Noname manuscript No.
(will be inserted by the editor)

Solving a Class of Simulation-based Optimization Problems Using "Optimality in Probability"

Jianfeng Mao · Christos G. Cassandras

Received: date / Accepted: date

Abstract We consider a class of simulation-based optimization problems using *optimality in probability*, an approach which yields what is termed a "champion solution". Compared to the traditional *optimality in expectation*, this approach favors the solution whose actual performance is more likely better than that of any other solution; this is an alternative complementary approach to the traditional optimality sense, especially when facing a dynamic and nonstationary environment. Moreover, using *optimality in probability* is computationally promising for a class of simulation-based optimization problems, since it can reduce computational complexity by orders of magnitude compared to general simulation-based optimization methods using *optimality in expectation*. Accordingly, we have developed an "Omega Median Algorithm" in order to effectively obtain the champion solution and to fully utilize the efficiency of well-developed off-line algorithms to further facilitate timely decision making. An inventory control problem with nonstationary demand is included to illustrate and interpret the use of the Omega Median Algorithm, whose performance is tested using simulations.

Keywords Simulation-based Optimization · Optimality in Probability · Nonstationary Inventory Control

The authors' work is supported in part by NSFC under grant U1733102, by CUHK(SZ) under grant PF.01.000404, by ATMRI under grant M4061216.057, by NTU under grant M58050030, by AcRF under grant RG 33/10 M52050117, by NSF under grants CNS-1239021, ECCS-1509084, CNS-1645681, and IIP-1430145, by AFOSR under grant FA9550-15-1-0471, by ONR under grant N00014-09-1-1051, and by Bosch and the MathWorks.

Jianfeng Mao
School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China 518172
E-mail: jfmao@cuhk.edu.cn

Christos G. Cassandras
Division of Systems Engineering, Boston University, Brookline, MA 02446, USA
E-mail: cgc@bu.edu

1 Introduction

In discrete event systems, we are often faced with a class of stochastic optimization problems that involve only parametric optimization and no structural changes to the underlying systems. In such cases, optimality in expectation is commonly adopted with problems formulated as

$$\min_{u\in\Phi} E[J(u,\omega)] \tag{1}$$

where u is the decision variable, Φ is the feasible decision space of u, and ω is used to index sample paths resulting from different realizations of a collection of random variables that affect the performance $J(u,\omega)$.

In the context of discrete event systems, we commonly face a dynamic stochastic process, in which u is an event-triggered online control action and $J(u,\omega)$ is the actual performance of u over a certain sample path ω. For example, in the on-line inventory control problem later considered in Section 3, u is the order quantity decided at the beginning of each period, ω is a sample path constructed by a sequence of demands, and $J(u,\omega)$ is the corresponding operating cost, including setup cost, holding cost and shortage cost.

Since it is typically impossible to derive the closed form of $E[J(u,\omega)]$ in (1), simulation-based optimization methods need to be employed to obtain a near-optimal solution. In what follows, we define an "evaluation" as an operation of calculating the value of $J(u,\omega)$ for a specific u over a specific sample path ω. In general, simulation-based optimization methods include two major operations:

1. **Solution Assessment**: Implement M evaluations for a specific u over M sample paths and estimate the expected performance of solution u, $E[J(u,\omega)]$, by sample average approximation, *i.e.*, $\sum_{i=1}^{M} J(u,\omega_i)\Big/M$;
2. **Search Strategy**: Use the sample average approximation in 1) to rank solutions and search for better solutions in promising areas according to gradient information (if possible) or certain partition structures.

Let I denote the total number of solutions explored in a simulation-based method and C denote the complexity of an evaluation. Then, the total complexity can be measured by the computational effort of implementing $M \cdot I$ evaluations, that is, $O(M \cdot I \cdot C)$ (M is not necessarily a constant throughout the entire search process). To get a near optimal (or good enough) solution, we need to implement more evaluations to refine solution assessment, *i.e.,* larger M, and explore a greater number of solutions, *i.e.,* larger I. Since both M and I can be very large in solving a general simulation-based optimization problem using optimality in expectation, this approach is computationally intensive or even intractable for many applications in practice.

A number of simulation-based optimization methods have been developed over the past few decades. Computational effort can be reduced by either using a smaller number M of evaluations in assessment, such as Ordinal Optimization [18] and Optimal Computing Budget Allocation [9], or by reducing I in search, such as Nested Partitions [29] and COMPASS [20], or by both ways, such as Perturbation Analysis [17] and Retrospective Optimization [10,23]. Moreover, to further improve computational efficiency, these methods may be applied to certain approximations of the original systems with little loss of accuracy in the optimization solutions, such as the use of Stochastic Flow Models [8,37] and Hindsight Optimization [11, 36]. Since these methods still need to employ sample average approximations to assess every explored solution (or estimate its performance gradient), their complexity can still be approximated as $O(M \cdot I \cdot C)$ with either smaller M or smaller I or both. In practice, timely decision making is usually preferable or required in a dynamic environment. The heavy computational burden of those methods using optimality in expectation limits their applications in such situations.

Moreover, we argue that optimality in expectation is not truly "optimal" in certain cases since the expected performance is not exactly the actual performance, but only a promising guess. This kind of optimality is generally suitable for a stationary environment, in which probability distributions remain unchanged over time and the objective value is the average performance over the long term. However, in practice we often face a nonstationary environment, as in the inventory control problem included in the paper, in which nonstationary demand is a common occurrence in industries with short product life cycles, seasonal patterns, varying customer behavior, or other factors [27]. When we continually or periodically make decisions, the probability distributions used are only valid for a short term and need to be occasionally updated. Clearly, optimality in expectation does not necessarily lead to the "best" solution in this case.

In this paper, we propose an alternative sense of optimality, "optimality in probability", which favors a solution that has a *higher chance* to get a better actual performance. The best solution using optimality in probability, termed "Champion Solution", is defined as the one whose actual performance is more likely better than that of any other solution. *Optimality in probability* is an alternative complementary approach to optimality in expectation, especially when facing a dynamic and nonstationary environment. Moreover, using *optimality in probability* is computationally promising for a class of simulation-based optimization problems, since it can reduce computational complexity by orders of magnitude compared to general simulation-based optimization methods using *optimality in expectation.* Accordingly, we develop an "Omega Median Algorithm" to obtain the champion solution without iteratively searching for better solutions based on sample average approximations, a process which is computationally intensive and commonly required when seeking optimality in expectation. Furthermore, although it is quite challenging to solve many stochastic optimization problems, their corresponding deterministic versions, which can be regarded as optimization problems defined over a single sample path, have been efficiently solved by certain off-line algorithms. The Omega Median Algorithm is able to fully utilize the efficiency of these well-developed off-line algorithms to further facilitate timely decision making, which is clearly preferable in a dynamic environment with limited computational resources. It should be noted that, although an analytical solution of single sample path optimization problems is quite helpful in improving computational efficiency, it is not required for the implementation of the Omega Median Algorithm.

In the rest of the paper, we first introduce the champion solution and then develop an efficient simulation-based optimization method, termed Omega Median Approximation in Section 2. We then consider a nonstationary inventory control problem in Section 3. Numerical results are given in Section 4 to demonstrate the performance of the champion solution. We close with conclusions in Section 5.

2 Champion Solution

The "Champion Solution" is the best solution using optimality in probability and defined for general stochastic minimization problems as follows, where $\Pr[\cdot]$ is the usual notation for "probability":

Definition 1 The champion solution is a solution u^c such that

$$\Pr\left[J(u^c,\omega) \le J(u,\omega)\right] \ge 0.5, \quad \forall\, u \in \Phi, \tag{2}$$

where $J(u,\omega)$ is the actual performance of u over a certain sample path ω.

Remark: A natural question which immediately arises is "why do we select 0.5?" rather than some $q > 0.5$ and define the champion solution as u' below such that

$$\Pr\left[J(u',\omega) \le J(u,\omega)\right] \ge q, \quad \forall\, u \in \Phi, \tag{3}$$

which looks even better than u^c in (2). However, a definition using $q > 0.5$ is not meaningful for the large majority of stochastic problems with continuous random variables. Generally speaking, if the sample path ω is constructed with continuous random variables ω and continuous functions $J(u,\omega)$, we can have for $u' \ne u^c$:

$$\Pr\left[J(u',\omega) < J(u^c,\omega)\right] = \Pr\left[J(u',\omega) \le J(u^c,\omega)\right]. \tag{4}$$

From (3), we have $\Pr\left[J(u',\omega) \le J(u^c,\omega)\right] \ge q$. Combining it with (4), we have

$$\Pr\left[J(u^c,\omega) \le J(u',\omega)\right] \le 1 - q,$$

which contradicts (2) if $q > 0.5$. Even if there might exist some u' that satisfies (3), it will be still the same as u^c defined in (2). Therefore, we will set 0.5 instead of some $q > 0.5$ in the definition of champion solution.

The NBA Finals can be used as an example to illustrate the champion solution. The champion team (the champion solution) will be determined from two teams (solutions) based on the results in 7 games (sample-paths). The champion solution is the team (solution) that wins more games (performs better in more sample-paths). Ideally, if there is an infinite number of games (sample-paths), then the champion solution is the team with winning ratio of more than 50%.

For cases with more than two solutions, we interpret the champion solution through the example of presidential elections originally used for Arrow's Impossibility Theorem in social choice theory [1]. Imagine we have three candidates (solutions) A, B and C. Each voter (sample-path) will rank the three candidates according to his or her own preference. Now, we randomly pick three voters' preference lists (sample-paths) as shown in the following table, where $A \succ B$ means A is preferred over B.

	Voter 1	Voter 2	Voter 3
Preference	$A \succ B \succ C$	$B \succ C \succ A$	$C \succ B \succ A$

Based on the the three voters' preferences, we can estimate that

- A : $\Pr[A \succ B] = 33\%$, $\Pr[A \succ C] = 33\%$;
- B : $\Pr[B \succ A] = 67\%$, $\Pr[B \succ C] = 67\%$;
- C : $\Pr[C \succ A] = 67\%$, $\Pr[C \succ B] = 33\%$.

Clearly, B should be the president (the champion solution) because B gets a higher preference (performs better) than all the other candidates (solutions) from the majority of voters (sample-paths).

2.1 Optimality in Expectation vs. Optimality in Probability

The champion solution favors the winning ratio instead of the winning scale. That is why we call it "Champion Solution". We can still use the example of NBA Finals to illustrate the new sense of optimality and compare it with the traditional one. Imagine it was finished in 6 games and the results are shown in the following table.

	Game 1	Game 2	Game 3	Game 4	Game 5	Game 6
A	107	103	84	106	90	98
B	100	97	103	104	101	95

Team A is the champion (the champion solution) because Team A won more games than Team B. However, we can also find out that the average score of Team B, 100, is higher than 98, the one of Team A, which implies that Team B is actually better than Team A in the sense of "Optimality in Expectation" commonly adopted in the literature.

Clearly, the champion solution is the best solution in a different sense of optimality, termed "Optimality in Probability" here, which may be a better optimality sense than the traditional "Optimality in Expectation" in some applications, such as the NBA Finals.

Generally, the champion solution and the traditional optimal solution are not the same, but they coincide under the following "**Non-singularity Condition**" as shown in [25]:

$$\Pr\left[J(u',\omega) \le J(u'',\omega)\right] \ge 0.5 \implies E\left[J(u',\omega)\right] \le E\left[J(u'',\omega)\right], \quad \forall u', u'' \in \Phi$$

The interpretation of the Non-singularity Condition is that if u' is more likely better than u'' (in the sense of resulting in lower cost), then the expected cost under u' will be lower than the one under u''. This is consistent with common sense in that any solution A more likely better than B should result in A's expected performance being better than B's. Only "singularities" such as $J(u',\omega) \gg J(u'',\omega)$ with an unusually low probability for some (u',u'') can affect the corresponding expectations so that this condition may be violated. It is straightforward to verify this Non-singularity Condition for several common cases. For example, consider $J(u,\omega) = (u-\omega)^2$, where ω is a uniform random variable over $[a,b]$. The function satisfies the Non-singularity Condition and the solution $(a+b)/2$ achieves both optimality in probability and in expectation.

In addition, even though decision makers may prefer "optimality in expectation" in their applications, the champion solution still has a very promising performance if the corresponding problem does not exhibit significant singularities because it can beat all the other solutions with a probability greater than 0.5.

2.2 A Condition for the Existence of a Champion Solution

A champion solution may not always exist for a general stochastic optimization problem. If there are only two feasible solutions, as in the NBA Finals, a champion solution can be obviously guaranteed. However, this is not the case even for as few as three feasible solutions. Recalling the example of presidential elections, what if Voter 3 changes his or her preference as shown in the following table?

	Voter 1	Voter 2	Voter 3
Preference	$A \succ B \succ C$	$B \succ C \succ A$	$C \succ A \succ B$

This time we have

- A : $\Pr[A \succ B] = 67\%$, $\Pr[A \succ C] = 33\%$;
- B : $\Pr[B \succ A] = 33\%$, $\Pr[B \succ C] = 67\%$;
- C : $\Pr[C \succ A] = 67\%$, $\Pr[C \succ B] = 33\%$.

No candidate can be elected as president (the champion solution) because no one can be preferred over all the other candidates (solutions) from the majority of voters (sample-paths); this is in fact the case addressed in Arrow's paradox [1].

In the following, we will establish a sufficient existence condition, which can be utilized later in the inventory problem considered in the next section. To accomplish that, we first define the concepts of "ω-problem", "ω-solution" and "ω-median" for the class of stochastic optimization problems in (1). (As these definitions are based on or related to a single sample-path ω, we name their initials as ω-.)

Definition 2 An ω-problem is the deterministic optimization problem defined over a single sample-path ω, *i.e.*,

$$\min_{u \in \Phi} J(u, \omega).$$

where $\Phi \subseteq \mathbb{R}$ is the constraint set of u and $J(\cdot, \omega) : \Phi \mapsto \mathbb{R}$ is a scalar function of u.

Definition 3 The ω-solution is the solution u^ω such that

$$u^\omega = \min_{\hat{u}} \left\{ \hat{u} : J(\hat{u}, \omega) = \min_{u \in \Phi} J(u, \omega) \right\}. \quad (5)$$

Remark: Although $\hat{u}$ is a minimizer of $J(u, \omega)$ and may not be unique. u^ω is defined as the smallest one picked from these minimizers to guarantee the uniqueness of u^ω. We will impose the regularity assumptions that the minimizer $\hat{u}$ of $J(u, \omega)$ exists and u^ω is measurable. Then u^ω is a random variable related to sample-path ω.

Definition 4 An ω-median is a median of the probability distribution of ω-solution u^ω, *i.e.*, the solution u^m such that

$$\Pr[u^\omega \leq u^m] \geq 0.5 \quad \text{and} \quad \Pr[u^\omega \geq u^m] \geq 0.5 \tag{6}$$

Remark: The two probabilities in (6) are the cumulative distribution function (*cdf*) and complementary cumulative distribution function (*ccdf*) of u^ω respectively. Both probabilities can be strictly more than 0.5 at the same time if u^ω is not continuous. Moreover, the ω-median may not be unique for cases such that the pdf or pmf of u^ω is 0 for some values of u^ω.

Theorem 1 *If $J(u, \omega)$ is a scalar unimodal function of u for any ω, then an ω-median is a champion solution.*

Proof Since $J(u, \omega)$ is a scalar unimodal function of u for any ω, we have

$$J(u', \omega) \leq J(u'', \omega), \quad \text{for any} \quad u'' < u' < u^\omega; \tag{7}$$

and

$$J(u', \omega) \leq J(u'', \omega), \quad \text{for any} \quad u^\omega < u' < u''. \tag{8}$$

Assume u^m is the ω-median. For any solution $u > u^m$, we have

$$\begin{aligned} &\Pr[J(u^m, \omega) \leq J(u, \omega)] \\ =&\Pr[J(u^m, \omega) \leq J(u, \omega) | u^\omega \leq u^m] \Pr[u^\omega \leq u^m] \\ +&\Pr[J(u^m, \omega) \leq J(u, \omega) | u^\omega > u^m] \Pr[u^\omega > u^m] \end{aligned} \tag{9}$$

From (8), if $u > u^m$ and $u^m \geq u^\omega$, then $J(u^m, \omega) \leq J(u, \omega)$, which implies that

$$\Pr[J(u^m, \omega) \leq J(u, \omega) | u^\omega \leq u^m] = 1 \tag{10}$$

Since u^m is the ω-median, we have $\Pr[u^\omega \leq u^m] \geq 0.5$. Combining it with (9) and (10), we have

$$\begin{aligned} &\Pr[J(u^m, \omega) \leq J(u, \omega)] \\ \geq&0.5 + \Pr[J(u^m, \omega) \leq J(u, \omega) | u^\omega > u^m] \Pr[u^\omega > u^m] \\ \geq&0.5 \end{aligned}$$

The case of $u < u^m$ can be similarly proved. Therefore, u^m satisfies the definition of champion solution

$$\Pr[J(u^m, \omega) \leq J(u, \omega)] \geq 0.5, \quad \text{for any } u \in \Phi.$$

which implies u^m is a champion solution. □

2.3 A Condition for the Uniqueness of a Champion Solution

The champion solution may not be unique in general. The uniqueness can be guaranteed if the following conditions can be satisfied as shown in Theorem 2.

Theorem 2 *Let $J(u, \omega)$ be a scalar strictly unimodal function of u for any ω,* i.e.,

$$J(u', \omega) < J(u'', \omega), \forall u^\omega < u' < u'' \quad \text{and} \quad J(u', \omega) < J(u'', \omega), \forall u'' < u' < u^\omega,$$

where $u^\omega = \arg\min_{u \in \Phi} J(u, \omega)$. If there exists some u^m such that

$$\Pr[u^\omega \leq u^m - \epsilon] < 0.5 \quad \text{and} \quad \Pr[u^\omega \geq u^m + \epsilon] < 0.5 \quad \text{for every } \epsilon > 0 \tag{11}$$

then the champion solution is u^m and unique.

Proof We will only prove the result for cases such that u^ω is a continuous random variable. The discrete case can be similarly proved. The uniqueness of a champion solution can be shown by proving the following two parts: [a] the champion solution must be some u^m satisfying (11); [b] the solution u^m satisfying (11) is unique.
Part [a]: Assume on the contrary that there exists some champion solution u' such that

$$\Pr[u^\omega \leq u' - \epsilon] \geq 0.5 \quad \text{or} \quad \Pr[u^\omega \geq u' + \epsilon] \geq 0.5 \quad \text{for some } \epsilon > 0; \tag{12}$$

which implies that

$$\Pr[u^\omega \geq u'] < 0.5 \quad \text{or} \quad \Pr[u^\omega \leq u'] < 0.5 \quad . \tag{13}$$

From (13), without loss of generality, assume $\Pr[u^\omega \geq u'] < 0.5$. Then there exists some $\delta > 0$ such that

$$\Pr[u^\omega \geq u' - \delta] < 0.5 \tag{14}$$

It holds that

$$\begin{aligned} &\Pr[J(u', \omega) \leq J(u' - \delta, \omega)] \\ =&\Pr[J(u', \omega) \leq J(u' - \delta, \omega) | u^\omega \geq u' - \delta] \Pr[u^\omega \geq u' - \delta] \\ +&\Pr[J(u', \omega) \leq J(u' - \delta, \omega) | u^\omega < u' - \delta] \Pr[u^\omega < u' - \delta] \end{aligned} \tag{15}$$

From the definition of u^ω and $J(u, \omega)$ is a scalar strictly unimodal function of u for any ω, it must satisfy that $J(u', \omega) > J(u' - \delta, \omega)$ if $u^\omega < u' - \delta < u'$, which implies that

$$\Pr[J(u', \omega) \leq J(u' - \delta, \omega) | u^\omega < u' - \delta] = 0$$

Combining this with (14) and (15), we have

$$\begin{aligned} &\Pr[J(u', \omega) \leq J(u' - \delta, \omega)] \\ =&\Pr[J(u', \omega) \leq J(u' - \delta, \omega) | u^\omega \geq u' - \delta] \Pr[u^\omega \geq u' - \delta] \\ \leq&\Pr[u^\omega < u' - \delta] < 0.5 \end{aligned} \tag{16}$$

Since u' is a champion solution that should have

$$\Pr[J(u', \omega) \leq J(u, \omega)] \geq 0.5, \quad \text{for any } u \in \Phi. \tag{17}$$

which contradicts (16). Therefore, the champion solution must be some u^m satisfying (11).

Part[b]: Without loss of generality, assume on the contrary that there exists some $u' > u^m$ that also satisfies (11). From (11), we have

$$\Pr[u^\omega < u^m + \epsilon] \geq 0.5 \quad \text{for every } \epsilon > 0.$$

Combining this with $u' > u^m$, there exists some $\epsilon > 0$ such that

$$\Pr[u^\omega < u' - \epsilon] \geq \Pr[u^\omega < u^m + \epsilon] \geq 0.5$$

which contradicts the assumption that u' also satisfies (11). Thus, the solution u^m that satisfies (11) is unique.

The result follows from Part [a] and Part [b] above. □

2.4 Multinomial Optimal Solution vs. Champion Solution

From Theorems 1 and 2, the champion solution can be obtained using the ω-median if the corresponding conditions are satisfied. For example, if u^ω is integer-valued and satisfies the probability mass function (pmf) and cumulative density function (cdf) as shown in Figure 1, then the champion solution is 49, the ω-median marked as the bold line labeled with "CS" in this case.

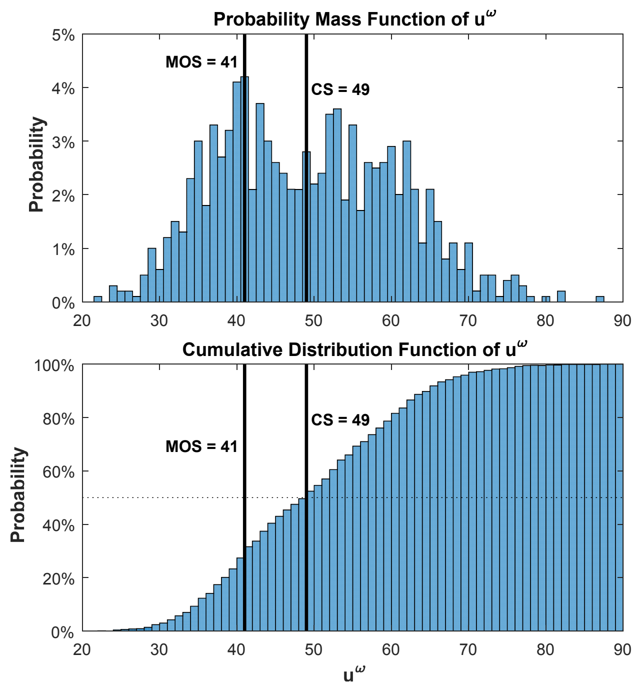

Fig. 1 The pmf and cdf of u^ω

Another interesting solution with a different optimality type can also be derived based on the pmf of u^ω, that is, the solution obtained using multinomial selection [16,26,34]. For convenience, this solution is termed as *Multinomial Optimal Solution* (MOS) in the rest of paper. The MOS is a solution with the highest probability of being the actual best among all the solutions, that is, the solution u^{MOS} such that

$$u^{MOS} = \arg\max_{\hat{u}\in\Phi} \left\{ \Pr\left[J(\hat{u},\omega) \le J(u,\omega), \forall u \in \Phi\right] \right\} \tag{18}$$

According to the definition of ω-solution u^ω in (5), if u^ω is a discrete random variable and $J(u,\omega)$ is strictly unimodal, we have

$$\Pr\left[u^\omega = \hat{u}\right] = \Pr\left[J(\hat{u},\omega) \le J(u,\omega), \forall u \in \Phi\right]$$

Combining it with (18), we have

$$u^{MOS} = \arg\max_{\hat{u}\in\Phi} \Pr\left[u^\omega = \hat{u}\right]. \tag{19}$$

As shown in Figure 1, since u^ω can only be integer-valued and it achieves the highest probability of 4.2% at $u^\omega = 41$, the MOS is 41 (marked as the bold line labeled with "MOS"). Clearly, MOS is not the same as CS for general cases. Their performance difference will be demonstrated in the section of numerical results below for the inventory control problem.

Furthermore, it should be noted that the pdf and cdf of u^ω in Figure 1 can only be estimated through many replications. More replications is needed for a more accurate estimation of MOS and CS. Since MOS is the solution with the highest probability of $\Pr\left[u^\omega\right]$ as in (19), a good estimation of MOS requires a good estimation of entire probability distribution of u^ω, which consume a great number of Monte Carlo simulations. Moreover, the estimation of MOS is sensitive and may vary considerably as the number of replications increases.

The champion solution can be estimated through the median of the probability distribution of u^ω if the conditions in Theorem 1 can be satisfied. A good estimation of CS does not require a good estimation of the entire pdf of u^ω. Besides, the estimation of CS is not sensitive and gradually changes as the number of replications increases. In the following section, we will develop an algorithm to obtain an estimation of CS and prove that this estimation can approach an actual CS exponentially fast as the number of replications increases.

2.5 Omega Median Algorithm

Theorem 1 provides a sufficient existence condition for a champion solution for a class of simulation-based optimization problems. If it is satisfied, then a champion solution is guaranteed and can be efficiently obtained by computing the ω-median. We can efficiently obtain an estimate of the ω-median using the Omega Median Algorithm (OMA) in Table 1 even though the closed form of the *cdf* and *ccdf* of u^ω cannot be derived in the class of stochastic optimization problems in (1).

Table 1 Omega Median Algorithm

Step 1:	Randomly generate M sample-paths $\omega^1, ..., \omega^M$;
Step 2:	Obtain the ω-solutions, u^{ω_i}, by solving the deterministic ω-problems $\min_{u\in\Phi} J(u,\omega_i)$ for $i = 1, ..., M$;
Step 3:	Find the median solution $\hat{u}^m$ from $u^{\omega_1}, ..., u^{\omega_M}$.

The median solution $\hat{u}^m$ derived in Step 3 of OMA is an estimator of the ω-median. Let $\mathbf{1}(\cdot)$ denote an indicator function and

$$G_M(u) \equiv \frac{1}{M}\sum\nolimits_{j=1}^{M} \mathbf{1}(u^{\omega_j} \leq u);$$
$$\bar{G}_M(u) \equiv \frac{1}{M}\sum\nolimits_{j=1}^{M} \mathbf{1}(u^{\omega_j} \geq u).$$

Then, $G_M(u)$ and $\bar{G}_M(u)$ are the unbiased estimates of the *cdf* and *ccdf* of u^ω respectively. It can be easily verified that the median solution $\hat{u}^m$ is the solution that satisfies

$$G_M(\hat{u}^m) \geq 0.5 \quad \text{and} \quad \bar{G}_M(\hat{u}^m) \geq 0.5\ .$$

For any given u, based on the strong law of large numbers, $G_M(u)$ and $\bar{G}_M(u)$ converge to $\Pr[u^\omega \leq u]$ and $\Pr[u^\omega \geq u]$ respectively *w.p.*1 (with probability 1) as $M \to +\infty$. Thus, $\hat{u}^m$ also converges to an ω-median u^m *w.p.*1 as $M \to +\infty$.

Furthermore, $\hat{u}^m$ can approach an ω-median u^m exponentially fast as M increases as shown in Theorems 3 and 4 below, which enables us to estimate the ω-median with a smaller number M of sample paths.

Let U^m denote the set of ω-medians satisfying (6) and $\hat{U}^m_M$ denote the set of medians based on estimated cdf and ccdf as shown below:

$$U^m = \{u^m : \Pr[u^\omega \leq u^m] \geq 0.5, \quad \Pr[u^\omega \geq u^m] \geq 0.5\}$$
$$\hat{U}^m_M = \{\hat{u}^m : G_M(\hat{u}^m) \geq 0.5, \quad \bar{G}_M(\hat{u}^m) \geq 0.5\}$$

Theorem 3 *If* $\inf_{u^m\in U^m} \Pr(u^\omega = u^m) > 0$*, then there always exists some constant* C *such that*

$$\Pr[\hat{U}^m_M \cap U^m \neq \varnothing] \geq 1 - 2e^{-CM}$$

Proof Without loss of generality, assume $\inf_{u^m\in U^m} \Pr(u^\omega = u^m) = c > 0$, $\Pr(u^\omega < u^m) = p_1$ and $\Pr(u^\omega > u^m) = p_2$. From the definition of ω-median, we have $p_1 + c \geq 0.5$ and $p_2 + c \geq 0.5$ for any $u^m \in U^m$. Combining it with $p_1 + c + p_2 = 1$ and $c > 0$, we have

$$p_1 < 0.5, \quad p_2 < 0.5.$$

The event $[\hat{U}^m_M \cap U^m \neq \varnothing]$ is equivalent to the event $[G_M(u^m) \geq 0.5$ and $\bar{G}_M(u^m) \geq 0.5 \mid u^m \in U^m]$, which can be further equivalently reduced to $[L_M(u^m) < 0.5$ and $\bar{L}_M(u^m) < 0.5 \mid u^m \in U^m]$, where

$$L_M(u) = \frac{1}{M}\sum_{j=1}^{M} \mathbf{1}(u^{\omega_j} < u),$$

$$\bar{L}_M(u) = \frac{1}{M}\sum_{j=1}^{M} \mathbf{1}(u^{\omega_j} > u).$$

Therefore, we have

$$\begin{aligned}
&\Pr[\hat{U}_M^m \cap U^m \neq \varnothing] \\
=&\Pr[L_M(u^m) < 0.5 \text{ and } \bar{L}_M(u^m) < 0.5 \mid u^m \in U^m] \\
=&1 - Pr[L_M(u^m) > 0.5 \text{ or } \bar{L}_M(u^m) > 0.5 \mid u^m \in U^m] \\
=&1 - \big(\Pr[L_M(u^m) > 0.5 \mid u^m \in U^m] + Pr[\bar{L}_M(u^m) > 0.5 \mid u^m \in U^m]\big)
\end{aligned} \tag{20}$$

Clearly, $\mathbf{1}(u^{\omega_j} < u^m), j = 1, ..., M$ are *i.i.d.* 0-1 random variables and $E[\mathbf{1}(u^{\omega_j} < u^m) \mid u^m \in U^m] = p_1$. Then, based on the Chernoff-Hoeffding Theorem [19], we have for any $\epsilon > 0$

$$\Pr[L_M(u^m) \geq p_1 + \epsilon \mid u^m \in U^m] \leq e^{-D(p_1+\epsilon||p_1)M}$$

where $D(x||y) = x \log \frac{x}{y} + (1-x) \log \frac{1-x}{1-y}$. Similarly, we can also have

$$\Pr[\bar{L}_M(u^m) \geq p_2 + \epsilon \mid u^m \in U^m] \leq e^{-D(p_2+\epsilon||p_2)M}$$

Combining the two inequalities above with $p_1 < 0.5$ and $p_2 < 0.5$, we further have

$$\Pr[L_M(u^m) > 0.5 \mid u^m \in U^m] \leq \Pr[L_M(u^m) \geq 0.5 \mid u^m \in U^m] \leq e^{-D(0.5||p_1)M}$$

$$\Pr[\bar{L}_M(u^m) > 0.5 \mid u^m \in U^m] \leq \Pr[\bar{L}_M(u^m) \geq 0.5 \mid u^m \in U^m] \leq e^{-D(0.5||p_2)M}$$

Combining them with (20), we can finally have

$$\begin{aligned}
\Pr[\hat{U}_M^m \cap U^m \neq \varnothing] &\geq 1 - e^{-D(0.5||p_1)M} - e^{-D(0.5||p_2)M} \\
&\geq 1 - 2e^{-CM}
\end{aligned}$$

where $C = \min\big(D(0.5||p_1), D(0.5||p_2)\big)$ □

Theorem 4 *If* $\Pr(u^\omega = u^m) = 0$ *for* $u^m \in U^m$*, then for any* $\epsilon > 0$*, there always exists* $C > 0$ *such that*

$$\begin{aligned}
&\Pr\big[\, |G_M(u^m) - 0.5| < \epsilon \;\big|\; u^m \in U^m\big] \geq 1 - 2e^{-CM}, \\
&\Pr\big[\, |\bar{G}_M(u^m) - 0.5| < \epsilon \;\big|\; u^m \in U^m\big] \geq 1 - 2e^{-CM}.
\end{aligned}$$

Proof From $\Pr(u^\omega = u^m) = 0$ and the definition of U^m, we have

$$\Pr[u^\omega \leq u^m \mid u^m \in U^m] = 1 - \Pr[u^\omega \geq u^m \mid u^m \in U^m] = 0.5$$

which implies that

$$E\big[G_M(u^m) \mid u^m \in U^m\big] = 0.5$$

Since $\mathbf{1}(u^{\omega_j} \leq u^m), j = 1, ..., M$ are *i.i.d.* 0-1 random variables and $E[\mathbf{1}(u^{\omega_j} < u^m) \mid u^m \in U^m] = 0.5$, based on the Chernoff-Hoeffding Theorem [19], we have for any $\epsilon > 0$

$$\begin{aligned}
&\Pr[G_M(u^m) \geq 0.5 + \epsilon \mid u^m \in U^m] \leq e^{-D(0.5+\epsilon||0.5)M} \quad \text{and} \\
&\Pr[G_M(u^m) \leq 0.5 - \epsilon \mid u^m \in U^m] \leq e^{-D(0.5-\epsilon||0.5)M}
\end{aligned}$$

where $D(x||y) = x\log\frac{x}{y} + (1-x)\log\frac{1-x}{1-y}$. Therefore, we have

$$\begin{aligned}
&\Pr\left[\ |G_M(u^m) - 0.5| < \epsilon \ \middle|\ u^m \in U^m\right] \\
=&1 - \Pr[G_M(u^m) \geq 0.5 + \epsilon \mid u^m \in U^m] - \Pr[G_M(u^m) \leq 0.5 - \epsilon \mid u^m \in U^m] \\
\geq&1 - e^{-D(0.5+\epsilon||0.5)M} - e^{-D(0.5-\epsilon||0.5)M} \\
\geq&1 - 2e^{-CM}.
\end{aligned}$$

where $C = \min\big(D(0.5+\epsilon||0.5), D(0.5-\epsilon||0.5)\big)$.

It can be similarly proved that $\Pr\left[\ |\bar{G}_M(u^m) - 0.5| < \epsilon \ \middle|\ u^m \in U^m\right] \geq 1 - 2e^{-CM}$. □

Theorem 3 corresponds to the case that u^ω is discrete and Theorem 4 is mainly for the case that u^ω is continuous. Theorem 3 has a stronger sense of convergence than Theorem 4, which implies that $\hat{u}^m$ converges faster in discrete cases than in continuous ones.

3 An Application: Inventory Control with Nonstationary Demand

To illustrate and interpret the use of the Omega Median Algorithm, we consider an on-line periodic review inventory control problem with nonstationary demand as depicted in Figure 2 as a discrete event system (DES), in which fixed setup cost and full backlogging are adopted. The following notation will be used in the rest of the paper:

- x_i = Inventory level in period i;
- d_i = Demand in period i;
- u_i = Order quantity in period i;
- h = Holding cost rate for inventory;
- p = Penalty cost rate for backlog;
- K = Fixed setup cost per order;
- $\delta(u_i) = \begin{cases} 1 & u_i > 0 \\ 0 & u_i = 0 \end{cases}$.

The one-period demand d_i is nonstationary, *i.e.*, its corresponding probability distribution is arbitrary and allowed to vary and correlate over periods i.

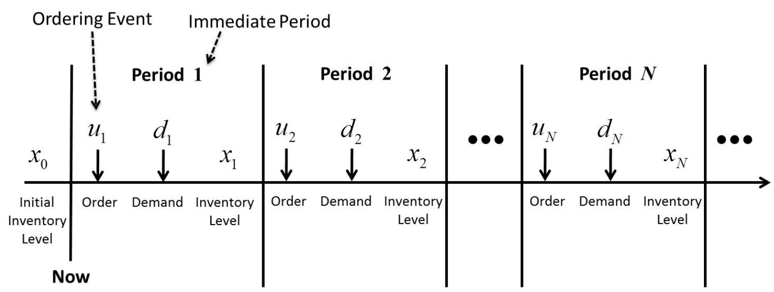

Fig. 2 On-line Inventory Control Process

An ordering event may be triggered at the beginning of a period, namely, an order of u_i items may be placed in period i. A fixed setup cost K will be triggered if $u_i > 0$. The inventory level x_i is counted after the one-period demand d_i, *i.e.*, $x_i = x_{i-1} + u_i - d_i$, which results in the maintenance cost of period i (either holding or shortage cost) defined below,

$$H(x_i) = h \cdot \max(x_i, 0) + p \cdot \max(-x_i, 0). \tag{21}$$

The average operating cost in each period, including both maintenance cost and setup cost, determines the system performance.

The static (s, S) policy is an optimal policy for the cases with stationary demands using optimality in expectation, which has been extensively studied for the inventory systems with setup cost [28,22,32,40,5,3]. Once the two thresholds (s, S) are optimally determined, the corresponding optimal ordering quantity can be simply derived as $u_i = S - x_{i-1}$ if $x_{i-1} \leq s$ and $u_i = 0$ otherwise. Several efficient methods have been developed in [33,39,14] to find the optimal static (s, S) policy for the stationary cases. When nonstationary demand processes arise, the static (s, S) policy is not optimal [3]: the optimal order decisions cannot be simply derived by optimizing the two static thresholds (s, S). Even though for several special classes of nonstationary demand in [41,15], the two-threshold policy can still be optimal but no more static, *i.e.*, the optimal policy becomes (s_i, S_i) varying over period i. The specialty and computational inefficiency limit the application of the two-threshold policy for the nonstationary cases.

Some efforts have been made towards the nonstationary inventory control with fixed setup cost [2,7,6,21] and without setup cost [31,24,27]. A heuristic similar to Silver-Meal heuristics [30] is proposed in [2] and requires to explicitly compute the probability distributions of cumulative demands, which is not plausible for demands with complicated patterns. In [7,21], static-dynamic uncertainty approaches were developed for a class of nonstationary demand that still require the assumption of mutually independent demands over periods. In [6], nonstationary cases are approximated by averaging demands over periods and then a stationary policy is computed by utilizing the algorithm in [39], which will be benchmarked against the proposed Omega Median Algorithm in the numerical results section below.

Although general simulation-based methods can still be utilized to determine the best order decision using optimality in expectation, they are computationally intensive or even intractable as analyzed in Section 3.3. Instead, we pursue the best solution in the sense of optimality in probability, namely, the "Champion Solution", which is a very attractive alternative when facing a nonstationary environment.

In the on-line inventory control process depicted in Fig 2, we make an order decision at the beginning of each period. The rolling horizon method can be applied, in which we look ahead N periods and the actual performance over a specific N-period sample path $\omega = \{d_1, d_2, ..., d_N\}$ can be defined as the total cost:

$$\begin{aligned} J_N(u_1, u_2, ..., u_N, \omega) &= \sum\nolimits_{i=1}^{N} \big(H(x_i) + K \cdot \delta(u_i)\big) \\ s.t.\ \ x_i &= x_{i-1} - d_i + u_i,\ i = 1, ..., N. \end{aligned} \tag{22}$$

where $H(x_i) + K \cdot \delta(u_i)$ is the operating cost in period i, including maintenance cost and setup cost.

Since only the immediate-period order decision, u_1, is required each time, we will focus on u_1 and optimally determine $u_2, ..., u_N$ based on the choice of u_1. Then, the actual performance over a specific N-period sample path ω becomes solely associated with u_1 as follows:

$$\begin{aligned} J_N(u_1, \omega) = & \big(H(x_1) + K \cdot \delta(u_1)\big) \\ & + \min_{u_2,\ldots,u_N} \sum\nolimits_{i=2}^{N} \big(H(x_i) + K \cdot \delta(u_i)\big) \\ & s.t. \;\; x_i = x_{i-1} - d_i + u_i, \; i = 1, ..., N. \end{aligned} \tag{23}$$

In the ideal case of looking ahead for an infinite horizon, the actual performance over a specific sample path ω can be formulated as the infinite-horizon average cost:

$$J(u_1, \omega) \equiv \lim_{N \to +\infty} \frac{1}{N} \big\{ J_N(u_1, \omega) \big\} \tag{24}$$

We aim at the champion solution using the actual (as opposed to expected) performance function in (24).

3.1 Existence of Champion Solution

The inventory control problem can be solved by sequentially answering the two questions below:

Question 1: Whether to order (Yes or No);

Question 2: How many items to order if "Yes" to Question 1.

Since Question 1 has only two options, its champion solution can be guaranteed and easily obtained as follows,

$$\begin{cases} \text{Yes} & \text{if } \Pr[u_1^\omega > 0] \geq 50\% \\ \text{No} & \text{otherwise.} \end{cases}$$

where u_1^ω is the ω-solution of minimizing $J(u_1, \omega)$ in (24) and $\Pr[u_1^\omega > 0]$ is the probability to place a positive order.

Question 2 is conditioned on "Yes" to Question 1, which implies that $u_1 > 0$ in Question 2. In the following, we will verify the existence of a champion solution for $u_1 > 0$ with the help of the lemma below.

Lemma 1 *$J_N(u_1, \omega)$ in (23) is strictly K-convex in u_1 for $u_1 > 0$, that is, for any $0 < u_1 < u_1' < u_1''$, it holds that*

$$K + J_N(u_1'', \omega) > J_N(u_1', \omega) + \left(\frac{u_1'' - u_1'}{u_1' - u_1}\right)\left(J_N(u_1', \omega) - J_N(u_1, \omega)\right).$$

Proof It can be proved that $L_N(x_1, \omega)$ is K-convex in x_1 using a similar way as shown in Section 4.2 in [4] (The definition of K-convex can be found in [4,28]). Combining it with $x_1 = u_1 + x_0 - d_1$, $L_N(u_1 + x_0 - d_1, \omega)$ is also K-convex in u_1.

From the definition of $H(x)$ in (21), $H(x_1)$ is strictly convex in x_1, which implies $H(u_1 + x_0 - d_1)$ is also strictly convex in u_1.

Recalling the definition of $J_N(u_1, \omega)$ in (23). From $u_1 > 0$, we have

$$J_N(u_1, \omega) = H(u_1 + x_0 - d_1) + K + L_N(u_1 + x_0 - d_1, \omega)$$

Combining it with the fact that $H(u_1 + x_0 - d_1)$ is strictly convex in u_1 and $L_N(u_1 + x_0 - d_1, \omega)$ is K-convex in u_1, we have $J_N(u_1, \omega)$ is strictly K-convex in u_1 for $u_1 > 0$. □

Based on Lemma 1 and the definition of $J(u_1, \omega)$ in (24), we prove the following theorem.

Theorem 5 $J(u_1, \omega)$ *is strictly convex in* u_1 *for* $u_1 > 0$.

Proof From Lemma 1, $J_N(u_1, \omega)$ is strictly K-convex in u_1 for $u_1 > 0$, that is, it satisfies that for any $0 < u_1 < u_1' < u_1''$

$$K + J_N(u_1'', \omega) > J_N(u_1', \omega) + (\frac{u_1'' - u_1'}{u_1' - u_1})(J_N(u_1', \omega) - J_N(u_1, \omega)).$$

Then we apply limit operator at both sides and can have

$$\lim_{N\to+\infty} \frac{K + J_N(u_1'', \omega)}{N} > \lim_{N\to+\infty} \frac{J_N(u_1', \omega)}{N} + (\frac{u_1'' - u_1'}{u_1' - u_1}) \lim_{N\to+\infty} \frac{(J_N(u_1', \omega) - J_N(u_1, \omega))}{N}$$

which implies that for any $0 < u_1 < u_1' < u_1''$,

$$J(u_1'', \omega) > J(u_1', \omega) + (\frac{u_1'' - u_1'}{u_1' - u_1})(J(u_1', \omega) - J(u_1, \omega)).$$

The inequality above is equivalent to the definition of strictly convex function, that is, $J(u_1, \omega)$ is strictly convex in u_1 for $u_1 > 0$. □

Theorem 5 implies that $J(u_1, \omega)$ is strictly unimodal for $u_1 > 0$, which satisfies the sufficient existence condition identified in Theorem 1 and paves the way to the uniqueness of a champion solution using Theorem 2. Therefore, a champion solution can be guaranteed to address Question 2 and can be obtained using OMA, which may be also unique if the probabilistic condition in (11) can be verified in the simulation results.

3.2 Implementation of OMA

Although d_i, $i = 1, 2, \ldots$, is nonstationary, we can still estimate their probability distributions based on the most recently updated information. Sample paths can then be randomly generated in Step 1 of OMA using these estimates.

Step 2 of OMA determines the major portion of its computational complexity, which can be largely reduced if we manage to find an efficient algorithm to solve the corresponding ω-problems. In the context of this inventory control problem, the ω-problem is to find the ω-solution u_1^ω of minimizing $J(u_1, \omega)$ in (24). This ω-solution u_1^ω can be well approximated by minimizing $J_N(u_1, \omega)$ in (23) with a large enough N. Furthermore, it can be easily verified that, if $u_1^*, ...u_N^*$ can minimize $J_N(u_1, ..., u_N, \omega)$ in (22), then u_1^* can also minimize $J_N(u_1, \omega)$ in (23). Therefore, we can finally obtain the ω-solution u_1^ω by minimizing $J_N(u_1, ..., u_N, \omega)$ in (22) with a sufficiently large N.

The problem of minimizing $J_N(u_1, ..., u_N, \omega)$ in (22) is closely related to the following problem in (25), which is a dynamic lot-sizing problem with backlogging as defined in the literature. Several methods have been developed to solve this type of problems. The seminal work was the one developed in [35] to solve the case without backlogging. Then in [38], although backlogging is considered, it is required to generate dominant set and its size grows exponentially with respect to N. Finally, in [12,13], highly efficient algorithms were developed to solve the dynamic lot-sizing problem for both cases without and with backlogging.

$$\begin{aligned} &\min_{u_1,\dots,u_N} \sum\nolimits_{i=1}^{N} \left\{ H(x_i) + K \cdot \delta(u_i) \right\} \\ &s.t.\ \ x_i = x_{i-1} - d_i + u_i,\ i = 1, ..., N; \\ &\qquad \sum\nolimits_{i=1}^{N} u_i + x_0 = \sum\nolimits_{i=1}^{N} d_i. \end{aligned} \tag{25}$$

The only difference between the two problems results from the second constraint, which can be interpreted as the condition of "zero inventory at last". Since profits earned from sales are not included in the objective, it would never be optimal to place a new order at the last period which would mostly end up with a negative inventory level. The terminal effect of "ordering nothing at last" and "ending with negative inventory" are quite undesirable. Solving the problem in (25) instead with the extra second constraint can be very helpful in approximating the ω-solution when using a relatively small N. Since the problem in (25) has been well studied in [13], we can efficiently solve each ω-problem with complexity $O(N \log N)$ for general cases.

The remaining Step 3 of OMA can be trivially fulfilled once we have M ω-solutions.

3.3 Complexity Analysis

Clearly, the complexities of Step 1 and 3 of OMA are $O(MN)$ and $O(M)$ respectively. With the help of the algorithm in [13], the complexity of Step 2 is $O(M \cdot N \log N)$. Thus, we can finally efficiently obtain a champion solution of the nonstationary inventory control problem in complexity $O(M \cdot N \log N)$ by applying OMA.

If we try a general simulation-based optimization method using optimality in expectation, then we need to solve the following stochastic optimization problem

(26) at each decision point:

$$\begin{aligned} \min_{u_1} \ \bar{J}_N(u_1) = E\Big\{ & \big(H(x_1) + K \cdot \delta(u_1)\big) \\ + \min_{\mu_2,\dots,\mu_N} & E\Big\{ \sum\nolimits_{i=2}^{N} \big(H(x_i) + K \cdot \delta(u_i)\big) \Big\} \Big\} \\ s.t. \ \ & x_i = x_{i-1} - d_i + u_i, \ i = 1, \dots, N; \\ & u_i = \mu_i(x_{i-1}), \ i = 2, \dots, N. \end{aligned} \tag{26}$$

where $\mu_i(\cdot)$ is the feedback control policy to determine u_i based on the state x_{i-1}. Clearly, even for a given u_1, computing $\bar{J}_N(u_1)$ is a notoriously hard dynamic programming problem. Although a heuristic termed "Hindsight Optimization" [11] can be employed to approximate the second term in the objective of (26) as the expected hindsight-optimal value below,

$$E\left\{ \min_{u_2,\dots,u_N} \sum\nolimits_{i=2}^{N} \Big(H(x_i) + K \cdot \delta(u_i)\Big) \right\},$$

still requires a complexity of $O(M \cdot N \log N)$ to assess a specific choice of u_1. Moreover, it needs to go through a search process to get a near optimal u_1. If there are a total of I solutions explored in the process, then the total computational complexity is $O(M \cdot I \cdot N \log N)$, which is an order of magnitude higher than that of OMA.

4 Numerical Results

We illustrate the performance of champion solution through numerical examples. The following parameters are identical to those used in [40]:

- Fixed Setup Cost $K = 64$;
- Holding Cost Rate $h = 1$;
- Penalty Cost Rate $p = 9$.

The mean value μ_i will be randomly picked from a set of numbers between 10 and 75 in increments of 5, that is, $\{10, 15, 20, ..., 70, 75\}$. The champion solution will be benchmarked against the (s, S) policy and the multinomial optimal solution for both stationary and nonstationary cases. Before proceeding to the comparisons, we will first demonstrate the approximation of ω-median and its convergence rate with respect to the number of replications.

4.1 ω-median Approximation

An example of estimating the ω-median is shown in Figure 3, in which $M = 200$ sample-paths are generated. The ω-solutions are obtained by solving 200 corresponding ω-problems through the algorithm in [13]. The solid line in Figure 3 is the *cdf* function of the ω-solution constructed based on these sample-paths. The estimate of the ω-median is $u^m = 78$, which is indicated through the dashed line.

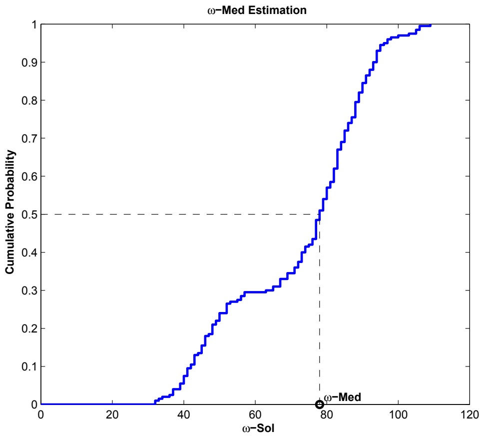

Fig. 3 ω-median Approximation

4.2 Convergence of ω-median in M

The convergence of the ω-median in the number of sample-paths M is shown in Figure 4, in which M varies from 10 to 1000 in increments of 10. It can be seen that the estimate of the ω-median quickly converges within 100 replications, which supports the result in Theorem 3.

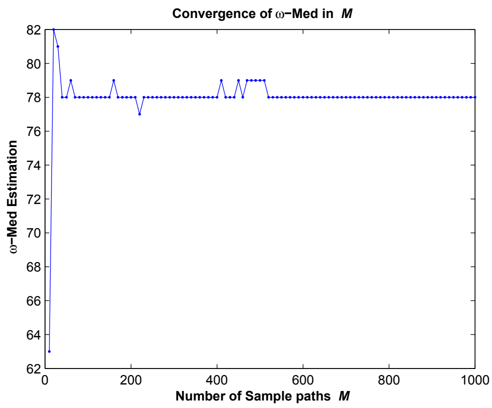

Fig. 4 Convergence of ω-median in M

4.3 Stationary Cases

We set $\mu = 20$ to simulate stationary cases. Then the optimality in expectation can be achieved using the optimal static policies (s^*, S^*), which have been exactly derived by using the algorithm in [39] for stationary cases with different μ. This provides us an opportunity to benchmark the performance of the champion solution against the best solution in the sense of "optimality in expectation". In the following, we will compare the actual performance of the three methods described below in 1000 randomly generated instances.

1. Method **SS**: Order decisions are directly obtained according to the optimal static policy ($s^* = 14, S^* = 62$) as obtained in [39];
2. Method **MS**: Order decisions are derived by using multinomial selection with $M = 1000$ sample paths at the beginning of each period, namely, the estimates of multinomial optimal solutions;
3. Method **CS**: Order decisions are obtained by using the ω-median approximation with $M = 1000$ sample paths at the beginning of each period, namely, the estimates of champion solutions.

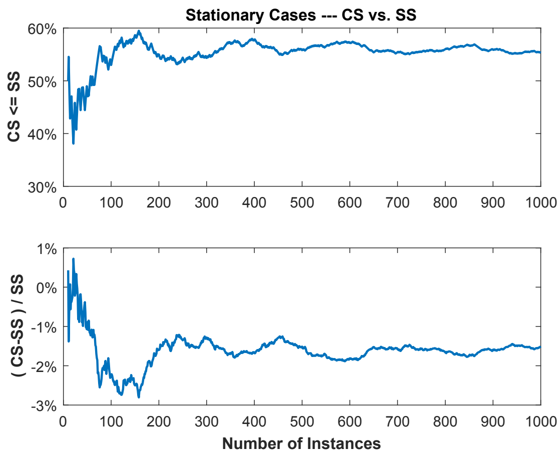

Fig. 5 Stationary Cases: CS *vs.* SS

Pairwise comparisons are carried out between these methods. The comparison between **CS** and **SS** is depicted in Figure 5. The upper plot shows the percentage of instances that **CS** is no worse than **SS** changes as more instances are simulated. The percentage is 55.30% after finishing 1000 instances. The lower plot shows that the mean fractional actual cost difference changes along with more instances simulated. The fractional actual cost difference is calculated as $\frac{(C_{cs}-C_{ss})}{C_{ss}}$, where C_{ss} and C_{cs} are the costs of using the methods **SS** and **CS** respectively. Within 1000 instances, the mean cost of **CS** is 1.51% less than the one of **SS**. Based on the numerical results above, the performance difference between **CS** and **SS** is very small and the champion solution can perform as well as the optimal (s^*, S^*) policy in the stationary cases.

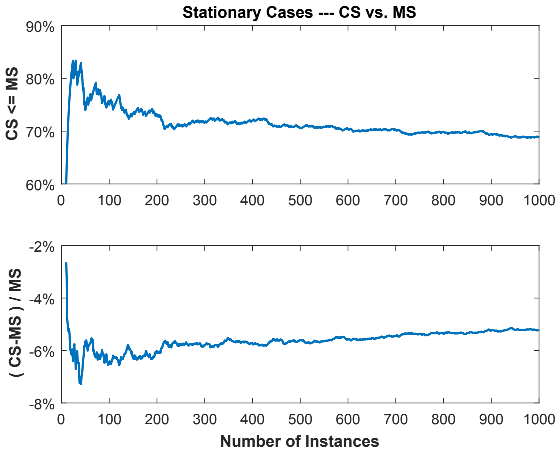

Fig. 6 Stationary Cases: CS *vs.* MS

The comparison between **CS** and **MS** is depicted in Figure 6. Its upper and lower plots and the ones in the following figures are similarly defined as in Figure 5. Based on the 1000 instances simulated, the percentage of instances that **CS** is no worse than **MS** is 68.80% and the mean cost of **CS** is 5.21% less than the one of **MS**. Therefore, the champion solution (CS) more likely performs better than the multinomial optimal solution (MOS) and CS is also about 5% better than MOS in cardinal value.

We can zoom in on a certain instance to get a more detailed analysis about the performance difference between MOS and CS. Figure 1 is actually an example of the estimated pmf and cdf of u^ω based on 1000 replications for the inventory control problem. As shown before, CS is 49 and MOS is 41 in this case. We start with the comparison between the MOS, *i.e.*, 41 and its neighboring solution of 42. Since the inventory control problem satisfies the condition in Theorems 1 and 2, we can use the similar reasoning adopted in the proof of theorems to derive the following two probability:

$$\Pr\left[J(u^{MOS},\omega) < J(42,\omega)\right] = \Pr\left[u^\omega \leq u^{MOS} = 41\right] = 31.6\%$$

$$\Pr\left[J(42,\omega) < J(u^{MOS},\omega)\right] = \Pr\left[u^\omega \geq= 42\right] = 68.4\%$$

which means that the solution of 42 is better than MOS with a probability of 68.4% and worse than MOS with a probability of 31.6%. However, $\Pr\left[u^\omega = 42\right] = 2.1\%$, which is a lot smaller than $\Pr\left[u^\omega = u^{MOS} = 41\right] = 4.2\%$ as shown in Figure 1. Therefore, the solution $\hat{u}$ with inferior performance based on the probability of $\Pr\left[J(\hat{u},\omega) \leq J(u,\omega), \forall u \in \Phi\right]$ may still be better than MOS in the majority of cases. The comparisons can be furthered. Although 42 seems better than MOS in more cases, we can similarly derive that

$$\Pr\left[J(42,\omega) < J(43,\omega)\right] = \Pr\left[u^\omega \leq 42\right] = 33.7\%$$

$$\Pr\left[J(43,\omega) < J(42,\omega)\right] = \Pr\left[u^\omega \geq 43\right] = 66.3\%$$

$$\Pr\left[J(43,\omega) < J(u^{MOS},\omega)\right] \geq \Pr\left[u^{\omega} \geq 43\right] = 66.3\%$$

which implies that the solution of 43 is better than both MOS and 42 with probability greater than or equal to 66.3%. Similar comparison results can be derived until the champion solution of 49, which can be better than all of these solutions in more cases.

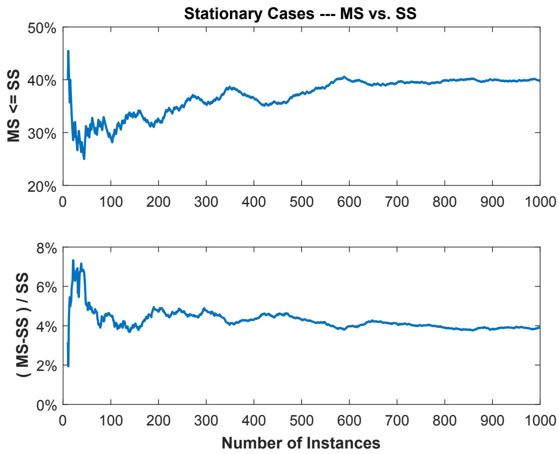

Fig. 7 Stationary Cases: MS *vs.* SS

The comparison between **MS** and **SS** is depicted in Figure 7. Based on the 1000 instances simulated, the percentage of instances that **MS** is no worse than **SS** is 39.80% and the mean cost of **MS** is 3.91% higher than the one of **SS**. Therefore, the (s^*, S^*) policy more likely performs better than MOS and MOS is also about 4% worse than the (s^*, S^*) policy in cardinal value.

To summerize, for the stationary cases, the (s^*, S^*) policy and the champion solution perform similarly and they are all better than the multinomial optimal solution.

4.4 Nonstationary Cases

Based on historical data, practitioners can usually observe and estimate some demand pattern over periods, that is, a sequence of different expected demand μ_i for period i, before placing orders. We set different μ_i for each period to simulate the situation in the nonstationary cases. In particular, to reflect different demand pattern observed, we randomly select μ_i from the values listed in $\{10, 15, 20, ..., 70, 75\}$. We again generate 1000 instances to compare the three methods:

1. Method **SS**: Order decisions are directly obtained according to a heuristic nonstationary policy (s_i, S_i) for each period i. A common heuristic method is to determine (s_i, S_i) according to μ_i in the corresponding period i as if demands are stationary with the mean value of μ_i. For example, if $\mu_1 = 15, \mu_2 = 30, \mu_3 = 20, ...$, then we can look up the table obtained in [39] to find their corresponding

optimal values, choose ($s_1 = 10, S_1 = 49$), ($s_2 = 23, S_2 = 66$), ($s_3 = 14, S_3 = 62$), ..., to apply in period $1, 2, 3, ...$, respectively. Clearly, this heuristic (s_i, S_i) policy is not optimal for the nonstationary case.
2. Method **MS**: Order decisions are similarly obtained based on the multinomial solutions as the **MS** used for stationary cases.
3. Method **CS**: Order decisions are similarly obtained based on the champion solutions as the **CS** used for stationary cases.

The comparison between **CS** and **SS** is shown in Figure 8. Based on the 1000 instances simulated, the percentage of instances that **CS** is no worse than **SS** is 95.30% and the mean cost of **CS** is 16.43% less than the one of **SS**. Therefore, the champion solution performs better than the heuristic (s, S) policy in almost all of the instances and CS is also about 16% better than the heuristic (s, S) policy in cardinal value.

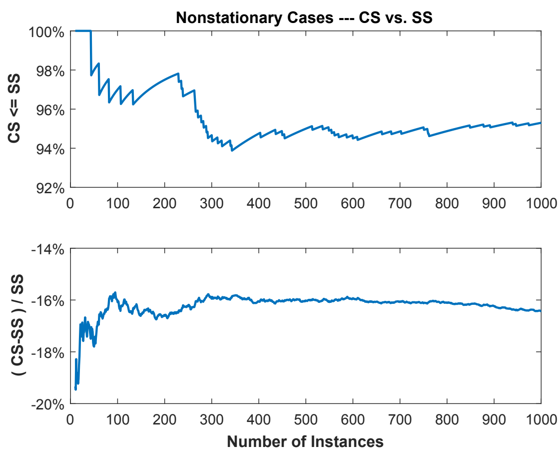

Fig. 8 Nonstationary Cases: CS *vs.* SS

The comparison between **CS** and **MS** is depicted in Figure 9. Based on the 1000 instances simulated, the percentage of instances that **CS** is no worse than **MS** is 72.40% and the mean cost of **CS** is 5.03% less than the one of **SS**. Therefore, CS much more likely performs better than MOS and it is also about 5% better than MOS in cardinal value.

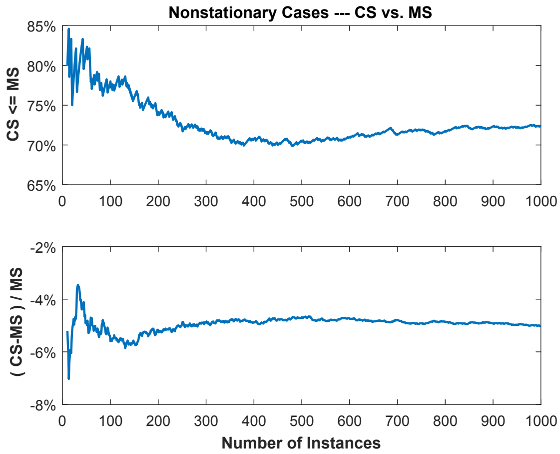

Fig. 9 Nonstationary Cases: CS *vs.* MS

The comparison between **MS** and **SS** is depicted in Figure 10. Based on the 1000 instances simulated, the percentage of instances that **MS** is no worse than **SS** is 88.40% and the mean cost of **MS** is 12.00% less than the one of **SS**. Therefore, MOS performs better than the heuristic (s, S) policy in most of the instances and MOS is also about 12% better than the heuristic (s, S) policy in cardinal value.

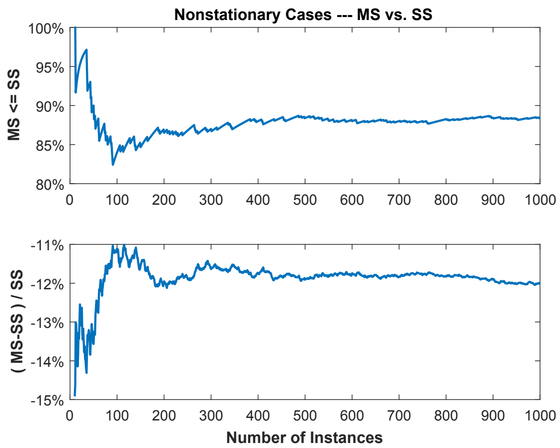

Fig. 10 Nonstationary Cases: MS *vs.* SS

To summarize, for the nonstationary cases, the champion solution still performs better than the multinomial optimal solution and they are all much better than the heuristic (s, S) policy.

5 Conclusion

An alternative optimality sense, optimality in probability, is proposed in this paper. The best solution using optimality in probability is termed a "Champion Solution" whose actual performance is more likely better than that of any other solution. A sufficient existence and uniqueness condition for the champion solution are proved for a class of simulation-based optimization problems. A highly efficient method, the Omega Median Algorithm (OMA), is developed to compute the champion solution without iteratively exploring better solutions based on sample average approximations. OMA can reduce the computational complexity by orders of magnitude compared to general simulation-based optimization methods using optimality in expectation.

The champion solution becomes particularly meaningful when facing a nonstationary environment. As shown in the application of inventory control with nonstationary demand, the solution using optimality in expectation is not necessarily optimal and is computationally intractable in a dynamic environment. The champion solution is a good alternative and computationally promising. Its corresponding solution algorithm, OMA, can fully utilize the efficiency of existing well-developed off-line algorithms to further facilitate timely decision making, which is preferable in a dynamic environment with limited computing resources. Moreover, even for some stationary scenarios as shown in the numerical results, the "Champion Solution" can still achieve a performance comparable to the one using optimality in expectation.

It is nontrivial to show the existence of a champion solution and OMA cannot be directly applied when facing general cases with multiple decision variables. Some partial decomposition methods can be utilized to reduce the original problem into scalar optimization problems, which is quite common for dynamic programming problems with separable cost functions. Nonetheless, the existence issue is still troublesome and limits the application of champion solutions. A possible generalized version of champion solution is one defined as the solution u^c that achieves the maximum of $q(\cdot)$:

$$q(u^c) = \max_{\hat{u}} \Big\{ q(\hat{u}) = \max_{q} \big\{ q : \Pr\left[J(\hat{u},\omega) \le J(u,\omega) \right] \ge q,\ \forall\ u \in \Phi \big\} \Big\}$$

or equivalently,

$$q(u^c) = \max_{\hat{u}} \Big\{ q(\hat{u}) = \min_{u \in \Phi} \big\{ \Pr\left[J(\hat{u},\omega) \le J(u,\omega) \right] \big\} \Big\}$$

It can be easily verified that the champion solution defined in (2) is a special case of this generalized version. Moreover, the existence of this generalized champion solution can be guaranteed for general cases. We will aim at generalizing the concept of champion solution and extend it to a wider class of multidimensional stochastic optimization problems in future work.

References

1. Kenneth J. Arrow. *Social Choice and Individual Values.* Yale University Press, 1963.
2. R. G. Askin. A procedure for production lot sizing with probabilistic dynamic demand. *AIIE Transactions*, 12(2):132–137, 1981.

3. S. Axsäter. *Inventory Control (Second Edition)*. Springer, 2006.
4. D. P. Bertsekas. *Dynamic Programming and Optimal Control Vol.1 (Second Edition)*. Athena Scientific, 2000.
5. D. Beyer and S. P. Sethi. The classical average-cost inventory models of iglehart (1963) and veinott and wagner (1965) revisited. *Journal of Optimization Theory and Applications*, 101(3):523–555, 1999.
6. Srinivas Bollapragada and Thomas E. Morton. A simple heuristic for computing nonstationary (s,S) policies. *Operations Research*, 47(4):576–584, 1999.
7. James H. Bookbinder and Jin-Yan Tan. Strategies for the probabilistic lot-sizing problem with service-level constraints. *Management Science*, 34(9):1096–1108, 1988.
8. C. G. Cassandras, Y. Wardi, B. Melamed, G. Sun, and C. G. Panayiotou. Perturbation analysis for on-line control and optimization of stochastic fluid models. *IEEE Trans. on Automatic Control*, 47(8):1234–1248, 2002.
9. C. H. Chen and L. H. Lee. *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*. World Scientific Publishing Co., 2011.
10. H. Chen. *Stochastic Root Finding in System Design*. Ph.D. Thesis. Purdue University, West Lafayette, Indiana, USA., 1994.
11. E. K. P. Chong, R. L. Givan, and H. S. Chang. A framework for simulation-based network control via hindsight optimization. In *Proceedings of the 39th IEEE Conference on Decision and Control*, pages 1433–1438, 2000.
12. AWI Federgruen and Michal Tzur. A simple forward algorithm to solve general dynamic lot sizing models with n periods in $o(n \log n)$ or $o(n)$ time. *Management Science*, 37(8):909–925, 1991.
13. AWI Federgruen and Michal Tzur. The dynamic lot-sizing model with backlogging: A simple $o(n \log n)$ algorithm and minimal forecast horizon procedure. *Naval Research LOgistics*, 40(4):459–478, 1993.
14. M.C. Fu. Sample path derivatives for (s,S) inventory systems. *Operations Research*, 42(2):351–364, 1994.
15. G. Gallego and Ö. Özer. Integrating replenishment decisions with advanced demand information. *Management Science*, 47(10):1344–1360, 2001.
16. D. Goldsman and B. L. Nelson. Comparing systems via simulation. *Chapter 8 in The Handbook of Simulation (J. Banks, ed.)*, pages 273–306, 1998.
17. Y. C. Ho and X. R. Cao. *Perturbation Analysis of Discrete-Event Dynamic Systems*. Kluwer Academic Publisher, Boston, 1991.
18. Yu-Chi Ho, Qian-Chuan Zhao, and Qing-Shan Jia. *Ordinal Optimization: Soft Optimization for Hard Problems*. Springer Science & Business Media, 2008.
19. W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, March 1963.
20. L. Jeff Hong and Barry L. Nelson. Discrete optimization via simulation using compass. *Operations Research*, 54(1):115–129, 2006.
21. Z. Hua, J. Yang, F. Huang, and X. Xu. A static-dynamic strategy for spare part inventory systems with nonstationary stochastic demand. *Journal of the Operational Research Society*, 60:1254 –1263, 2009.
22. Donald L. Iglehart. Optimality of (s, S) policies in the infinite-horizon dynamic inventory problem. *Management Science*, 9(2):259–267, 1963.
23. J. Jin. *Simulation-Based Retrospective Optimization of Stochastic Systems*. Ph.D. Thesis. Purdue University, West Lafayette, Indiana, USA., 1998.
24. Retsef Levi, Martin Pal, Robin O. Roundy, and David B. Shmoys. Approximation algorithms for stochastic inventory control models. *Mathematics of Operations Research*, 32(2):284–302, 2007.
25. J. Mao and C. G. Cassandras. On-line optimal control of a class of discrete event systems with real-time constraints. *Journal of Discrete Event Dynamic Systems*, 20(2):187–213, 2010.
26. J. O. Miller, B. L. Nelson, and C. H. Reilly. Efficient multinomial selection in simulation. *Naval Research Logistics*, 45:459–482, 1998.
27. John J. Neale and Sean P. Willems. Managing inventory in supply chains with nonstationary demand. *Interfaces*, 39(5):388–399, 2009.
28. Herbert Scarf. The optimality of (s,S) policies in the dynamic inventory problem. *Technical Report, Applied Mathematics and Statistics Laboratory, Stanford University*, 1959.
29. L. Shi and S. Olafsson. Nested partitions method for global optimization. *Operations Research*, 48(3):390–407, 2000.

30. E. A. Silver and H. C. Meal. A heuristic selecting lot-size requirements for the case of a deterministic time varying demand rate and discrete opportunities for replenishment. *Production Inventory Management*, 14(9):64–74, 1973.
31. James T. Treharne and Charles R. Sox. Adaptive inventory control for nonstationary demand and partial information. *Management Science*, 48(5):607–624, 2002.
32. Arthur F. Veinott. On the optimality of (s, S) inventory policies: New conditions and a new proof. *SIAM Journal on Applied Mathematics*, 14(5):1067–1083, 1966.
33. Artur F. Veinott and Harvey M. Wagner. Computing optimal (s,S) inventory policies. *Management Science*, 11(5):525–552, 1966.
34. H. Vieira, S. M. Sanchez, P. J. Sanchez, and K.H. Kienitz. A restricted multinomial hybrid selection procedure. *ACM Transactions on Modeling and Computer Simulation*, 24(2):Article No. 10, 2014.
35. Harvey M. Wagner and Thomson M. Whitin. Dynamic version of the economic lot size model. *Management Science*, 5(1):89–96, 1958.
36. G. Wu, E. K. P. Chong, and R. L. Givan. Burst-level congestion control using hindsight optimization. *IEEE Transactions on Automatic Control, special issue on Systems and Control Methods for Communication Networks*, 47(6):979–991, 2002.
37. Chen Yao and C. G. Cassandras. A solution to the optimal lot sizing problem as a stochastic resource contention game. *IEEE Trans. on Automation Science and Engineering*, 9(2):250–264, 2012.
38. Willard I. Zangwill. A deterministic multi-period production scheduling model with backlogging. *Management Science*, 13(1):105–119, 1966.
39. Y. Zheng and A. Federgruen. Finding optimal (s,S) policies is about as simple as evaluating a single policy. *Operations Research*, 39(4):654–665, 1991.
40. Y. S. Zheng. A simple proof for optimality of (s,S) policies in infinite-horizon inventory systems. *Journal of Applied Probability*, 28:802–810, 1991.
41. P. H. Zipkin. *Foundations of Inventory Management*. McGraw-Hill, Singapore, 2000.