIT labs played a role in the earlier days, specifically Lincoln Labs and Draper Lab. Harold Kushner and Rajan Suri mentioned them (respectively) in their interviews. Harold Kushner first thought about simulation optimization while at Lincoln Labs, where he was "part of a group with lots of freedom," and the director of the group encouraged him to look into Robbins' work. That was the beginning of his foray into stochastic approximation, which led among other things to two of the most well-known books in the area and also led him to brief forays into perturbation analysis, to be described in more detail in the next section on gradient estimation. Also deferred to there will be the origins of perturbation analysis itself, which were based on work done on transfer lines for FIAT under a contract to Draper Labs.

Substantial developments along theoretical fronts include more sophisticated and generally applicable proof techniques. These include the ordinary differential equation (ODE) approach, the martingale approach, and the weak convergence approach; see Kushner and Clark (1979), Kushner and Yin (1997), Kushner and Yin (2003). Asymptotic convergence rate results were another theoretical advance, and central limit theorem results for the iterates were first established by Fabian (1968), Fabian (1971), which spawned much follow-up theoretical work. Another more recent significant advance, both in terms of practical significance and theoretical importance, was iterate averaging, introduced in Ruppert (1988), Polyak and Juditsky (1992). Stochastic approximation has a huge body of literature, and many books summarize both the theory and numerous applications, e.g., Pflug (1996), Spall (2003), so the focus here will be on developments closest to and most relevant to and/or driven by the WSC community.

Probably the earliest direct introduction of SA to the WSC community was through the two 1986 WSC papers (Glynn 1986a; Glynn 1986b). Prior to that, it seems hard to argue that the WSC community was aware of these methods. The latter paper was the first to apply SA algorithms in a setting familiar to the WSC community, using both finite differences and LR/SF gradient estimates for provably convergent algorithms; shortly thereafter, the first convergence proof of an SA algorithm using IPA gradient estimates was provided in Fu (1990). The proof required updates at regenerative points, which was relaxed by Chong and Ramadge (1992), Chong and Ramadge (1993) (for a regenerative G/G/1 queue). Chong (2013)[p.77] states "In fact, even an algorithm that updates after every customer behaves similarly to one that updates after every busy period." This was the idea behind the so-called "single-run optimization" (an idea suggested independently by Larry Ho and Marc Meketon, according to Suri and Zazanis (1988)[page 58]) whereby SA was employed in the optimization of steady-state discrete-event dynamic systems. Whereas at that time, it might take simulating thousands of customers to reach steady state just to estimate the stationary waiting time in a single-server queue, empirical results from single-run optimization, first introduced at WSC by

Suri and Leung (1987) and later more formalized in Suri and Leung (1989), indicated that simulation optimization could be accomplished on approximately the same time scale, i.e., rather than simply having a single sensitivity estimate at the end of a simulation run, the parameter(s) could also be optimized by the completion of the single run. This will be described in more detail in the next section.

Outside of the WSC community, the focus was on applications in systems and control and more sophisticated and general proof techniques, whereas the simulation community seemed more interested in fresh approaches to the algorithms themselves, although still desiring provable convergence guarantees. Sigrun Andradóttir made numerous contributions to stochastic approximation in the simulation optimization setting, many of which were first introduced at WSC, e.g., Andradóttir (1990), Andradóttir (1991), including new types of projections in the constrained setting. Another form of averaging introduced at WSC in the Robbins-Monro framework (Fu and Ho 1988) averaged the gradients to smooth the step size sequence.

As mentioned earlier, one of the potentially most relevant algorithms to simulation optimization is SPSA, proposed in Spall (1992) (although a related algorithm called random directions stochastic approximation (RDSA) is mentioned in Kushner and Clark 1979). SPSA is particularly suited for high-dimensional problems based on "black box" simulations, i.e., those settings in which single-run gradient estimates are not available. The general SPSA algorithm was introduced to the WSC community in Spall (1994), and applied specifically to simulation optimization of discrete-event systems in Fu and Hill (1997), which received the 1998 *IIE Transactions* Best Paper Award. Under the setting where direct gradients such as PA or LR/SF are available, Chau, Qu, and Fu (2014), Chau, Fu, and Qu (2017) introduce an algorithm called STAR-SPSA, which uses a Secant Tangents AveRaged (STAR) gradient that is a convex combination of a direct gradient estimate and an SP gradient.

Another novel SA idea introduced first at WSC was the MRAS conversion of discrete optimization problems via SA and parameters of distributions (Hu and Hu 2010).

WSC's contribution to SA shows no signs of a slow down, as a quick perusal of recent titles shows three papers each in the last two years with "stochastic approximation" in the title.

4 GRADIENT ESTIMATION

Being able to estimate the gradients without the need for resimulation, i.e., a single simulation replication would return both an estimate of performance along with estimates of sensitivities with respect to parameters of the underlying system, was a watershed development in simulation optimization research (at least in the admittedly biased opinion of one of the co-authors). Through the last four decades, WSC has served to highlight many of these advances, whose historical developments are detailed in this section. This section is probably the most specialized or narrow in the paper, again reflecting the clear bias of one of the co-authors, and can be easily skipped without loss of flow if the reader is not interested in the details of these developments; however, there may be more stories (controversy?) here than elsewhere in the paper.

As highlighted earlier in the discussion of stochastic approximation, a key component of gradient-based simulation optimization is the presence of a "good" gradient estimator, where good means ideally unbiased, low variance, and computationally efficient. The "brute-force" method of estimating gradients is using finite differences, which results in an estimator that is generally biased, grows in computation linearly with the dimension of the underlying parameter (a notable exception being the simultaneous perturbation approach described in the previous section), and may suffer from high variance if the difference size is too small. However, the chief advantage of techniques based on finite differences over the techniques discussed in this section is that they require no knowledge of the system being simulated, which is treated as a black box. Technical details summarizing the various approaches can be found in Fu (2006), Fu (2015b).

One of the most important contributions of the perturbation analysis approach was the focus on sample path analysis, which provided another link between the (stochastic) simulation and applied probability communities. Such connections included later explorations of the relationship between gradient estimation and Malliavin calculus, which was especially explored in the finance setting. One thing is almost sure tautologically speaking: the set of simulationists who had heard of Malliavin calculus fifty years ago was

Table 3: INFORMS Simulation Society Outstanding Publication Award winners.

Year	relationship to gradient estimation
1990	1988 MS article on IPA
1992	1991 book on IPA
1998	1997 book on SPA
2005	2004 simulation book w/ chapter on gradient estimation
2008	2007 simulation book w/ chapter on gradient estimation
2012	2009 OR and MS articles for quantile & CVaR sensitivities

of measure zero. In any case, a testimony to the impact that the approach had on the field is David Yao's opening sentence in his Foreword to the book, *Stochastic Simulation Optimization for Discrete Event Systems: Perturbation Analysis, Ordinal Optimization and Beyond* (Chen, Jia, and Lee 2013):

> "One of the best ideas in the field of systems and control over the last three decades is *perturbation analysis* (PA)... Like all great ideas, PA has an elegant simplicity, built upon first principles."

A sampling of key concepts in gradient estimation research (whether from the systems and control community or applied probability community), some of which are closely related to sample path analysis, include the following:

- differentiating a random variable (Suri 1983);
- differentiating a sample path vs. the underlying measure (distribution);
- structural vs. distributional parameters (push in and push out methods);
- theory for unbiasedness: interchange of integration and differentiation (limits), uniform integrability, dominated convergence theorem, role of continuity (a.s. and Lipschitz), (generalized) mean-value theorem;
- insight as to when various methods may fail: discontinuities for IPA; depdendent support for the likelihood ratio/score function (LR/SF) method (not a problem for weak derivatives (WD));
- the role of representation in estimator unbiasedness and other statistical properties such as variance;
- continuous flow (approximations or exact) models of discrete-event & hybrid systems;
- idea of "phantoms" (Brémaud and Vázquez-Abad 1992) and relationship to SPA and WD methods;
- SPA: degenerated nominal path and perturbed path, most easily seen in (s,S) inventory systems.

Some random reminiscences, milestones, victories, regrets/setbacks:

- The early days of perturbation analysis were tumultuous at times. Due to the early promise of the approach, many claims were made that were later not realizable, at least not by IPA. At some point, many researchers became aware that it simply came down to the mathematical question of exchanging derivative (limit) and expectation (integral) operations (Cao 1985), but it was more of a practical matter to determine how this could be reasonably checked for some classes of systems. The IBM technical report (Heidelberger 1986) had some interesting and valid points (counterexamples) but also generated some bad feelings (bruised egos?) that were eventually patched up and culminated in the 1988 paper that appeared in *Management Science*, which also received the Outstanding Publication Award (by the TIMS College on Simulation, predecessor to the INFORMS Simulation Society) in 1990. The paper by Suri and Zazanis (1988) contained the first analytical proof of strong consistency of an IPA estimator for a well-known queueing system, specifically the M/G/1 queue, but it took about four years to make it through the editorial process. The work by Paul Glasserman in his PhD dissertation (published in Glasserman (1991a), Glasserman (1991b), Glasserman (1991c)) further cemented a rigorous foundation for IPA by capturing the intuition behind discontinuities due to system structure through the introduction of the elegant "commuting condition" in the generalized semi-Markov process (GSMP) framework.

- The "failure" of IPA to be a panacea resulted in what the first author referred to as PA "alphabet soup" beginning in the late 1980s into the next decade, including smoothed perturbation analysis (SPA), finite perturbation analysis (FPA), augmented infinitesimal perturbation analysis (APA), rare perturbation analysis (RPA), discontinuous perturbation analysis (DPA), keeping dedicated researchers up many a late night. Perhaps the most significant of these developments was the introduction of SPA in Gong and Ho (1987), which built upon work of Zazanis and Suri (1986) (finally published much later as Zazanis and Suri (1994)) for the single-server queue, by proposing the idea of using conditioning (conditional Monte Carlo) as a general approach to "smooth" a discontinuous sample performance. This was formalized in the GSMP framework in Glasserman and Gong (1990), Fu and Hu (1992), Fu and Hu (1997).
- In 1988, out of the blue Rajan Suri received an (unsolicited) Research Award of $150,000 from Ford Motor Company "in recognition of outstanding contributions made to the field of Perturbation Analysis of Discrete Event Systems." The award was sponsored by Ford Motors Research Division as recognition that this approach had potential for their industry.
- One interviewee recalled a 1992 IIASA conference in Austria, with Larry Ho, Georg Pflug, and Reuven Rubinstein, all in the same room speaking about perturbation analysis, the method of weak derivatives, and the score function method, respectively.
- Another controversy concerned the name of the other "competing" technique: the likelihood ratio or score function method. Reuven Rubinstein was very adamant about using the term he coined, which was the score function method, and chafed when he heard the other term used. Both terms are in current use.
- From the abstract of the paper by Aleksandrov, Sysoyev, and Shemeneva (1968): "Optimization of the parameters of a system subject to random disturbances is considered. It is shown that when the Monte Carlo method is used to determine the average performance criterion of a system, information is simultaneously obtained about the gradient of the average performance criterion with respect to the optimal parameters and about the matrix of second derivatives. Thus, after performing a series of samples for determining the average performance criterion, we can find the direction of system parameter variation such that the criterion increases the fastest. This result is applied to optimization of the parameters..." Turns out they have proposed exactly the same LR/SF estimator later rediscovered by our community nearly two decades later.
- The first author recalls his PhD student days drawing "hump" diagrams all the time on the blackboard (especially with fellow PhD students Paul Glasserman, Weibo Gong, Jian-Qiang Hu, Shu Li, Pirooz Vakili, and Leyuan Shi) – these are sample paths of discrete-event systems, since they are piecewise constant with jumps at discrete time points upon occurrences of events; a typical illustration of this is provided in Suri (1989) for a single-server queue.
- Whereas IPA faced the hurdle of general applicability, the LR/SF method faced a separate challenge of high variance that happened to be exacerbated by the nature of the application domain in which it was first applied – queueing systems where the parameter appears in a common distribution. In the case of the limitations of IPA, as one of the co-authors often notes and mentioned earlier, there is always a silver lining in research: the limitations of IPA meant that there were immense research opportunities to develop other techniques (e.g., SPA).

The 1990s saw the publication of many research monographs focused on gradient estimation techniques with an eye towards simulation optimization, beginning in 1991 with *Perturbation Analysis of Discrete Event Dynamic Systems* (Ho and Cao 1991) and *Gradient Estimation Via Perturbation Analysis* (Glasserman 1991a), followed shortly by *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method* (Rubinstein and Shapiro 1993) and *Realization Probabilities* (Cao 1994), *Optimization of Stochastic Models: The Interface Between Simulation and Optimization* (Pflug 1996) and *Conditional Monte Carlo: Gradient Estimation and Optimization Applications* (Fu and Hu 1997).

In terms of applications, gradient estimation started with queueing systems, motivated by manufacturing systems, as Ho and Cao (1983), Ho, Cao, and Cassandras (1983) introduced the first IPA for queueing networks, including optimization. Much of Rajan Suri's efforts in the 1980s were spent on extending this reach to computer/communication networks. The first application areas beyond queueing were inventory and finance, both of which arguably have found more success than in the original queueing domain, which may be more a reflection of the application domain. In terms of inventory control applications, the first work in the area was carried out in the mid-1990s: Fu (1994) for (s,S) inventory policies, published in *Operations Research*, and Glasserman and Tayur (1995) for base-stock inventory policies, published in *Management Science*. Another interesting contrast is that clearly the focus in these two papers is on simulation optimization, because the parameters control the ordering policy, and hence were clearly what are now referred to as "structural" parameters, similar to the buffer sizes in the original perturbation analysis papers. However, later on, in the queueing domain, the focus shifted to distributional parameters (e.g., interarrival and service time distribution parameters). Both of the two examples mentioned in the introduction were inventory control applications, with the Caterpillar supply chain case building upon the work in Glasserman and Tayur (1995), and the Kroger drug inventory example related to Fu and Healy (1997).

Application to finance was also begun in the mid-1990s, with the first works being Fu and Hu (1995), Broadie and Glasserman (1996), applying perturbation analysis (where IPA was called the "pathwise method" in the latter work) and the likelihood ratio method to estimating the "Greeks" in options pricing. Later, Malliavin calculus and the method of weak derivatives were also applied to the same. The Lanchester Prize-winning book by Glasserman (2004) highlights the use of such sensitivity estimates, which are central to hedging financial derivatives.

One interviewee lamented the extraordinary time and effort spent on gradient estimation, and judging by the number of papers spent on the area just on the single-server queue, there is definitely an argument to be made there, but one of the co-authors is guilty of being part of that, and he has to admit it was a lot of fun at the time, and still is even today. Another interviewee expressed regret that he did not implement the gradient estimation techniques in software, a feeling shared by many researchers in the field. In fact, gradient estimation is still a very active area of research, with some of the most recent work including the application of MVD to quantile estimation (Heidergott and Volk-Makarewicz 2016), a generalization of both the IPA and LR method to handle structural parameters in discontinuous sample performance functions (Peng, Fu, and Hu 2016a; Peng, Fu, and Hu 2016b; Peng, Fu, Hu, and Heidergott 2017), and application of IPA to stochastic fluid models and hybrid systems (Cassandras and Wardi 2013).

5 RESPONSE SURFACE METHODOLOGY

A central idea in response surface methods is to develop and use a simplified model of the objective function that can be evaluated separately from, and with less computation than, the simulation model. This simplified model is best viewed as an approximation for the true objective function. Such models can be local, in that they only apply in a neighborhood of a current point, or they can be global, in that they apply over the entire search space. Barton and Meckesheimer (2006) and Kleijnen (2015) introduce these methods and their literature, which is extensive. The methods trace their origins at least to Box and Wilson (1951) and Box (1954).

The WSC paper by Mihram (1970) proposed a response surface method for simulation optimization. Biles (1973) is the first WSC paper that actually developed algorithms using response-surface methodology. This paper illustrated and compared a few different algorithms for simulation optimization. The follow-up paper (Biles 1974) promotes an algorithm that iterates between two steps: first an experimental design is used to determine an approximation for the gradient, and, second, a separate experimental design is used to perform a line search in the negative direction of the estimated gradient. The 1974 paper was the one most commonly figured amongst our interviewees as the earliest explicit instance of "simulation optimization" they could recall. (Of course, optimization of noisily observed functions, not explicitly centered on simulation, dates back at least to the work of Box and Wilson 1951 and Robbins and Monro 1951).

A modern incarnation of methods based on local metamodeling is the family of trust-region methods. The first of these of which we are aware is Deng and Ferris (2006), which was followed by STRONG algorithms as developed in Chang, Hong, and Wan (2007), Chang and Wan (2009), and, most recently, ASTRO-DF (Shashaani, Hunter, and Pasupathy 2016).

The methods described to date are primarily built on *local* metamodels. Global metamodeling techniques explored for use with simulation include methods like radial basis functions and regression, but the method of Kriging has seen the most use. Kriging was originally developed for deterministic optimization problems. The central idea is to model the function one wants to optimize as a random realization from a Gaussian random field. One can then predict the value of the function at any desired point through its conditional distribution, conditional on the observations already obtained.

Perhaps the pioneer in bringing Kriging to simulation was Jack Kleijnen. He invited Mitchell and Morris (1992) to the 1992 Winter Simulation Conference to present on Kriging. Articles on the use of Kriging in simulation followed, but not until much later (van Beers and Kleijnen 2003; Huang et al. 2006). The first papers that applied Kriging to simulation mostly ignored the fact that simulation does not provide exact function values. Some work mentioned the issue and appealed to existing results for allowing for measurement error in Kriging that represents a partial, but not ideal, adaptation to account for simulation noise. Ankenman, Nelson, and Staum (2008), Ankenman, Nelson, and Staum (2010) later showed how to explicitly model the simulation noise in Kriging models. Innovations in the use of Kriging for simulation optimization continue, e.g., Scott, Frazier, and Powell (2011), Chen, Ankenman, and Nelson (2013), Qu and Fu (2014), where the latter two combine the use of stochastic kriging with direct gradient estimates.

6 SAMPLE-AVERAGE APPROXIMATION

Sample average approximation (SAA) is not an algorithm, but rather a principle through which algorithms can be designed. The central problem addressed is to $\min_x f(x)$, where $f(x) = Ef(x;\xi)$. Here $f(x;\xi)$ is real valued, x is a decision variable, and ξ is a random quantity. (Functionals other than expected values can be handled seamlessly with SAA.) The principle behind SAA is to first generate a sample $(\xi_1, \xi_2, \ldots, \xi_n)$ where each ξ_i is distributed according to ξ, and then to fix that sample throughout, and select X_n to minimize $n^{-1}\sum_{i=1}^n f(x,\xi_i)$. In other words, we fix the sample, and then solve the sample equivalent of the original optimization problem. This principle does not say *how* to solve the sampled optimization problem, just that one should solve it. The theory of SAA elucidates the qualities of the resulting solution X_n, which is a random variable, because it is a function of the sample itself.

The origins of SAA are difficult to pin down, partly because it is closely related to the principle of maximum likelihood for estimating parameters in statistical models, partly because it is related to the more general principle of M-estimation, e.g., Serfling (1980), and partly because the idea is very natural, so that it may have been used in an informal way for a long time before being formalized. Alex Shapiro, a key figure in the development of the method, said that the idea may even be as old as Monte Carlo itself. For example, the same general principle is described through the names "sample path optimization" by Gürkan, Özge, and Robinson (1994), Plambeck et al. (1996), and Robinson (1996), and by retrospective optimization (or approximation) by Healy and Schruben (1991), Chen and Schmeiser (1994), Chen and Schmeiser (2001), Pasupathy and Schmeiser (2004).

SAA was popularized and significantly developed by Shapiro through many articles. Important surveys include Shapiro (2003), Shapiro, Dentcheva, and Ruszczyński (2009), and Kim, Pasupathy, and Henderson (2015) is an accessible introduction. The name SAA was coined in Kleywegt, Shapiro, and Homem-de-Mello (2001). It is closely related to, but significantly more general than, the score-function method developed in Rubinstein and Shapiro (1990). Reuven Rubinstein, one of the co-authors of this work, was a key figure in the development of several approaches to simulation optimization, including the cross-entropy method. Indeed, in an interview, Shapiro credited Reuven Rubinstein as an important source of ideas in simulation optimization. Reuven's contribution lay in identifying exciting and interesting research ideas, and he successfully partnered with top researchers to pursue those ideas.

In our interview, Shapiro advanced his perspective that the primary successes of SAA lie in "static" problems, where a single (in time) decision must be made. In contrast, SAA has had less impact in multistage stochastic programming problems, a particular subclass of problems where decisions must be made sequentially in time. More generally, he believes that successfully solving dynamic problems, like those currently tackled by methods such as approximate dynamic programming, e.g., Powell (2011), remains an important and challenging goal. On a personal note, one of us (Henderson) was excited upon first learning about SAA, because it spoke to a desire to exploit structure in simulation optimization problems.

We believe that SAA has had much more impact in stochastic linear programming than in more general simulation optimization problems. Why? Stochastic linear programs automatically possess convexity properties that make the sampled problem relatively tractable. This is not automatically true of more general simulation optimization problems, as discussed at length in Kim, Pasupathy, and Henderson (2015).

In our interview with Shapiro, he contrasted sample-average approximation with stochastic approximation, and expressed admiration for the insights of Arkadii Nemirovski that have helped to demonstrate the power and scalability of careful instantiations of stochastic approximation for high-dimensional optimization problems. Some of these insights can be seen in Nemirovski et al. (2009).

We remain convinced that SAA will play a major role in further developments in simulation optimization, perhaps through the teaching of a disciplined approach to modeling that retains tractability (convexity/differentiability) as long as possible before sacrificing those qualities for model fidelity.

7 SELECTION METHODS

A broad variety of methods are available for determining which of a finite number of solutions is the best, when the number of solutions is small enough that *all* solutions can be simulated *at least to some degree.* The methods vary in terms of the guarantees, if any, that they return on the selected solution, and the manner in which they determine the sample size for each solution.

Ranking and Selection methods are designed in such a way that they provide a statistical guarantee on the result that is returned. Bechhofer (1954) and Gupta (1956) are usually credited as the originating work for this suite of techniques. Of course, with any "new" idea, these articles were influenced by earlier papers. For example, Bechhofer cites Paulson (1949). Paulson expressed dissatisfaction with hypothesis tests that try to determine whether all solutions have the same mean, or not. When this hypothesis is rejected, what then? This same motivation appears at the outset of Bechhofer's paper. Paulson went further in suggesting a few different types of statistical guarantee that might be sought, although he did not show how to ensure them through practical sampling rules, and perhaps this is why Bechhofer and Gupta's works are considered the originating papers; they provided the first provable guarantees.

Ranking and selection methods have been intensively studied. We do not attempt to present a survey here, instead referring readers to the excellent introduction Kim and Nelson (2006), and to the book-length treatments Gupta and Panchapakesan (1979) and Bechhofer, Santner, and Goldsman (1995).

Ranking and selection methods were identified as potentially being useful in simulation optimization very early. Indeed, Conway (1963) mentioned the potential of Bechhofer (1954) in this sense, and both Conway, Johnson, and Maxwell (1959) and Burdick and Naylor (1966) reinforced the point that ANOVA was inappropriate in a simulation context where one *knows* that solutions have different objective values, and that what is more important is identifying strong designs.

However, it was not until much later that these methods came to be used widely in simulation. Barry Nelson and David Goldsman were leaders in the effort to translate these ideas into simulation usage. A key hurdle was the modest limit on the number of solutions k for which these methods could still deliver guarantees within a reasonable sampling budget. Values of $k \leq 10$ were typical, but insufficient for simulation applications. An important step forward was the advent of efficient screening, whereby clearly inferior solutions could be quickly eliminated from consideration (Nelson et al. 2001). A key idea there is to allocate part of the probability of failure to the possibility that the best solution is screened out in a screening phase, and to follow a screening phase by a selection over the surviving solutions. These screening procedures allow one to exploit a literature on *sequential* selection procedures that began in the 1960s.

In our, admittedly biased, view, there have been three central developments in this sphere in the last 15 years.

First, Boesel, Nelson, and Kim (2003) advocated the use of ranking and selection algorithms to "clean up" after search algorithms have visited many potential solutions in an attempt to identify an optimal solution. Clean up offers a statistical guarantee that the solution reported as best upon conclusion is indeed the best of those solutions that were visited within the search. It does not make a guarantee relative to solutions that were not visited during the search. Clean up prevents many common errors of misinterpretation on the part of users of simulation software, and has become an integral part of high-quality simulation optimization software. This paper won the INFORMS Simulation Society's best publication award in 2006.

Second, parallel implementations of ranking and selection procedures are now available, and have been designed in such a way as to scale to huge numbers of solutions; see Luo and Hong (2011), (Luo et al. 2015; Ni et al. 2017) for the most prominent examples of this work. This work has helped to advance a perspective that parallel computing should be our default mindset as we design and implement algorithms. One of us (Henderson) was inspired to adopt this perspective during a workshop organized by Michael Fu and Barry Nelson immediately prior to the 2010 Winter Simulation Conference.

Third, building on ideas introduced by Chen (1994) and Dai (1996) in ordinal optimization and simulation budget allocation discussed later in this section, Glynn and Juneja (2004) ushered in a stream of work that views selection problems from the perspective of large-deviations theory. This work has led to new algorithms, e.g., Hunter and Pasupathy (2013), but it also provides a cautionary note about how successful selection algorithms can be when simulation outputs do not satisfy regularity conditions such as being normally distributed, or having sufficiently light tails; see Glynn and Juneja (2015).

While the term "ranking and selection" came directly to the simulation community from statistics, the theoretical computer science community has more recently, in parallel, been investigating "best arm" algorithms for multi-armed bandit problems (Bubeck and Cesa-Bianchi 2012). Most of this literature focuses on a "regret" formulation, in which one accumulates rewards from sampling systems online, i.e., as the exploration (sampling) proceeds. In contrast, the "pure exploratory" bandit problem (Mannor and Tsitsiklis 2004) has the same goal as selection algorithms, namely to select the best from a finite number of noisily-observed solutions *after* a period of exploration. Best-arm algorithms work under different regularity assumptions to the ranking-and-selection community, often assuming bounded outputs, or sub-Gaussian outputs (with an explicit bound on the tail decay rates of the random variables). There seems to have been little cross-fertilization between the best-arm community and those working in ranking and selection, which is unfortunate given that both streams of work trace their origins to Bechhofer (1954). Some researchers are now starting to explore the interface.

Ranking and selection methods tend to be quite conservative in the statistical guarantees that they provide, in that they usually exceed the desired level of accuracy by a wide margin. This conservativeness arises from the mathematical techniques used to ensure the guarantee will hold, and while not intended, may be unavoidable. The conservativeness has the effect of leading to larger sample sizes than are strictly necessary to meet the statistical guarantees. Other selection methods have arisen that place efficiency over statistical guarantees, in that they do not offer an ironclad guarantee, but they are designed to quickly (with small sample sizes) identify high-quality solutions.

Perhaps the most prominent of these methods is the optimal computing budget allocation (OCBA). This method sequentially selects sample sizes so as to maximize an approximation to a statistical guarantee, interpreted in a certain Bayesian sense. The method originated in the PhD dissertation of Chen (1994), presented at WSC by Chen and Dai (1996) and Chen, Chen, Dai, and Yücesan (1997), and has since grown into a large collection of algorithms for different variants of the selection problem; see Chen and Lee (2010).

Other selection methods have been developed that have various efficiency goals. For example, the knowledge gradient (Frazier 2009) works within a Bayesian structure, and at each stage selects a single solution, in a greedy fashion, to next sample. It selects that solution to sample that would maximally increase the expected quality of the final selected solution, assuming that one will stop and select a solution after the next sample is obtained. The central principle in this method is closely related to that of the expected value of information approach (Chick and Inoue 2001) that is also reviewed in Chick (2006). These Bayesian methods that work with a loss function are, at their core, applications of decision science to the selection problem; a very early reference in this vein is Bross (1950), which refers to the author's PhD thesis. An overview of these methods can be found in Chen, Chick, and Lee (2015).

When the number of solutions is finite but so large that not all systems can be simulated, a different suite of techniques are appropriate. *Search* now becomes an essential part of the optimization procedure, in that some method for sequentially selecting new solutions to explore is needed. In this sphere, the dominant methods are random-search algorithms and metaheuristics as surveyed in, e.g., Andradóttir (2006), Andradóttir (2015), Ólafsson (2006). Variations of these methods have been adopted by commercial simulation software, perhaps due to the ease with which they can be applied to a variety of problem structures. Some effort has been made to develop hybrid methods that combine some elements of random search with the clean-up ideas mentioned earlier (Hong and Nelson 2009). Section 8 discusses the use of metaheuristics.

The ideas in ordinal optimization originated with the work of Larry Ho and his lab at Harvard in the 1980s; see Ho, Zhao, and Jia (2008), Chen, Jia, and Lee (2013) for surveys, and Ho, Cao, and Cassandras (1983). The goal in that early work was to develop search algorithms for optimizing discrete-event systems. There are, perhaps, two key ideas in this work. The first is that it is easier to determine *which* of two solutions is the better solution than it is to determine *how much* better one solution is than the other. This idea can be formalized through Chernoff bounds applied in the case of light-tailed random responses (Dai 1996). The second is that of "goal softening," meaning that one might seek a solution that is *close* to optimal rather than optimal. The second idea has parallels in the "probably approximately correct" goal in best-arm algorithms, and equivalently in the "good selection" goal in ranking and selection. These two ideas continue to heavily influence simulation optimization work today (Chen, Jia, and Lee 2013).

8 SOFTWARE

Although simulation optimization has been around for arguably the lifetime of the WSC, software has lagged research results. The specialized nature of many algorithms has made it either too difficult, or commercially unattractive, to implement them in software designed for general modeling. Moreover, it can be difficult to interface the algorithms with the model execution code. As a result, commercial software vendors have tended to shy away from algorithms that require some knowledge of the inner workings of the model, instead focusing almost exclusively on "black box" algorithms based on heuristics from deterministic optimization. Perhaps the most prominent of these black-box algorithms are scatter search and tabu search as used in OptQuest (OptTek 2017) and evolutionary algorithms as used in SimRunner/ProModel Optimization Suite (ProModel 2017) and ExtendSim Optimizer (ExtendSim 2017). Ranking & selection techniques are also amenable to implementation in general-purpose packages, e.g., KN (Kim and Nelson 2001) is implemented in Simio (2017). Due to the commercial nature of these software packages, details of specific algorithms are proprietary. We will describe instead some historical developments of note related to software developments.

OptQuest is perhaps the most widely available simulation optimization package, possibly because the developers worked directly with simulation software developers on the interface between OptQuest and the simulation software. OptQuest was the first and core commercial product of OptTek Systems, which was founded in 1992 (initially under the name "Optimization Technologies," which was shortened three years later) as a result of interactions with the simulation and consulting firm Decisioneering. Those interactions led the three founders – University of Colorado Business Professors at the time, Jim Kelly, Manuel Laguna and Fred Glover – to conclude in a note we received directly from them, "that this 'simulation optimization' area was susceptible to being treated by innovations they had introduced in the field of metaheuristics — higher level methods for solving problems that could not be approached by traditional methods." Asked why he thought it took so long for simulation optimization to become embedded in software, CEO Jim Kelly replied that he believed that the focus of most simulation users was on modeling and analysis rather than decision making, so the software developers at the time reflected the preferences of the users. This in some ways parallels the development of descriptive and predictive analytics versus prescriptive analytics.

At a 2000 WSC panel discussion (Fu et al. 2000) "Integrating Optimization and Simulation: Research and Practice," the panel and audience asked for ways that the gap between research and practice might be bridged, and one of the panelists suggested that perhaps the simulation community should build a testbed of simulation optimization problems on which algorithms could be tested, along the lines of one used by the combinatorial optimization community. This need was reinforced in the paper by Fu (2002) and in a commentary on that paper (Glynn 2002). Shane Henderson and Raghu Pasupathy introduced such a testbed in 2006, which is first described in their WSC proceedings paper (Pasupathy and Henderson 2006). A followup WSC paper (Pasupathy and Henderson 2011) described advances in standardizing the interface with the testbed problems, as well as an extension to also include algorithms. SimOpt focuses attention on the empirical performance of algorithms for problems that range in scale from simple single-variable problems to problems with hundreds of decision variables.

Many researchers have lamented the lack of implementation of gradient estimation techniques, and hence gradient-based search, into widely available software, but there have been recent successful efforts to incorporate other state-of-the-art simulation optimization advances into software, including the implementation of R&S in commercial software packages such as Simio as mentioned earlier. Another (non-commercial) example (Hong and Nelson 2009; Xu, Nelson, and Hong 2010) is Industrial Strength COMPASS (ISC), which is "open source code for maximizing or minimizing the expected value of a single performance measure generated by a stochastic simulation with respect to integer ordered decision variables subject to linear-integer constraints." ISC is based on an enhanced version of COMPASS, work clearly coming out of the simulation research community, first published in Hong and Nelson (2006). ISC has shown itself to be competitive with OptQuest in empirical testing, while at the same time providing statistical guarantees as to the quality of the found solution(s).

9 REFLECTIONS

In reflecting on the history of simulation optimization, several themes seem, to us, to stand out.

Various subdisciplines within simulation optimization appear to have evolved somewhat independently. Although various research groups knew about each other's work, there appears to have been only modest cross fertilization. Perhaps this is a function of the origins of various ideas. For example, stochastic approximation and gradient estimation emerged, to a large extent, from systems and control; ranking & selection and response-surface methodology primarily emerged from the statistics and design-of-experiments communities; and sample-average approximation emerged, mostly, from the math programming community. More recently, computer science is contributing through analyses of stochastic gradient descent methods, multi-armed bandit formulations and reinforcement learning and approximate dynamic programming.

In this same vein, commercial software has developed almost completely independently from the work of academic researchers. This may be partly because software vendors have needed to provide a general-purpose simulation optimization capability that handles the lack of structure present in models built by users who model first and think about optimization only later. Jim Kelly emphasized this latter point to us, in that simulation classes teach how to model, and tend to ignore the tractability of the resulting simulation optimization problems. Most academics, true to the incentive structures within which they work, have attempted to work on problems where mathematical analysis plays a large role. This has usually limited their efforts to well-structured problems, or to problems of a modest scale that do not necessarily reflect practice. Of course, the two prominent examples discussed in the introduction demonstrate the practical value of this basic research. Moreover, response-surface methodology has been used to great effect in practical problems. There are positive indications that the crossover is increasing, as discussed in §8 above.

We offer the following suggestions for future work, based on what we have learned herein.

- The plethora of simulation-optimization algorithms that are available now perhaps reflects compartmentalized development. Careful empirical comparisons of algorithms could help cross-fertilize between these different schools of thought, which is one of the central goals of SimOpt (Henderson and Pasupathy 2017).
- The fundamental work on which many academics engage should be complemented by work on realistic, large-scale problem instances, to assist with cross-fertilization between academia and practice.
- Academics should make additional efforts to reach out to commercial simulation software developers to learn what is needed in practice, and to explore potential joint work. Academics might also explore developing and sharing (free) software that showcases their algorithms. SimOpt (Henderson and Pasupathy 2017) is a positive development in this regard.
- Users might be taught how to build models with optimization in mind. Rather than having a single model to capture a given situation, a collection of models could be employed. Simple models that yield tractable simulation optimization problems might help with "global search," providing initial solutions to local search algorithms built over more granular simulation models.
- We see great potential for simulation optimization to leverage the large-scale parallel computing facilities that are now widely accessible.
- Recent developments in distributionally robust optimization that are emerging from the deterministic optimization community could have large impact in simulation optimization applications; see Bayraksan's comments in Fu et al. (2014) and Lam (2016).

The overall impression we obtain in reflecting on the progress to date on simulation optimization is that the field has had great impact, and is in an excellent position to further increase that impact over the coming years.

10 TIMELINE

1951 Box and Wilson (1951) on design of experiments.
1951 Robbins and Monro (1951) on stochastic root finding.
1952 Kiefer and Wolfowitz (1952) on finding the maximum of a regression function.
1954 Box (1954) on response surfaces.
1954 Bechhofer (1954) on ranking and selection.
1956 Gupta (1956) on ranking and selection.
Late 1950s Kushner explores Robbins-Monro work at Lincoln Labs at the suggestion of Robert Sitler.
1959 Conway, Johnson, and Maxwell (1959) discuss simulation optimization in a section entitled "Problems of Experimental Design," saying "For this problem, there exists a considerable literature, some controversy, and an increasing amount of interest."
1963 Conway (1963) updates Conway, Johnson, and Maxwell (1959), discusses the use of common random numbers, and again highlights the limitations of ANOVA, pointing instead to ranking and selection as a way forward.
1964 Hammersley and Handscomb (1964) does not appear to mention simulation optimization.
1966 Burdick and Naylor (1966) reviews several response surface ideas.
1968 Schriber introduces a simulation elective for MBAs at the University of Michigan.
1973 Biles (1973, 1974) plants simulation optimization seeds at WSC.

1975 Kleijnen (1975) includes early references on the use of response-surface methodology in simulation.
1977 Simulation optimization increases in prominence at WSC, with articles by, e.g., Jim Swain and Dennis Pegden.
1979 Ho, Eyler, and Chien (1979) generally regarded as the birth of perturbation analysis (PA).
1981 Schruben and Cogliano (1981) introduces idea of frequency domain experimentation (at WSC).
1986+ The likelihood ratio/score function method for gradient estimation is introduced and formalized by Glynn (1986a), Glynn (1986b), Glynn (1987), Reiman and Weiss (1989), Rubinstein (1986), apparently independently of its earlier introduction in Aleksandrov, Sysoyev, and Shemeneva (1968).
1987 Gong and Ho (1987) introduces smoothed perturbation analysis (SPA).
1989 Pflug (1989) introduces the idea of weak derivatives for gradient estimation.
1988-92 Ruppert (1988) and Polyak and Juditsky (1992) independently introduce the idea of iterate averaging in stochastic approximation.
1991 Healy and Schruben (1991) introduces "retrospective optimization," one origin of SAA in the WSC community; also introduced as "sample-path optimization" in Gürkan, Özge, and Robinson (1994).
1992 Spall (1992) introduces SPSA.
1992 Ho, Sreenivas, and Vakili (1992) introduces ordinal optimization.
1992 Optimization Technologies Inc. (later shortened to OptTek) founded, introducing OptQuest.
1994/95 Chen (1994, 1995) introduces OCBA.
1997/1999 The cross-entropy method is introduced; originally designed for rare event simulation (Rubinstein 1997), later adapted for (deterministic) combinatorial optimization (Rubinstein 1999).
2001 Kleywegt, Shapiro, and Homem-de-Mello (2001) coins the term "SAA."
2006 Pasupathy and Henderson (2006) introduces SimOpt.
2007 Hu, Fu, and Marcus (2007) introduces Model Reference Adaptive Search.
2007-2009 Industrial Strength Compass introduced.
2011 Simulation optimization becomes a separate track from analysis methodology at WSC.

ACKNOWLEDGMENTS

The authors are grateful for the time that the following persons spent with them, in person, on the phone, and via e-mail exchanges: Russell Barton, Fred Glover, Bernd Heidergott, Jim Henriksen, Yu-Chi (Larry) Ho, Jim Kelly, Jack Kleijnen, Harold Kushner, Barry Nelson, Dennis Pegden, Bruce Schmeiser, Lee Schruben, Tom Schriber, Alex Shapiro, Rajan Suri, Jim Swain, Michael Zazanis.

This work was partially supported by National Science Foundation grants CMMI-1537394, CMMI-1362303, and CMMI-1434419, by Air Force Office of Scientific Research (AFOSR) Grant FA9550-15-10050, and Army Research Office W911NF-17-1-0094.

REFERENCES

Aleksandrov, V. M., V. I. Sysoyev, and V. V. Shemeneva. 1968. ''Stochastic Optimization''. *Engineering Cybernetics* 5:11–16.

Alrefaei, M. H., and S. Andradóttir. 1999. ''A Simulated Annealing Algorithm with Constant Temperature for Discrete Stochastic Optimization''. *Management Science* 45 (5): 748–764.

Andradóttir, S. 1990. ''A New Algorithm for Stochastic Optimization''. In *Proceedings of the 1990 Winter Simulation Conference*, edited by O. Balci, R. P. Sadowski, and R. E. Nance, 364–366. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Andradóttir, S. 1991. ''A Projected Stochastic Approximation Algorithm''. In *Proceedings of the 1991 Winter Simulation Conference*, edited by B. L. Nelson, G. M. Clark, and W. D. Kelton, 954–957. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Andradóttir, S. 1995. ‘‘A Method for Discrete Stochastic Optimization”. *Management Science* 44:1946–1961.

Andradóttir, S. 2006. ‘‘An Overview of Simulation Optimization via Random Search”. In *Simulation*, edited by S. G. Henderson and B. L. Nelson, Handbooks in Operations Research and Management Science, 617–631. Amsterdam: Elsevier.

Andradóttir, S. 2015. ‘‘A Review of Random Search Methods”. In *Handbook of Simulation Optimization*, edited by M. C. Fu, Volume 216 of *International Series in Operations Research & Management Science*, Chapter 10, 277–292. New York: Springer.

Ankenman, B., B. L. Nelson, and J. Staum. 2008. ‘‘Stochastic Kriging for Simulation Metamodeling”. In *Proceedings of the 2008 Winter Simulation Conference*, edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, and J. W. Fowler, 362–370. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Ankenman, B., B. L. Nelson, and J. Staum. 2010. ‘‘Stochastic Kriging for Simulation Metamodeling”. *Operations Research* 58:371–382.

Banks, J., J. S. Carson II, B. L. Nelson, and D. M. Nicol. 2010. *Discrete-Event System Simulation*. 5th ed. Upper Saddle River, NJ: Prentice-Hall.

Barton, R. R., and J. S. Ivey Jr.. 1996. ‘‘Nelder-Mead Simplex Modifications for Simulation Optimization”. *Management Science* 42:954–973.

Barton, R. R., and M. Meckesheimer. 2006. ‘‘Metamodel-Based Simulation Optimization”. In *Simulation*, edited by S. G. Henderson and B. L. Nelson, Handbooks in Operations Research and Management Science, 535–574. Elsevier.

Bechhofer, R. E. 1954. ‘‘A Single-Sample Multiple-Decision Procedure for Ranking Means of Normal Populations with Known Variances”. *Annals of Mathematical Statistics* 25:16–39.

Bechhofer, R. E., T. J. Santner, and D. Goldsman. 1995. *Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons*. New York: Wiley.

Biles, W. E. 1973. ‘‘Constrained Sequential-Block Search in Simulation Experimentation”. In *Proceedings of the 1973 Winter Simulation Conference*, edited by J. Sussman and A. C. Hoggatt, 227–241. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Biles, W. E. 1974. ‘‘A Gradient-Recession Search Procedure for Simulation Experimentation”. In *Proceedings of the 1974 Winter Simulation Conference*, edited by M. F. Morris, H. Steinberg, and H. J. Highland, 491–497. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Boesel, J., B. L. Nelson, and S. Kim. 2003. ‘‘Using Ranking and Selection to ‘Clean Up’ After Simulation Optimization”. *Operations Research* 51:814–825.

Box, G. E. P. 1954. ‘‘The Exploration and Exploitation of Response Surfaces. Some General Considerations and Examples”. *Biometrics* 10:16–60.

Box, G. E. P., and K. B. Wilson. 1951. ‘‘On the Experimental Attainment of Optimum Conditions”. *Journal of the Royal Statistical Society. Series B* 13:1–45.

Brémaud, P., and F. J. Vázquez-Abad. 1992. ‘‘On the Pathwise Computation of Derivatives with Respect to the Rate of a Point Process: The Phantom RPA Method”. *Queueing Systems: Theory and Applications* 10:249–270.

Broadie, M., and P. Glasserman. 1996. ‘‘Estimating Security Price Derivatives Using Simulation”. *Management Science* 42 (2): 269–285.

Bross, I. 1950. ‘‘Two-Choice Selection”. *Journal of the American Statistical Association* 45 (252): 530–540.

Bubeck, S., and N. Cesa-Bianchi. 2012. ‘‘Regret Analysis of Stochastic and Nonstochastic Multi-Armed Bandit Problems”. *Foundations and Trends in Machine Learning* 5 (1): 1–122.

Burdick, D. S., and T. H. Naylor. 1966. ‘‘Design of Computer Simulation Experiments for Industrial Systems”. *Communications of the ACM* 9 (5): 329–339.

Cao, X. R. 1985. ‘‘Convergence of Parameter Sensitivity Estimates in a Stochastic Experiment”. *IEEE Transactions on Automatic Control* AC-30:845–853.

Cao, X.-R. 1994. *Realization Probabilities: The Dynamics of Queuing Systems*. Springer-Verlag, Boston.
Cassandras, C. G., and Y. Wardi. 2013. ''The IPA Calculus for Hybrid Systems''. In *Stochastic Simulation Optimization for Discrete Event Systems: Perturbation Analysis, Ordinal Optimization And Beyond*, edited by C.-H. Chen and L.-H. Lee, Chapter 1, 3–23. Hackensack, NJ: World Scientific.
Chan, W.-K., and L. W. Schruben. 2003. ''Properties of Discrete Event Systems From Their Mathematical Programming Representations''. In *Proceedings of the 2003 Winter Simulation Conference*, edited by S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, 496–502. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
Chan, W.-K., and L. W. Schruben. 2006. ''Response Gradient Estimation Using Mathematical Programming Models of Discrete-Event System Sample Paths''. In *Proceedings of the 2006 Winter Simulation Conference*, edited by L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 272–278. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
Chan, W.-K., and L. W. Schruben. 2008. ''Optimization Models of Discrete-Event System Dynamics''. *Operations Research* 56 (5): 1218–1237.
Chang, K. H., L. J. Hong, and H. Wan. 2007. ''Stochastic Trust Region Gradient-Free Method (STRONG)-A New Response-Surface-Based Algorithm in Simulation Optimization''. In *Proceedings of the 2007 Winter Simulation Conference*, edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 346–354. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
Chang, K.-H., and H. Wan. 2009. ''Stochastic Trust Region Response Surface Convergent Method for Generally Distributed Response Surface''. In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 563–573. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
Chau, M., and M. C. Fu. 2015. ''An Overview of Stochastic Approximation''. In *Handbook of Simulation Optimization*, edited by M. C. Fu, Volume 216 of *International Series in Operations Research & Management Science*, Chapter 6, 149–178. Springer New York.
Chau, M., M. C. Fu, and H. Qu. 2017. ''Multivariate Stochastic Approximation Using a Secant-Tangents AveRaged (STAR) Gradient''. *Operations Research*:submitted.
Chau, M., H. Qu, and M. C. Fu. 2014. ''A New Hybrid Stochastic Approximation Algorithm''. In *Proceedings of the 12th International Workshop on Discrete Event Systems (WODES)*, 241–246.
Chen, C.-H. 1994. *An Efficient Approach for Discrete Event System Decision Problems*. Ph. D. thesis, Harvard University, MA.
Chen, C.-H. 1995. ''An Effective Approach to Smartly Allocate Computing Budget for Discrete Event Simulation''. In *Proceedings of the 34th IEEE Conference on Decision and Control*, 2598–2605. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
Chen, C.-H., S. E. Chick, and L. H. Lee. 2015. ''Ranking and Selection: Efficient Simulation Budget Allocation''. In *Handbook of Simulation Optimization*, edited by M. C. Fu, Volume 216 of *International Series in Operations Research & Management Science*, Chapter 3, 45–80. New York: Springer.
Chen, C.-H., Q.-S. Jia, and L. H. Lee. 2013. *Stochastic Simulation Optimization for Discrete Event Systems: Perturbation Analysis, Ordinal Optimization and Beyond*. Hackensack, NJ: World Scientific.
Chen, C.-H., and L.-H. Lee. 2010. *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*. Hackensack, NJ: World Scientific.
Chen, H., and B. W. Schmeiser. 1994. ''Retrospective Approximation Algorithms for Stochastic Root Finding''. In *Proceedings of the 1994 Winter Simulation Conference*, edited by J. D. Tew, S. Manivannan, D. A. Sadowski, and A. F. Seila, 255–261. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
Chen, H., and B. W. Schmeiser. 2001. ''Stochastic Root Finding via Retrospective Approximation''. *IIE Transactions* 33:259–275.

Chen, H. C., C. H. Chen, L. Dai, and E. Yücesan. 1997. ''New Development of Optimal Computing Budget Allocation for Discrete Event Simulation''. In *Proceedings of the 1997 Winter Simulation Conference*, edited by S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson, 334–341. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chen, H. C., and L. Dai. 1996. ''A Gradient Approach of Smartly Allocating Computing Budget for Discrete Event Simulation''. In *Proceedings of the 1996 Winter Simulation Conference*, edited by J. M. Charnes, D. J. Morrice, D. T. Brunner, and J. J. Swain, 398–405. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chen, X., B. E. Ankenman, and B. L. Nelson. 2013. ''Enhancing Stochastic Kriging Metamodels with Gradient Estimators''. *Operations Research* 61 (2): 512–528.

Chick, S. E. 2006. ''Subjective Probability and Bayesian Methodology''. In *Simulation*, edited by S. G. Henderson and B. L. Nelson, Handbooks in Operations Research and Management Science. Amsterdam: Elsevier.

Chick, S. E., and K. Inoue. 2001. ''New Two-Stage and Sequential Procedures for Selecting the Best Simulated System''. *Operations Research* 49:732–743.

Chong, E. K. P. 2013. ''Infinitesimal Perturbation Analysis and Optimization Algorithms''. In *Stochastic Simulation Optimization for Discrete Event Systems: Perturbation Analysis, Ordinal Optimization And Beyond*, edited by C.-H. Chen and L.-H. Lee, Chapter 5, 75–95. Hackensack, NJ: World Scientific.

Chong, E. K. P., and P. J. Ramadge. 1992. ''Convergence of Recursive Optimization Algorithms using Infinitesimal Perturbation Analysis Estimates''. *Discrete Event Dynamic Systems* 1 (4): 339–372.

Chong, E. K. P., and P. J. Ramadge. 1993. ''Optimization of Queues using an Infinitesimal Perturbation Analysis-Based Stochastic Algorithm with General Update Times''. *SIAM J. Control and Optimization* 31 (3): 698–732.

Conway, R. W. 1963. ''Some Tactical Problems in Digital Simulation''. *Management Science* 10 (1): 47–61.

Conway, R. W., B. M. Johnson, and W. L. Maxwell. 1959. ''Some Problems of Digital Systems Simulation''. *Management Science* 6 (1): 92–110.

Dai, L. 1996. ''Convergence Properties of Ordinal Comparison in the Simulation of Discrete Event Dynamic Systems''. *Journal of Optimization Theory and Applications* 91:363–388.

Deng, G., and M. C. Ferris. 2006. ''Adaptation of the UOBYQA Algorithm for Noisy Functions''. In *Proceedings of the 2006 Winter Simulation Conference*, edited by L. F. Perrone, B. Lawson, J. Liu, and F. Wieland, 312–319. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Emery, J. C. 1969. ''Job Shop Scheduling By Means of Simulation and an Optimum-Seeking Search''. In *Third Conference on the Applications of Simulation*, edited by A. Ockene and P. J. Kiviat, 363–372.

ExtendSim 2017. "ExtendSim Optimizer". https://www.extendsim.com/downloads/papers/ExtendSim_Optimizer.pdf Accessed April 26, 2017.

Fabian, V. 1968. ''On Asymptotic Normality in Stochastic Approximation''. *The Annals of Mathematical Statistics* 39:1327–1332.

Fabian, V. 1971. ''Stochastic Approximation''. In *Optimization Methods in Statistics*, edited by J. S. Rustagi. Academic Press, New York.

Frazier, P. I. 2009. *Knowledge-gradient methods for statistical learning*. Ph. D. thesis, Princeton, NJ.

Fu, M. C. 1990. ''Convergence of a Stochastic Approximation Algorithm for the $GI/G/1$ Queue Using Infinitesimal Perturbation Analysis''. *Journal of Optimization Theory and Applications* 65 (1): 149–160.

Fu, M. C. 1994. ''Optimization via Simulation: A Review''. *Annals of Operations Research* 53:199–247.

Fu, M. C. 2002. ''Optimization for Simulation: Theory Vs. Practice (Feature Article)''. *INFORMS Journal on Computing* 14 (3): 192–215.

Fu, M. C. 2006. ''Gradient Estimation''. In *Handbooks in Operations Research and Management Science: Simulation*, edited by S. G. Henderson and B. L. Nelson, Chapter 19, 575–616. Elsevier.

Fu, M. C. 2015a. *Handbook of Simulation Optimization*. International Series in Operations Research & Management Science. Springer New York.

Fu, M. C. 2015b. ''Stochastic Gradient Estimation''. In *Handbook of Simulation Optimization*, International Series in Operations Research & Management Science, Chapter 5, 109–150. Springer New York.

Fu, M. C., S. Andradóttir, J. S. Carson, F. Glover, C. R. Harrell, Y. C. Ho, J. P. Kelly, and S. M. Robinson. 2000. ''Integrating Optimization and Simulation:research and Practice''. In *Proceedings of the 2000 Winter Simulation Conference*, edited by J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, 610–616. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Fu, M. C., G. Bayraksan, S. G. Henderson, B. L. Nelson, W. B. Powell, I. O. Ryzhov, and B. Thengvall. 2014. ''Simulation Optimization: A Panel on the State of the Art in Research and Practice''. In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, D. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 3696–3706. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Fu, M. C., and K. J. Healy. 1997. ''Techniques for Simulation Optimization: An Experimental Study on an (S,S) Inventory System''. *IIE Transactions* 29 (3): 191–199.

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., and Y. C. Ho. 1988. ''Using Perturbation Analysis for Gradient Estimation, Averaging, and Updating in a Stochastic Approximation Algorithm''. In *Proceedings of the 1988 Winter Simulation Conference*, edited by M. A. Abrams, P. L. Haigh, and J. C. Comfort, 509–517. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Fu, M. C., and J. Q. Hu. 1992. ''Extensions and Generalizations of Smoothed Perturbation Analysis in a Generalized Semi-Markov Process Framework''. *IEEE Transactions on Automatic Control* 37 (10): 1483–1500.

Fu, M. C., and J. Q. Hu. 1995. ''Sensitivity Analysis for Monte Carlo Simulation of Option Pricing''. *Probability in the Engineering and Informational Sciences* 9 (3): 417–446.

Fu, M. C., and J. Q. Hu. 1997. *Conditional Monte Carlo: Gradient Estimation and Optimization Applications*. Boston: Kluwer Academic Publishers.

Garzia, R. F., P. H. Schmidt, and M. R. Garzia. 1980. ''Simulation, Mathematical Modeling and Optimization in Industrial Design: When and How to Apply It?''. In *Proceedings of the 1980 Winter Simulation Conference*, edited by T. I. Ören, C. M. Shub, and P. F. Roth, 325–340. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Ghadimi, S., and G. Lan. 2015. ''Stochastic Approximation Methods and Their Finite-Time Convergence Properties''. In *Handbook of Simulation Optimization*, edited by M. C. Fu, Volume 216 of *International Series in Operations Research & Management Science*, Chapter 7, 179–206. Springer New York.

Glasserman, P. 1991a. *Gradient Estimation via Perturbation Analysis*. Kluwer Academic Publishers, Boston.

Glasserman, P. 1991b. ''Structural Conditions For Perturbation Analysis Derivative Estimation: Finite-Time Performance Indices''. *Operations Research* 39:724–738.

Glasserman, P. 1991c. ''Structural Conditions for Perturbation Analysis of Queueing Systems''. *Journal of the ACM* 38 (4): 1005–1025.

Glasserman, P. 2004. *Monte Carlo Methods in Financial Engineering*. New York: Springer-Verlag.

Glasserman, P., and W. Gong. 1990. ''Smoothed Perturbation Analysis for a Class of Discrete Event Systems''. *IEEE Transactions on Automatic Control* AC-35:1218–1230.

Glasserman, P., and S. Tayur. 1995. ''Sensitivity Analysis for Base-Stock Levels in Multi-Echelon Production-Inventory Systems''. *Management Science* 41:263–281.

Glynn, P. W. 1986a. ''Optimization of Stochastic Systems''. In *Proceedings of the 1986 Winter Simulation Conference*, edited by J. R. Wilson, J. O. Henriksen, and S. D. Roberts, 52–59. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Glynn, P. W. 1986b. ''Stochastic Approximation for Monte Carlo Optimization''. In *Proceedings of the 1986 Winter Simulation Conference*, edited by J. R. Wilson, J. O. Henriksen, and S. D. Roberts, 356–365. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Glynn, P. W. 1987. ''Likelihood Ratio Gradient Estimation: An Overview''. In *Proceedings of the 1987 Winter Simulation Conference*, edited by A. Thesen, H. Grant, and W. D. Kelton, 366–375. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Glynn, P. W. 2002. ''Additional Perspectives on Simulation for Optimization''. *INFORMS Journal on Computing* 14:220–222.

Glynn, P. W., and S. Juneja. 2004. ''A Large Deviations Perspective on Ordinal Optimization''. In *Proceedings of the 2004 Winter Simulation Conference*, edited by R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 577–585. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Glynn, P. W., and S. Juneja. 2015. ''Ordinal Optimization - Empirical Large Deviations Rate Estimators, and Stochastic Multi-Armed Bandits''. Draft available on arXiv:1507.04564.

Gong, W. B., and Y. C. Ho. 1987. ''Smoothed Perturbation Analysis of Discrete-Event Dynamic Systems''. *IEEE Transactions on Automatic Control* AC-32:858–867.

Gong, W. B., Y. C. Ho, and W. Zhai. 1999. ''Stochastic Comparison Algorithm for Discrete Optimization with Estimations''. *SIAM Journal on Optimization* 10 (2): 384–404.

Gupta, S. S. 1956. *On a Decision Rule for a Problem in Ranking Means*. Ph. D. thesis, University of North Carolina, Chapel Hill, NC.

Gupta, S. S., and S. Panchapakesan. 1979. *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations*. New York: Wiley.

Gürkan, G., A. Y. Özge, and S. M. Robinson. 1994. ''Sample-Path Optimization in Simulation''. In *Proceedings of the 1994 Winter Simulation Conference*, edited by J. D. Tew, S. Manivannan, D. A. Sadowski, and A. F. Seila, 247–254. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Hammersley, J. M., and D. C. Handscomb. 1964. *Monte Carlo Methods*. London: Methuen.

Healy, K., and L. W. Schruben. 1991. ''Retrospective Simulation Response Optimization''. In *Proceedings of the 1991 Winter Simulation Conference*, edited by B. L. Nelson, W. D. Kelton, and G. Clark, 901–907. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Heidelberger, P. 1986. ''Limitations of Infinitesimal Perturbation Analysis''. Technical Report IBM Research Report RC 11891, IBM Watson Research Labs, Yorktown Heights, NY.

Heidergott, B., and W. Volk-Makarewicz. 2016. ''A Measure-Valued Differentiation Approach to Sensitivities of Quantiles''. *Mathematics of Operations Research* 41 (1): 293–317.

Heidergott, B. F. 2010. *Max-Plus Linear Stochastic Systems and Perturbation Analysis*. Springer.

Henderson, S. G. and R. Pasupathy 2017. "Simulation Optimization Library". http://www.simopt.org Accessed April 26, 2017.

Ho, Y. C., and X. R. Cao. 1983. ''Perturbation Analysis and Optimization of Queueing Networks''. *Journal of Optimization Theory and Applications* 40:559–582.

Ho, Y. C., and X. R. Cao. 1991. *Perturbation Analysis and Discrete Event Dynamic Systems*. Boston: Kluwer Academic.

Ho, Y. C., X. R. Cao, and C. G. Cassandras. 1983. ''Infinitesimal and Finite Perturbation Analysis for Queueing Networks''. *Automatica* 19:439–445.

Ho, Y. C., M. A. Eyler, and T. T. Chien. 1979. ''A Gradient Technique for General Buffer Storage Design in a Serial Production Line''. *International Journal of Production Research* 17:557–580.

Ho, Y. C., R. S. Sreenivas, and P. Vakili. 1992. ''Ordinal Optimization of DEDS''. *Journal of Discrete Event Dynamic Systems* 2 (2): 61–88.

Ho, Y. C., Q. C. Zhao, and Q. S. Jia. 2008. *Ordinal Optimization: Soft Optimization for Hard Problems*. New York: Springer.

Hong, L. J., and B. L. Nelson. 2006. ''Discrete Optimization via Simulation Using COMPASS''. *Operations Research* 54 (1): 115–129.

Hong, L. J. and B. L. Nelson 2009. "Industrial Strength Compass". http://users.iems.northwestern.edu/~nelsonb/ISC/ Accessed April 26, 2017.

Hong, L. J., B. L. Nelson, and J. Xu. 2015. ''Discrete Optimization via Simulation''. In *Handbook of Simulation Optimization*, edited by M. C. Fu, Volume 216 of *International Series in Operations Research & Management Science*, 9–44. Springer New York.

Hu, J., M. C. Fu, and S. I. Marcus. 2005. ''Simulation Optimization using Model Reference Adaptive Search''. In *Proceedings of the 2005 Winter Simulation Conference*, edited by C. Alexopoulos, K. Kang, W. R. Lilegdon, and D. Goldsman, 811–818. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Hu, J., M. C. Fu, and S. I. Marcus. 2007. ''A Model Reference Adaptive Search Method for Global Optimization''. *Operations Research* 55 (3): 549–568.

Hu, J., and P. Hu. 2010. ''An Approximate Annealing Search Algorithm to Global Optimization and Its Connection to Stochastic Approximation''. In *Proceedings of the 2010 Winter Simulation Conference*, edited by B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, E. Yücesan, and M. Fu, 1223–1234. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Huang, D., T. T. Allen, W. I. Notz, and N. Zheng. 2006. ''Global Optimization of Stochastic Black-Box Systems via Sequential Kriging Meta-Models''. *Journal of Global Optimization* 34 (3): 441–466.

Hunter, S. R., and R. Pasupathy. 2013. ''Optimal Sampling Laws for Stochastically Constrained Simulation Optimization on Finite Sets''. *INFORMS Journal on Computing* 25 (3): 527–542.

Kiefer, J., and J. Wolfowitz. 1952. ''Stochastic Estimation of the Maximum of a Regression Function''. *Annals of Mathematical Statistics* 23:462–266.

Kim, S., R. Pasupathy, and S. G. Henderson. 2015. ''A Guide to Sample Average Approximation''. In *Handbook of Simulation Optimization*, edited by M. C. Fu, Volume 216 of *International Series in Operations Research & Management Science*, Chapter 8, 207–244. Springer New York.

Kim, S.-H., and B. L. Nelson. 2001. ''A Fully Sequential Procedure for Indifference-Zone Selection in Simulation''. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 11 (3): 251–273.

Kim, S.-H., and B. L. Nelson. 2006. ''Selecting the Best System''. In *Simulation*, edited by S. G. Henderson and B. L. Nelson, Handbooks in Operations Research and Management Science, 501–534. Amsterdam: Elsevier.

Kleijnen, J. P. C. 1975. *Statistical Techniques in Simulation*. Dekker.

Kleijnen, J. P. C. 2008. *Design and Analysis of Simulation Experiments*. New York: Springer.

Kleijnen, J. P. C. 2015. ''Response Surface Methodology''. In *Handbook of Simulation Optimization*, edited by M. C. Fu, Volume 216 of *International Series in Operations Research & Management Science*, Chapter 4, 81–104. Springer New York.

Kleywegt, A. J., A. Shapiro, and T. Homem-de-Mello. 2001. ''The Sample Average Approximation Method for Stochastic Discrete Optimization''. *SIAM Journal on Optimization* 12:479–502.

Kushner, H. J., and D. Clark. 1979. *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. Springer-Verlag.

Kushner, H. J., and G. G. Yin. 1997. *Stochastic Approximation Algorithms and Applications*. Springer-Verlag.

Kushner, H. J., and G. G. Yin. 2003. *Stochastic Approximation and Recursive Algorithms and Applications*. 2nd ed. New York: Springer-Verlag.

Lam, H. 2016. ''Advanced Tutorial: Input Uncertainty and Robust Analysis in Stochastic Simulation''. In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 178–192. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Law, A. M. 2015. *Simulation Modeling and Analysis*. 5th ed. New York: McGraw-Hill.

Lefkowitz, R. M., and T. J. Schriber. 1971. ''Use of an External Optimizing Algorithm with a GPSS Model''. In *Proceedings of the 1971 Winter Simulation Conference*, edited by M. Araten and J. Sussman, 162–171. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.s

Luo, J., and L. J. Hong. 2011. ''Large-Scale Ranking and Selection using Cloud Computing''. In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu, 4051–4061. Piscataway, NJ: Institute of Electrical and Electronics Engineers, Inc.

Luo, J., L. J. Hong, B. L. Nelson, and Y. Wu. 2015. ''Fully Sequential Procedures for Large-Scale Ranking-And-Selection Problems in Parallel Computing Environments''. *Operations Research* 63:1177–1194.

Mannor, S., and J. N. Tsitsiklis. 2004, December. ''The Sample Complexity of Exploration in the Multi-Armed Bandit Problem''. *Journal of Machine Learning Research* 5:623–648.

Mihram, G. A. 1970. ''An Efficient Procedure for Locating the Optimal Simular Response''. In *Fourth Conference on the Applications of Simulation*, edited by P. J. Kiviat and M. Araten, 154–161.

Mitchell, T. J., and M. D. Morris. 1992. ''The Spatial Correlation Function Approach to Response Surface Estimation''. In *Proceedings of the 1992 Winter Simulation Conference*, edited by J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson, 565–571. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Nelson, B. L., J. Swann, D. Goldsman, and W. Song. 2001. ''Simple Procedures for Selecting the Best Simulated System When the Number of Alternatives Is Large''. *Operations Research* 49 (6): 950–963.

Nemirovski, A., A. Juditsky, G. Lan, and A. Shapiro. 2009. ''Robust Stochastic Approximation Approach to Stochastic Programming''. *SIAM Journal on Optimization* 19 (4): 1574–1609.

Ni, E. C., D. F. Ciocan, S. G. Henderson, and S. R. Hunter. 2017. ''Efficient Ranking and Selection in High Performance Computing Environments''. *Operations Research*:To appear.

Ólafsson, S. 2006. ''Metaheuristics''. In *Simulation*, edited by S. G. Henderson and B. L. Nelson, Handbooks in Operations Research and Management Science, 633–654. Amsterdam: Elsevier.

OptTek 2017. "OptQuest". http://www.opttek.com/products/optquest/ Accessed April 26, 2017.

Pasupathy, R., and S. G. Henderson. 2006. ''A Testbed of Simulation-Optimization Problems''. In *Proceedings of the 2006 Winter Simulation Conference*, edited by L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 255–263. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Pasupathy, R., and S. G. Henderson. 2011. ''SimOpt : A Library of Simulation Optimization Problems''. In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu, 4080–4090. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Pasupathy, R., and B. W. Schmeiser. 2004. ''Retrospective Approximation Algorithms for the Multidimensional Stochastic Root-Finding Problem.''. In *Proceedings of the 2004 Winter Simulation Conference*, edited by R. Ingalls, M. Rossetti, J. Smith, and B. Peters. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Paulson, E. 1949. ''A Multiple Decision Procedure for Certain Problems in the Analysis of Variance''. *Annals of Mathematical Statistics* 20:95–98.

Peng, Y., M. Fu, and J. Hu. 2016a. ''Estimating Distribution Sensitivity Using Generalized Likelihood Ratio Method''. In *Proceedings of the 13th International Workshop on Discrete Event Systems (WODES)*, 123–128. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Peng, Y., M. Fu, and J. Hu. 2016b. ''On the Regularity Conditions and Applications for Generalized Likelihood Ratio Method''. In In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 919–930. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Peng, Y., M. Fu, J. Hu, and B. Heidergott. 2017. ''A New Unbiased Stochastic Derivative Estimator for Discontinuous Sample Performances with Structural Parameters''. *Operations Research*: accepted.

Pflug, G. C. 1989. ''Sampling Derivatives of Probabilities''. *Computing* 42:315–328.

Pflug, G. C. 1996. *Optimization of Stochastic Models*. Boston: Kluwer Academic.
Pichitlamken, J., and B. L. Nelson. 2003. ''A Combined Procedure for Optimization via Simulation". *ACM Transactions on Modeling and Computer Simulation* 13 (2): 155–179.
Plambeck, E. L., B.-R. Fu, S. M. Robinson, and R. Suri. 1996. ''Sample-Path Optimization of Convex Stochastic Performance Functions". *Mathematical Programming* 75:137–176.
Polyak, B. T., and A. B. Juditsky. 1992. ''Acceleration of Stochastic Approximation By Averaging". *SIAM Journal on Control and Optimization* 30 (4): 838–855.
Powell, W. B. 2011. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Second ed. Hoboken, NJ: John Wiley & Sons.
ProModel 2017. "SimRunner". https://www.processmodel.com/learning-center/users-guide/chapter-9/ Accessed April 26, 2017.
Qu, H., and M. C. Fu. 2014. ''Gradient Extrapolated Stochastic Kriging". *ACM Transactions on Modeling and Computer Simulation* 24 (4).
Rao, U., A. Scheller-Wolf, and S. Tayur. 2000. ''Development of a Rapid-Response Supply Chain at Caterpillar". *Operations Research* 48 (2): 189–204.
Reiman, M., and A. Weiss. 1989. ''Sensitivity Analysis for Simulations via Likelihood Ratios". *Operations Research* 37:830–844.
Robbins, H., and S. Monro. 1951. ''A Stochastic Approximation Method". *Annals of Mathematical Statistics* 22:400–407.
Robinson, S. M. 1996. ''Analysis of Sample-Path Optimization". *Mathematics of Operations Research* 21:513–528.
Rubinstein, R. Y. 1986. *Monte Carlo Optimization, Simulation and Sensitivity of Queueing Networks*. John Wiley & Sons.
Rubinstein, R. Y. 1997. ''Optimization of Computer Simulation Models with Rare Events". *European Journal of Operations Research* 99:89–112.
Rubinstein, R. Y. 1999. ''The Cross-Entropy Method for Combinatorial and Continuous Optimization". *Methodology and Computing in Applied Probability* 2:127–190.
Rubinstein, R. Y., and A. Shapiro. 1990. ''Optimization of Static Simulation Models By the Score Function Method". *Mathematics and Computers in Simulation* 32:373–392.
Rubinstein, R. Y., and A. Shapiro. 1993. *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method*. John Wiley & Sons.
Ruppert, D. 1988, February. ''Efficient Estimations From a Slowly Convergent Robbins-Monro Process". Technical report, Cornell University Operations Research and Industrial Engineering, Ithaca, NY.
Schruben, L. W. 2000. ''Mathematical Programming Models of Discrete Event System Dynamics". In *Proceedings of the 2000 Winter Simulation Conference*, edited by J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, 381–385. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
Schruben, L. W., and V. Cogliano. 1981. ''Simulation Sensitivity Analysis: A Frequency Domain Approach". In *Proceedings of the 1981 Winter Simulation Conference*, edited by T. I. Ören, C. M. Delfosse, and C. M. Shub, 455–459. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
Scott, W., P. I. Frazier, and W. B. Powell. 2011. ''The Correlated Knowledge Gradient for Simulation Optimization of Continuous Parameters Using Gaussian Process Regression". *SIAM Journal on Optimization* 21 (3): 996–1026.
Serfling, R. J. 1980. *Approximation Theorems of Mathematical Statistics*. Wiley Series in Probability and Mathematical Statistics. New York: Wiley & Sons.
Shapiro, A. 2003. ''Monte Carlo Sampling Methods". In *Stochastic Programming*, edited by A. Ruszczynski and A. Shapiro, Handbooks in Operations Research and Management Science, 353–425. Amsterdam: Elsevier.

Shapiro, A., D. Dentcheva, and A. Ruszczyński. 2009. *Lectures on Stochastic Programming: Modeling and Theory*. MPS-SIAM Series on Optimization. Philadelphia, PA: SIAM-MPS.

Shashaani, S., S. R. Hunter, and R. Pasupathy. 2016. ''ASTRO-DF: Adaptive Sampling Trust-Region Optimization Algorithms, Heuristics, and Numerical Experience.''. In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 554–565. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Shi, L., and S. Ólafsson. 2000a. ''Nested Partitions Method for Global Optimization''. *Operations Research* 48:390–407.

Shi, L., and S. Ólafsson. 2000b. ''Nested Partitions Method for Stochastic Optimization''. *Methodology and Computing in Applied Probability* 2:271–291.

Siekman, P. 2000. ''New Victories in the Supply-Chain Revolution''. *Fortune* (October 30).

Simio 2017. "Simio". http://www.simio.com Accessed April 26, 2017.

Spall, J. C. 1992. ''Multivariate Stochastic Approximation using Simultaneous Perturbation Gradient Approximation''. *IEEE Transactions on Automatic Control* 37:332–341.

Spall, J. C. 1994. ''Developments in Stochastic Optimization Algorithms with Gradient Approximations Based on Function Measurements''. In *Proceedings of the 1994 Winter Simulation Conference*, edited by J. D. Tew, S. Manivannan, D. A. Sadowski, and A. F. Seila, 207–214. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Spall, J. C. 2003. *Introduction to Stochastic Search and Optimization*. New York, NY: Wiley-Interscience.

Sun, L., L. J. Hong, and Z. Hu. 2014. ''Balancing Exploitation and Exploration in Discrete Optimization via Simulation Through a Gaussian Process-Based Search''. *Operations Research* 62 (6): 1416–1438.

Suri, R. 1983. ''Implementation of Sensitivity Calculations on a Monte-Carlo Experiment''. *Journal of Optimization Theory and Applications* 40 (4): 625–630.

Suri, R. 1989. ''The State of the Art and Research Issues Illustrated Via the GI/G/1 Queue''. *Proceedings of the IEEE* 77 (1): 114–137.

Suri, R., and Y.-T. Leung. 1987. ''Single Run Optimization of a SIMAN Model for Closed Loop Flexible Assembly Systems''. In *Proceedings of the 1987 Winter Simulation Conference*, edited by A. Thesen, H. Grant, and W. D. Kelton, 738–748. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Suri, R., and Y.-T. Leung. 1989. ''Single Run Optimization of Discrete Event Simulations---An Empirical Study Using the M/M/l Queue''. *IIE Transactions* 21 (1): 35–49.

Suri, R., and M. A. Zazanis. 1988. ''Perturbation Analysis Gives Strongly Consistent Sensitivity Estimates for the $M/G/1$ Queue''. *Management Science* 34:39–64.

Swain, J. J. 2015. ''Simulated Worlds''. *OR/MS Today* 42 (5): 36–49.

Thompson, W. C. 1968. ''The Application of Simulation in Computer System Design and Optimization''. In *Second Conference on the Applications of Simulation*, edited by J. Reitman and A. Ockene, 286–290.

van Beers, W. C. M., and J. P. C. Kleijnen. 2003. ''Kriging for Interpolation in Random Simulation''. *Journal of the Operational Research Society* 54 (3): 255–262.

Wang, H., R. Pasupathy, and B. W. Schmeiser.. 2013. ''Integer-Ordered Simulation Optimization using R-SPLINE: Retrospective Search with Piecewise-Linear Interpolation and Neighborhood Enumeration''. *ACM Transactions on Modeling and Computer Simulation* 23 (3): Article 17.

WSC 2017. "WSC Archive". http://informs-sim.org Accessed April 26, 2017.

Xu, J., B. L. Nelson, and L. J. Hong. 2010. ''Industrial Strength COMPASS: A Comprehensive Algorithm and Software for Optimization via Simulation''. *ACM Transactions on Modeling and Computer Simulation* 20 (1): Article 3.

Xu, J., B. L. Nelson, and L. J. Hong. 2013. ''An Adaptive Hyperbox Algorithm for High-Dimensional Discrete Optimization via Simulation Problems''. *INFORMS Journal on Computing* 25 (1): 133–146.

Yan, D., and H. Mukai. 1992. ‘‘Stochastic Discrete Optimization”. *SIAM Journal of Control and Optimization* 30:594–612.

Zabinsky, Z. B. 2003. *Stochastic Adaptive Search for Global Optimization*. Kluwer.

Zabinsky, Z. B., and R. L. Smith. 1992. ‘‘Pure Adaptive Search in Global Optimization”. *Mathematical Programming* 53:323–338.

Zazanis, M. A., and R. Suri. 1986. ‘‘Perturbation Analysis of the GI/GI/1 Queue”. Technical report.

Zazanis, M. A., and R. Suri. 1994. ‘‘Perturbation Analysis of the GI/GI/1 Queue”. *Queueing Systems: Theory and Applications* 18:199–248.

Zhang, X., D. Meiser, Y. Liu, B. Bonner, and L. Lin. 2014. ‘‘Kroger Uses Simulation-Optimization to Improve Pharmacy Inventory Management”. *Interfaces* 44 (1): 70–84.

AUTHOR BIOGRAPHIES

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, A. James Clark School of Engineering, University of Maryland, College Park, where he has been since 1989. He attended his first WSC in 1988 and served as Program Chair for the 2011 WSC. He is a Fellow of INFORMS and IEEE. His e-mail address is mfu@umd.edu, and his Web page is https://www.rhsmith.umd.edu/directory/michael-fu.

SHANE G. HENDERSON is a professor in the School of Operations Research and Information Engineering at Cornell University. His research interests include discrete-event simulation and simulation optimization, and he has worked for some time with emergency services and bike sharing applications. He co-edited the Proceedings of the 2007 Winter Simulation Conference. His e-mail address is sgh9@cornell.edu, and his Web page is http://people.orie.cornell.edu/∼shane.