Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 17, No. 4, pp. 681–694, 2021

A SAMPLE AVERAGE APPROXIMATION METHOD BASED ON A GAP FUNCTION FOR STOCHASTIC MULTIOBJECTIVE OPTIMIZATION PROBLEMS*

Yong Zhao, Lin Chen and Xinmin Yang†

Abstract: In this paper, we consider the sample average approximation method for stochastic multiobjective optimization problems without the scalarization parameters. By virtue of the gap function, we transform stochastic multiobjective optimization problems into stochastic optimization reformulation problems. Some properties of the reformulation problems are discussed. Then, we propose a sample average approximation method for solving the reformulation problems, and the convergence and the rates of convergence of optimal values and optimal solutions of the approximation problems are investigated. Furthermore, the rates of convergence of the weakly Pareto optimal for sample average approximation multiobjective problem are discussed under the error bound condition.

Key words: *stochastic multiobjective optimization, gap function, sample average approximation method, convergence*

Mathematics Subject Classification: *90C15, 90C29*

1 Introduction

It is well-known that the multiobjective optimization form an important area of research in optimization theory and operational research for its significant applications in engineering, economics and social sciences, management and machine learning. In practice, multiobjective optimization problems often involve some stochastic data and this motivates one to consider stochastic multiobjective optimization problems.

In this paper, we consider the following stochastic multiobjective optimization problems

$$\begin{array}{ll} \min & F(x) = (\mathbb{E}[f_1(x,\xi(\omega))],\cdots,\mathbb{E}[f_m(x,\xi(\omega))]) \\ \text{s.t.} & x \in X, \end{array} \tag{1.1}$$

where X is a nonempty, compact and convex subset of $\mathbb{R}^n$, $f_i(\cdot,\cdot):\mathbb{R}^n\times\mathbb{R}^r\to\mathbb{R}, i=1,\cdots,m$, are real valued functions, $\xi:\Omega\to\Xi$ is a vector of random variables defined on the

*This work was supported by by the NSFC Grant (No. 12001072), the China Postdoctoral Science Foundation Project (No. 2019M653332), the Chongqing Natural Science Foundation Project (No. cstc2019jcyj-zdxmX0016, cstc2019jcyj-bshX0092), the Science and Technology Research Program of Chongqing Municipal Education Commission (No. KJQN201800744), the Group Building Scientific Innovation Project for universities in Chongqing (CXQT21021) and the open project of Key Laboratory No (CSSXKFKTQ202006), School of Mathematical Sciences, Chongqing Normal University.

†Corresponding author.

probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with support set $\Xi \subset \mathbb{R}^r$, and $\mathbb{E}[\cdot]$ denotes the expected value with respect to probability measure $\mathbb{P}$. For simplicity, we use ξ to denote either the random vector $\xi(\omega)$ or an element of $\mathbb{R}^r$ depending on the context.

Stochastic multiobjective optimization is a combination of multiobjective optimization and stochastic optimization, which is widely used in many fields such as economy, management and machine learning [14, 8, 11, 4, 5]. With the rapid development of machine learning, stochastic multiobjective optimization has been intensively investigated by many researchers in the field of optimization. However, compared with multiobjective optimization and stochastic optimization, the methodological and theoretical development of stochastic multiobjective optimization remains in its infancy. Compared to multiobjective optimization where we have access to the complete objective functions, in stochastic multiobjective optimization, only stochastic samples of objective functions are available for optimization. Compared to stochastic optimization, the fundamental challenge of stochastic multiobjective optimization is how to make appropriate tradeoff between different objectives. In particular, we should consider the conflict objective functions and accommodate the uncertainty in objective function in the algorithm design. Therefore, the analytical and algorithmic tools from multiobjective optimization and stochastic optimization cannot directly resolve this kind of problems when the distribution of random variable is unknown or it is difficult to obtain a closed form of the expected value of random functions. Therefore, some new methods are needed to deal with stochastic multiobjective optimization problems.

A popular numerical method in stochastic optimization is the Monte Carlo method where the expected value is approximated by its sample average approximation. Over the past years, sample average approximation method has also been increasingly investigated for solving stochastic multiobjective optimization. Let $\xi_1, \ldots, \xi_N$ be independently identically distributed sample of random variable ξ. It leads to the approximation of (1.1) as follows:

$$\begin{aligned} \min \quad & F^N(x) = \left(\frac{1}{N} \sum_{j=1}^{N} f_1(x, \xi_j), \ldots, \frac{1}{N} \sum_{j=1}^{N} f_m(x, \xi_j) \right) \\ \text{s.t.} \quad & x \in X, \end{aligned} \tag{1.2}$$

which results in a deterministic multiobjective optimization problem which can be solved by a standard solution technique for multiobjective optimization problem. At present, the main method to deal with (stochastic) multiobjective optimization is scalarization method. It converts (stochastic) multiobjective optimization into (stochastic) optimization problems, and then processes them according to (stochastic) optimization problems. Bonnel and Collonge [2] obtain the consistency of the weakly Pareto sets associated with the sample average approximation problem by virtue of the scalarization method, and the consistency results of optimal solutions for sample average approximation problem. Fliege and Xu [3] apply sample average approximation method to solve stochastic multiobjective optimization, and propose a smoothing infinity norm scalarization approach to solve the sample average approximation problem. Further, the convergence results of efficient solution of the sample average approximation problem are discussed. Kim and Ryu [6] apply the sample average approximation method to stochastic multiobjective optimization problems. Then, based on the product formulation scalarization method, they prove the convergence properties under a set of fairly general regularity conditions. Lin, Zhang and Liang [7] consider a class of stochastic multiobjective problems with complementarity constraints. Liu and Liang [10] study the stability analysis of stochastic multiobjective optimization problems with complementarity constraints when the underlying probability measure varies in some metric probability space. Pang, Meng and Wang [12] study asymptotic convergence of stationary points of stochastic multiobjective optimization problems with parametric variational inequality constraint via sample average approximation method. More advanced sample average approximation method for stochastic multiobjective optimization can be found in [5].

Note that for scalarizing the (stochastic) multiobjective optimization problems, some parameters are usually needed to be determined, but decide a suitable scalarization function with parameters needs a great insight into the problem structure, which is a very difficult task and particularly challenging when the complete objective functions are unavailable. In this paper, we first construct the gap function for stochastic multiobjective optimization problems without the scalarization parameters. By virtue of the gap function, we transform stochastic multiobjective optimization problems into a stochastic optimization reformulation problems. Some properties of the reformulation problems are discussed. Then, we propose a sample average approximation method for solving the reformulation problems, and the convergence and the rates of convergence almost surely, in mean and in probability for estimators of optimal values and optimal solutions of the approximation problems are investigated. Furthermore, the rates of convergence of the weakly Pareto optimal for sample average approximation multiobjective problem are discussed under the error bound condition.

2 Stochastic Optimization Reformulation

Throughout this paper, we use the following notation. Let Log(x) denote the function $\max\{1, \log x\}, x \geq 0$, and let LLog$(x)$ stand for Log(Log(x)). Further, we set for $N \in \mathbb{N}$,

$$a_N := \sqrt{2N\text{LLog}(N)} \quad \text{and} \quad b_N := \frac{a_N}{N} = \frac{\sqrt{2\text{LLog}(N)}}{\sqrt{N}}.$$

For each $i = 1, \ldots, m$, we denote $\widetilde{X}_i := f_i(\cdot, \xi) - \mathbb{E}[f_i(\cdot, \xi)]$ and $\sigma(\widetilde{X}_i) := \sup_{x \in X} (\mathbb{E}[\widetilde{X}_i^2(x)])^{\frac{1}{2}}$. Now, we introduce the concept of optimality for the stochastic multiobjective optimization problems (1.1).

Definition 2.1. A feasible solution $x^* \in X$ is said to be Pareto optimal of (1.1), if there is no $x \in X$ such that $F(x) \leq F(x^*)$ and $F(x) \neq F(x^*)$. Likewise, a feasible solution $x^* \in X$ is said to be weakly Pareto optimal of (1.1), if there is no $x \in X$ such that $F(x) < F(x^*)$.

We will denote the set of Pareto (resp. weakly Pareto) optimal of (1.1) by Sol(F, X) (resp.WSol(F, X)).

In the following we give the definition of gap function for (1.1).

Definition 2.2. A gap function for (1.1) is a function $\theta : \mathbb{R}^n \to \mathbb{R}$ such that

(i) $\theta(x) \geq 0$, for all $x \in X$,

(ii) $\theta(x) = 0$ if and only if $x \in$WSol(F, X).

Following [9, 16], we introduce a gap function for (1.1) as follows:

$$\theta(x) := \sup_{y \in X} \min_{i \in \{1,\ldots,m\}} \{\mathbb{E}[f_i(x, \xi)] - \mathbb{E}[f_i(y, \xi)]\}. \tag{2.1}$$

Therefore, in order to find the weakly Pareto optimal of (1.1), we may solve the following stochastic optimization problem:

$$\min_{x \in X} \theta(x). \tag{2.2}$$

We make some assumptions that will be used later on.

(A1) There is a vector $x_0 \in X$ such that $\mathbb{E}[f_i(x_0,\xi)] < \infty$ for each $i = 1, 2, \ldots, m$.

(A2) For each $i = 1, 2, \ldots, m$, there exists a measurable function $\kappa_i : \Xi \to \mathbb{R}_+$ such that $\mathbb{E}[\kappa_i(\xi)] < \infty$ and

$$|f_i(x_1,\xi) - f_i(x_2,\xi)| \leq \kappa_i(\xi)\|x_1 - x_2\|, \forall x_1, x_2 \in X,$$

for almost surely $\xi \in \Xi$.

If the optimal solution x^* of (2.2) exists and satisfies $\theta\,(x^*) > 0$, then x^* is not weakly Pareto optimal from the above conclusions. However, the following lemma shows that the global solutions of (2.2) are always weakly Pareto optimal for (1.1).

Lemma 2.3. *Let $\mathbb{E}[f_i(\cdot,\xi)] : X \longrightarrow \mathbb{R}$ be bounded functions for all $i = 1, \ldots, m$. If $x^* \in X$ is optimal solution for (2.2), then x^* is weakly Pareto optimal for (1.1).*

Proof. Let $f : X \longrightarrow \mathbb{R}$ and $g : X \longrightarrow \mathbb{R}$ be bounded functions. Then, it follows that

$$\sup_{x\in X}\,(f(x) + g(x)) \leq \sup_{x\in X} f(x) + \sup_{x\in X} g(x).$$

Thus, we have

$$\sup_{x\in X}\,(f(x) + g(x)) - \sup_{x\in X} f(x) \leq \sup_{x\in X} g(x). \tag{2.3}$$

Let x^* be an optimal solution for (2.2). Then, $\theta\,(x) \geq \theta\,(x^*)$ for all $x \in X$, which implies

$$\begin{aligned}
0 &\leq \theta(x) - \theta\,(x^*)\\
&= \sup_{y\in X}\min_{i\in\{1,\ldots,m\}}\{\mathbb{E}[f_i(x,\xi)] - \mathbb{E}[f_i(y,\xi)]\} - \sup_{y\in X}\min_{i\in\{1,\ldots,m\}}\{\mathbb{E}[f_i(x^*,\xi)] - \mathbb{E}[f_i(y,\xi)]\}\\
&\leq \sup_{y\in X}\left[\min_{i\in\{1,\ldots,m\}}\{\mathbb{E}[f_i(x,\xi)] - \mathbb{E}[f_i(y,\xi)]\} - \min_{i\in\{1,\ldots,m\}}\{\mathbb{E}[f_i(x^*,\xi)] - \mathbb{E}[f_i(y,\xi)]\}\right],
\end{aligned}$$

where the inequality comes from (2.3). Now, let $i^* \in \operatorname*{arg\,min}_{i\in\{1,\ldots,m\}}\{\mathbb{E}[f_i(x^*,\xi)] - \mathbb{E}[f_i(y,\xi)]\}$. Then, we obtain

$$\min_{i\in\{1,\ldots,m\}}\{\mathbb{E}[f_i(x^*,\xi)] - \mathbb{E}[f_i(y,\xi)]\} = \mathbb{E}[f_{i^*}(x^*,\xi)] - \mathbb{E}[f_{i^*}(y,\xi)],$$

which implies

$$\begin{aligned}
0 &\leq \theta(x) - \theta\,(x^*)\\
&\leq \sup_{y\in X}\left[\min_{i\in\{1,\ldots,m\}}\{\mathbb{E}[f_i(x,\xi)] - \mathbb{E}[f_i(y,\xi)]\} - \{\mathbb{E}[f_{i^*}(x^*,\xi)] - \mathbb{E}[f_{i^*}(y,\xi)]\}\right]\\
&\leq \sup_{y\in X}\{(\mathbb{E}[f_{i^*}(x,\xi)] - \mathbb{E}[f_{i^*}(y,\xi)]) - (\mathbb{E}[f_{i^*}(x^*,\xi)] - \mathbb{E}[f_{i^*}(y,\xi)])\}\\
&\leq \sup_{y\in X}\max_{i\in\{1,\ldots,m\}}\{(\mathbb{E}[f_i(x,\xi)] - \mathbb{E}[f_i(y,\xi)]) - (\mathbb{E}[f_i(x^*,\xi)] - \mathbb{E}[f_i(y,\xi)])\}\\
&= \max_{i\in\{1,\ldots,m\}}\{\mathbb{E}[f_i(x,\xi)] - \mathbb{E}[f_i(x^*,\xi)]\}.
\end{aligned}$$

Therefore, x^* is weakly Pareto optimal for (1.1). □

In the rest of this section, $\mathbb{P}(V)$ denotes the probability of an event V. In the following we provide an error bound of the gap function for the stochastic multiobjective optimization problems (1.1).

Theorem 2.4. *Suppose that each function $f_i(\cdot,\xi)$ $(i=1,\dots,m)$ is convex on X for almost every $\xi\in\Xi$. Let f_i $(i=1,\dots,m)$ be uniformly strongly convex on X with modulus $\mu_i>0$ over $V_i\subset\Xi$ with $\mathbb{P}(V_i)>0,\mu:=\min\limits_{1\le i\le m}\mu_i,\nu:=\min\limits_{1\le i\le m}\mathbb{P}(V_i)$. Then, we have*

$$\operatorname{dist}^2(x,\operatorname{WSol}(F,X))\le\frac{2}{\mu\nu}\theta(x),\forall x\in X. \tag{2.4}$$

Proof. Using the definitions of convexity and strong convexity, for any $x,y\in X,\alpha\in(0,1)$, we have

$$\begin{aligned}\mathbb{E}[f_i(\alpha x+(1-\alpha)y,\xi)]&=\int_{V_i}f_i(\alpha x+(1-\alpha)y,\xi)\mathbb{P}(d\xi)+\int_{\Xi\setminus V_i}f_i(\alpha x+(1-\alpha)y,\xi)\mathbb{P}(d\xi)\\&\le\int_{\Xi}(\alpha f_i(x,\xi)+(1-\alpha)f_i(y,\xi))\mathbb{P}(d\xi)-\frac{\alpha(1-\alpha)\mu_i}{2}\int_{V_i}\|x-y\|^2\mathbb{P}(d\xi)\\&=\alpha\mathbb{E}[f_i(x,\xi)]+(1-\alpha)\mathbb{E}[f_i(y,\xi)]-\frac{\alpha(1-\alpha)\mu_i}{2}\mathbb{P}(V_i)\|x-y\|^2\\&\le\alpha\mathbb{E}[f_i(x,\xi)]+(1-\alpha)\mathbb{E}[f_i(y,\xi)]-\frac{\alpha(1-\alpha)\mu\nu}{2}\|x-y\|^2,\end{aligned}$$

which implies that $\mathbb{E}[f_i(\cdot,\xi)]$ is strongly convex with modulus $\mu\nu>0$ for all $i\in\{1,\dots,m\}$. Then, it follows from Theorem 3.5 in [16] that the conclusion is true. □

Definition 2.5 ([13])**.** Assume that every function of the sequence $\{f_n(x)\}$ is lower semicontinuous and the function $f(x)$ is lower semicontinuous. We say that $\{f_n(x)\}$ epi-converges to f if for any x,

(i) for every sequence $\{x_n\}$ converging to x, it holds $\liminf\limits_{n\to\infty}f_n(x_n)\ge f(x)$;

(ii) there exists a sequence $\{x_n\}$ converging to x such that $\limsup\limits_{n\to\infty}f_n(x_n)\le f(x)$.

Definition 2.6 ([13])**.** Let $\{C_n\}$ be a sequence of closed sets in $\mathbb{R}^n$. The outer limit of $\{C_n\}$ is defined as follows:

$$\operatorname{ls}C_n=\left\{x\mid\exists\{x_{n_k}\}\text{ s.t. }\{x_{n_k}\}\in\{C_{n_k}\},x=\lim_{k\to\infty}x_{n_k}\right\}.$$

3 Convergence of Optimal Values and Optimal Solutions

Since problem (2.2) involves the mathematical expectation in the objective function and the distribution of the random variables may be unknown in practice or it is numerically too expensive to calculate the expected values, we apply sample average approximation techniques to deal with the expected value, and investigate its convergence. Let $\xi_1,\xi_2,\dots,\xi_N$ be independently and identically distributed samples drawn from Ξ. Then, we consider the sample average approximation problem

$$\begin{aligned}\min\quad&\theta_N(x)=\sup_{y\in X}\min_{i\in\{1,\dots,m\}}\left\{\frac{1}{N}\sum_{j=1}^N f_i(x,\xi_j)-\frac{1}{N}\sum_{j=1}^N f_i(y,\xi_j)\right\}\\\text{s.t.}\quad&x\in X,\end{aligned}\tag{3.1}$$

as an approximation to (2.2).

In what follows, we investigate the convergence of the approximation problem (3.1).

Let $f_{i,N}(x):=\frac{1}{N}\sum\limits_{j=1}^N f_i(x,\xi_j),F_i(x):=\mathbb{E}[f_i(x,\xi)],i=1,\dots,m.$

Lemma 3.1. *Assume that assumptions* (A1)-(A2) *hold. Then for each* $i = 1, 2, \dots, m$*, the following results hold:*

(a) $F_i(x)$ *is finite and Lipschitz continuous on* X*;*

(b) $\{f_{i,N}(x)\}$ *uniformly converges to* $F_i(x)$ *on* X*, that is,*

$$\lim_{N\to\infty} \max_{x\in X} |f_{i,N}(x) - F_i(x)| = 0.$$

Proof. From assumptions (A1) and (A2) that we have for each $i = 1, \dots, m$, $F_i(x) < +\infty$ for all $x \in X$. Moreover, (A2) implies that for each $i = 1, \dots, m$, F_i is Lipschitz continuous on X, which proves (a). Since X is a compact set, thus guarantees that (b) then follows from Theorem 7.48 of [15]. □

Theorem 3.2. *Assume that assumptions* (A1)-(A2) *hold. Then* $\theta_N(x)$ *uniformly converges to* $\theta(x)$ *with probability one.*

Proof. By the definitions of $\theta_N(x)$ and $\theta(x)$, we have

$$\begin{aligned}
|\theta_N(x) - \theta(x)| &= \left|\sup_{y\in X} \min_{i\in\{1,\dots,m\}} \{f_{i,N}(x) - f_{i,N}(y)\} - \sup_{y\in X} \min_{i\in\{1,\dots,m\}} \{F_i(x) - F_i(y)\}\right| \\
&\le \left|\sup_{y\in X} \max_{i\in\{1,\dots,m\}} \{f_{i,N}(x) - f_{i,N}(y) - F_i(x) + F_i(y)\}\right| \\
&\le \left|\sup_{y\in X} \left[\max_{i\in\{1,\dots,m\}} \{f_{i,N}(x) - F_i(x)\} + \max_{i\in\{1,\dots,m\}} \{F_i(y) - f_{i,N}(y)\}\right]\right| \\
&\le \max_{i\in\{1,\dots,m\}} |f_{i,N}(x) - F_i(x)| + \sup_{y\in X} \max_{i\in\{1,\dots,m\}} |f_{i,N}(y) - F_i(y)|.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\sup_{x\in X} |\theta_N(x) - \theta(x)| \le& \sup_{x\in X} \max_{i\in\{1,\dots,m\}} |f_{i,N}(x) - F_i(x)| \\
&+ \sup_{y\in X} \max_{i\in\{1,\dots,m\}} |f_{i,N}(y) - F_i(y)|. \qquad (3.2)
\end{aligned}$$

From Lemma 3.1, for each $i \in \{1, 2, \dots, m\}$, $\{f_{i,N}(x)\}$ uniformly converges to $F_i(x)$ on X. Then, for any $\epsilon > 0$, there exists N_0 such that for any $N > N_0$, we have

$$\sup_{x\in X} \max_{i\in\{1,\dots,m\}} |f_{i,N}(x) - F_i(x)| < \tfrac{\epsilon}{2}.$$

Thus, by (3.2), we have

$$\sup_{x\in X} |\theta_N(x) - \theta(x)| < \epsilon,$$

which implies that $\theta_N(x)$ uniformly converges to $\theta(x)$ on X with probability one. □

In the following we shall provide an upper bound of the distance between the weakly Pareto optimal of (1.2) and the true one WSol(F,X) in terms of error bound condition.

Theorem 3.3. *Let* $\{x_N\}$ *be a sequence of weakly Pareto optimal to* (1.2)*. Suppose that assumptions* (A1)-(A2) *hold and the error bound condition* (2.4) *holds with gap function* $\theta(\cdot)$*. Then*

(i) *any accumulation point* x^* *of sequence* $\{x_N\}$ *is weakly Pareto optimal to* (1.1)*;*

(ii)

$$\mathrm{dist}\left(x_N, \mathrm{WSol}(F,X)\right) \le \sqrt{\frac{2}{\mu\nu}}\left(\theta\left(x_N\right) - \theta_N\left(x_N\right)\right)^{\frac{1}{2}}.$$

Proof. (i) Denote x^* as an accumulation point of sequence $\{x_N\}$. From Theorem 3.2 that $\theta_N(x)$ uniformly converges to $\theta(x)$ on X, then $\theta(x^*)$=0 as $\theta_N(x_N) = 0$ and $x_N \to x^*$, which implies x^* is weakly Pareto optimal to (1.1).

(ii) According to the error bound condition (2.4), we have

$$\begin{aligned}\mathrm{dist}\left(x_N, \mathrm{WSol}(F,X)\right) &\le \sqrt{\frac{2}{\mu\nu}}\theta(x_N)^{\frac{1}{2}} \\ &= \sqrt{\frac{2}{\mu\nu}}\left(\theta(x_N) - \theta_N(x_N)\right)^{\frac{1}{2}},\end{aligned}$$

where the equality follows from the fact that, $\theta_N(x_N) = 0$ as x_N is weakly Pareto optimal to (1.2). □

Next we discuss the convergence of optimal values and optimal solutions of (3.1).

Theorem 3.4. *Assume that assumptions* (A1)-(A2) *hold, then*

$$\lim_{N\to\infty} \min_{x\in X} \theta_N(x) = \min_{x\in X} \theta(x),$$

and

$$\mathrm{ls}\{\mathrm{argmin}_{x\in X}\theta_N(x)\} \subset \mathrm{argmin}_{x\in X}\theta(x).$$

Proof. It follows from assumptions (A1) and (A2) that $\theta_N(x)$ epi-converges to $\theta(x)$ on X with probability one, and the remaining assertion then follows from Theorem 7.33 in [13]. □

4 Almost Sure Rates of Convergence of Optimal Values and Optimal Solutions

In this section, we discuss the rate of convergence in the almost sure sense for estimators of optimal values and optimal solutions of the approximation problem (3.1). From now on, we denote by θ^* and θ_N^* the optimal values of problems (2.2) and (3.1), respectively.

Lemma 4.1. *Suppose that assumptions* (A1)-(A2) *hold. Then, for any $\epsilon > 0$ and for each $i = 1, \ldots, m$, there exists $N_i^* = N_i^*(\epsilon)$ such that*

$$\sup_{x\in X} |f_{i,N}(x) - F_i(x)| \le (1+\epsilon) b_N \sigma(\widetilde{X}_i), \forall N \ge N_i^*,$$

almost surely. Furthermore, let $\sigma = \max_{i\in\{1,\ldots,m\}} \sigma(\widetilde{X}_i)$, *then*

$$\sup_{x\in X} \max_{i\in\{1,\ldots,m\}} |f_{i,N}(x) - F_i(x)| \le (1+\epsilon) b_N \sigma, \forall N \ge N^*, \tag{4.1}$$

where $N^* = \max_{i\in\{1,\ldots,m\}} N_i^*$.

Proof. It follows from Lemma 1 in [1] that for any $\epsilon > 0$ and for each $i = 1, \dots, m$, there exists N_i^* such that

$$\sup_{x \in X} |f_{i,N}(x) - F_i(x)| \le (1+\epsilon) b_N \sigma(\widetilde{X}_i), \forall N \ge N_i^*.$$

Therefore, for any $x \in X$ and for each $i = 1, \dots, m$, we have $|f_{i,N}(x) - F_i(x)| \le (1 + \epsilon) b_N \sigma(\widetilde{X}_i), \forall N \ge N_i^*$. This implies that for any $x \in X$,

$$\max_{i \in \{1,\dots,m\}} |f_{i,N}(x) - F_i(x)| \le (1+\epsilon) b_N \max_{i \in \{1,\dots,m\}} \sigma(\widetilde{X}_i), \forall N \ge N^*.$$

Therefore, we have

$$\sup_{x \in X} \max_{i \in \{1,\dots,m\}} |f_{i,N}(x) - F_i(x)| \le (1+\epsilon) b_N \sigma, \forall N \ge N^*.$$

□

Theorem 4.2. *Suppose that assumptions* (A1)-(A2) *hold. Then, for any $\epsilon > 0$, there exists N^* such that*

$$|\theta_N^* - \theta^*| \le 2(1+\epsilon) b_N \sigma, \forall N \ge N^*,$$

almost surely.

Proof. It follows from Lemma 4.1 and (3.2) that

$$|\theta_N^* - \theta^*| \le \sup_{x \in X} |\theta_N(x) - \theta(x)| \le 2 \sup_{x \in X} \max_{i \in \{1,\dots,m\}} |f_{i,N}(x) - F_i(x)| \le 2(1+\epsilon) b_N \sigma,$$

almost surely. □

Remark 4.3. Theorem 4.2 implies that the optimal value θ_N^* converges to θ^* almost surely at a rate of $\mathcal{O}(b_N)$.

Theorem 4.4. *Let $\{x_N\}$ be a sequence of solutions to* (3.1)*. Suppose that assumptions* (A1)-(A2) *hold. Then, for any $\epsilon > 0$, there exists N^* such that*

$$|\theta(x_N) - \theta^*| \le 4(1+\epsilon) b_N \sigma, \forall N \ge N^*,$$

almost surely.

Proof. Note that

$$\begin{aligned} |\theta(x_N) - \theta^*| &\le |\theta(x_N) - \theta_N(x_N)| + |\theta_N(x_N) - \theta^*| \\ &\le 2 \sup_{x \in X} |\theta_N(x) - \theta(x)|. \end{aligned} \tag{4.2}$$

Therefore, it follows from Lemma 4.1 and (3.2) that

$$|\theta(x_N) - \theta^*| \le 4 \sup_{x \in X} \max_{i \in \{1,\dots,m\}} |f_{i,N}(x) - F_i(x)| \le 4(1+\epsilon) b_N \sigma,$$

almost surely. □

Remark 4.5. Theorem 4.4 establishes the sequence of optimal solutions of the approximate problem (3.1) converges to an optimal solution of (2.2) almost surely at a rate of $\mathcal{O}(b_N)$.

5 Rates of Convergence of Optimal Values and Optimal Solutions in Mean

In the following we establish the rates of convergence in mean for estimators of optimal values and optimal solutions of the approximation problem (3.1). It follows from Proposition 2 in [1] that we have the following result.

Lemma 5.1. *Suppose that assumptions* (A1)-(A2) *hold. Then, for each* $i = 1, \ldots, m$, *we have*

$$\lim_{N\to\infty} \mathbb{E}\left[\frac{\sup_{x\in X} \|f_{i,N}(x) - F_i(x)\|}{b_N}\right] = 0,$$

i.e. $\mathbb{E}\big[\sup_{x\in X} \|f_{i,N}(x) - F_i(x)\|\big] = \mathbf{o}(b_N)$, *and* $\{f_{i,N}\}$ *converges to* F_i *at a rate of* $\mathbf{o}(b_N)$.

Remark 5.2. Lemma 5.1 implies that for any $\epsilon > 0$, there exists N^* such that

$$\frac{\mathbb{E}\Big[\sup_{x\in X} \|f_{i,N}(x) - F_i(x)\|\Big]}{b_N} < \epsilon, \ \forall\, i \in \{1, \ldots, m\}, N \geq N^*. \tag{5.1}$$

By the Proof of Theorem 4.2 and Lemma 5.1, we immediately obtain the following result for the convergence of optimal values.

Theorem 5.3. *Suppose that assumptions* (A1)-(A2) *hold. Then*

$$\mathbb{E}[|\theta_N^* - \theta^*|] = \mathbf{o}(b_N).$$

Remark 5.4. Theorem 5.3 states that θ_N^* is an asymptotically unbiased estimator of θ^*. It follows from Proposition 5.6 in [15] that $\mathbb{E}[\theta_N^*] \leq \mathbb{E}[\theta_{N+1}^*] \leq \theta^*$ for any N, which can be combined with the above result to obtain that for any $\epsilon > 0$, there exists N^* such that

$$\mathbb{E}[\theta_N^*] \leq \mathbb{E}[\theta_{N+1}^*] \leq \theta^* \leq \mathbb{E}[\theta_N^*] + \epsilon b_N, \forall N \geq N^*,$$

which implies that the optimal value θ^* is in an interval of known size.

By the Proof of Theorem 4.4 and Lemma 5.1, the following result shows that the sequence of optimal solutions of the approximate problem (3.1) converges to an optimal solution of (2.2) at a rate of $\mathbf{o}(b_N)$.

Theorem 5.5. *Let* $\{x_N\}$ *be a sequence of solutions to* (3.1). *Suppose that assumptions* (A1)-(A2) *hold. Then*

$$\mathbb{E}[|\theta(x_N) - \theta^*|] = \mathbf{o}(b_N).$$

6 Rates of Convergence of Optimal Values and Optimal Solutions in Probability

In this section, under some mild conditions, we shall infer the rates of convergence in probability for optimal values and optimal solutions of the approximation problem (3.1) by virtue of the obtained rates of convergence in mean. From Theorem 5.3 that we have the following result.

Proposition 6.1. *Suppose that assumptions* (A1)-(A2) *hold and let* $\delta > 0$ *be arbitrary. Then, we have*

$$\mathbb{P}\left(\frac{|\theta_N^* - \theta^*|}{b_N} > \delta\right) \to 0, \text{ as } N \to \infty.$$

In the following we discuss the rates of convergence in probability for optimal values and optimal solutions.

Theorem 6.2. *Suppose that assumptions* (A1)-(A2) *hold and let* $\delta > 0$ *be arbitrary. Then, the following results hold:*

(i) *For any* $\epsilon > 0$*, there exists* N^* *such that*

$$\mathbb{P}\Big(|\theta_N^* - \theta^*| \geq \delta\Big) \leq \frac{2\epsilon b_N}{\delta}, \text{ as } N \geq N^*.$$

(ii) *If* $\mathbb{E}[\|\widetilde{X}_i\|^s] < \infty$ *for* $s > 2$ *and for* $i = 1, \ldots, m$*, then there exists* N^* *such that for all* $N \geq N^*$

$$\mathbb{P}\Big(|\theta_N^* - \theta^*| \geq \delta\Big) \leq \exp\Big\{ - \frac{N\delta^2}{48\sigma^2}\Big\} + \frac{c}{N^{s-1}(\frac{\delta}{4})^s} \max_{1\leq i\leq m} \mathbb{E}[\|\widetilde{X}_i\|^s], \tag{6.1}$$

where c *is a positive constant. Moreover, let* $\{x_N\}$ *be a sequence of solutions to* (3.1)*, then with probability* $1 - \Big\{ - \frac{N\delta^2}{48\sigma^2}\Big\} - \frac{c}{N^{s-1}(\frac{\delta}{4})^s} \max_{1\leq i\leq m} \mathbb{E}[\|\widetilde{X}_i\|^s]$*,* x_N *becomes an approximate optimal solution of* (2.2)*.*

Proof. (i) For any $\delta > 0$, we have

$$\begin{aligned}\mathbb{P}\Big(|\theta_N^* - \theta^*| \geq \delta\Big) &\leq \mathbb{P}\Big(\sup_{x\in X} |\theta_N(x) - \theta(x)| \geq \delta\Big) \\ &\leq \mathbb{P}\Big(\sup_{x\in X} \max_{i\in\{1,\ldots,m\}} |f_{i,N}(x) - F_i(x)| \geq \frac{\delta}{2}\Big) \\ &\leq \frac{2\epsilon b_N}{\delta},\end{aligned}$$

where the third inequality follows from Markov's inequality and inequality (5.1).

(ii) It follows from Theorem 8 of [1] that for each $i = 1, \ldots, m$, there exists N_i^* such that for all $N \geq N_i^*$, we have

$$\mathbb{P}\Big(\sup_{x\in X} |f_{i,N}(x) - F_i(x)| \geq \delta\Big) \leq \exp\Big\{ - \frac{N\delta^2}{12\sigma^2(\widetilde{X}_i)}\Big\} + \frac{c}{N^{s-1}(\frac{\delta}{2})^s}\mathbb{E}[\|\widetilde{X}_i\|^s].$$

This implies that for all $N \geq N^* := \max_{i\in\{1,\ldots,m\}} N_i^*$, we have

$$\mathbb{P}\Big(\sup_{x\in X} \max_{i\in\{1,\ldots,m\}} |f_{i,N}(x) - F_i(x)| \geq \delta\Big) \leq \exp\Big\{-\frac{N\delta^2}{12\sigma^2}\Big\} + \frac{c}{N^{s-1}(\frac{\delta}{2})^s} \max_{1\leq i\leq m} \mathbb{E}[\|\widetilde{X}_i\|^s]. \tag{6.2}$$

We then obtain (6.1) similar to (i). If x_N is a solution to (3.1), then it follows from (4.2) and (6.2) that we have

$$\begin{aligned}\mathbb{P}\Big(|\theta(x_N) - \theta^*| \geq 4\delta\Big) &\leq \mathbb{P}\Big(\sup_{x\in X} \max_{i\in\{1,\ldots,m\}} |f_{i,N}(x) - F_i(x)| \geq \delta\Big) \\ &\leq \exp\Big\{ - \frac{N\delta^2}{12\sigma^2}\Big\} + \frac{c}{N^{s-1}(\frac{\delta}{2})^s} \max_{1\leq i\leq m} \mathbb{E}[\|\widetilde{X}_i\|^s],\end{aligned}$$

which implies that, with probability $1-\left\{-\frac{N\delta^2}{48\sigma^2}\right\}-\frac{c}{N^{s-1}(\frac{\delta}{4})^s}\max_{1\leq i\leq m}\mathbb{E}[\|\widetilde{X}_i\|^s]$, an optimal solution of (3.1) becomes a 4δ-approximate optimal solution of (2.2). □

Remark 6.3. The results of exponential rates of convergence in probability are obtained by a large deviation principle under strong exponential moment conditions (or boundedness condition) in [10, 3]. However, we obtain the exponential rates of convergence without a strong exponential moment conditions on the random variable are available. Therefore, the results improve the corresponding results in [10, 3].

7 Rates of Convergence of the Weakly Pareto Optimal for (1.2)

Next, we will discuss the rates of convergence of the weakly Pareto optimal of the approximate problem (1.2) under the error bound condition.

Theorem 7.1. *Let $\{x_N\}$ be a sequence of weakly Pareto optimal to* (1.2). *Suppose that assumptions* (A1)-(A2) *hold and the error bound condition* (2.4) *holds with gap function $\theta(\cdot)$. Then, for any $\epsilon>0$, there exists N^* such that*

$$\operatorname{dist}^2(x_N,\mathrm{WSol}(F,X))\leq\frac{4}{\mu\nu}(1+\epsilon)b_N\sigma,\forall N\geq N^*,$$

almost surely.

Proof. It follows from Theorem 3.3 and Theorem 4.2 that we have

$$\begin{aligned}\operatorname{dist}^2(x_N,\mathrm{WSol}(F,X))&\leq\frac{2}{\mu\nu}\left(\theta(x_N)-\theta_N(x_N)\right)\\&\leq\frac{2}{\mu\nu}\sup_{x\in X}|\theta_N(x)-\theta(x)|\leq\frac{4}{\mu\nu}(1+\epsilon)b_N\sigma,\end{aligned}$$

almost surely. □

Remark 7.2. Theorem 7.1 implies that the distance to the weakly efficient sets WSol(F,X) diminish almost surely at the rate of $\mathcal{O}(\sqrt{b_N})$.

If assumptions (A1)-(A2) are satisfied together with the error bound condition (2.4), then it is easy to obtain the rate of convergence of the weakly Pareto optimal to (1.2) in mean.

Theorem 7.3. *Let $\{x_N\}$ be a sequence of weakly Pareto optimal to* (1.2). *Suppose that assumptions* (A1)-(A2) *hold and the error bound condition* (2.4) *holds with gap function $\theta(\cdot)$. Then*

$$\mathbb{E}[\operatorname{dist}^2(x_N,\mathrm{WSol}(F,X))]\to 0,\ as\ N\to\infty,$$

and this implies that $\mathbb{E}[\operatorname{dist}^2(x_N,\mathrm{WSol}(F,X))]$ vanishes at the rate of $\mathbf{o}(b_N)$.

Proof. It follows from Theorem 7.1 that we have $\operatorname{dist}^2(x_N,\mathrm{WSol}(F,X))\to 0$, as $N\to\infty$, almost surely. Further, due to the compactness of X, we have $\operatorname{dist}^2(x_N,\mathrm{WSol}(F,X))\leq\operatorname{diam}^2(X)$. By Lebesgue's dominated convergence theorem, we obtain $\mathbb{E}[\operatorname{dist}^2(x_N,\mathrm{WSol}(F,X))]\to 0,\text{as }N\to\infty$. The remaining statements follows from Lemma 5.1 and Theorem 7.1. □

By the Proof of Theorem 6.2, it is easy to obtain the exponential rates of convergence of the weakly Pareto optimal to (1.2).

Theorem 7.4. *Let* $\{x_N\}$ *be a sequence of weakly Pareto optimal to* (1.2)*. Suppose that assumptions* (A1)-(A2) *hold and let* $\delta > 0$ *be arbitrary. Then, the following results hold:*

(i) *If the error bound condition* (2.4) *holds with gap function* $\theta(\cdot)$*, then for any* $\epsilon > 0$*, we have*

$$\mathbb{P}\Big(\operatorname{dist}^2(x_N, \mathrm{WSol}(F,X)) \geq \delta\Big) \leq \frac{4\epsilon b_N}{\delta\mu\nu}, \text{ as } N \geq N^*.$$

(ii) *If the error bound condition* (2.4) *holds with gap function* $\theta(\cdot)$*, and* $\mathbb{E}[\|\widetilde{X}_i\|^s] < \infty$ *for* $s > 2$ *and* $i = 1, \ldots, m$*, then there exists* N^* *such that for all* $N \geq N^*$

$$\mathbb{P}\Big(\operatorname{dist}^2(x_N, \mathrm{WSol}(F,X)) \geq \delta\Big) \leq \exp\Big\{-\frac{N\delta^2\mu^2\nu^2}{192\sigma^2}\Big\} + \frac{c}{N^{s-1}(\frac{\delta\mu\nu}{8})^s} \max_{1\leq i\leq m} \mathbb{E}[\|\widetilde{X}_i\|^s].$$

Proof. (i) From the proof of Theorem 7.1, we have

$$\begin{aligned}\mathbb{P}\Big(\operatorname{dist}^2(x_N, \mathrm{WSol}(F,X)) \geq \delta\Big) &\leq \mathbb{P}\Big(\theta(x_N) - \theta_N(x_N) \geq \frac{\mu\nu\delta}{2}\Big)\\ &\leq \mathbb{P}\Big(\sup_{x\in X}\max_{i\in\{1,\ldots,m\}} |f_{i,N}(x) - F_i(x)| \geq \frac{\mu\nu\delta}{4}\Big) \leq \frac{4\epsilon b_N}{\delta\mu\nu}.\end{aligned}$$

(ii) Similar to the proof of Theorem 6.2 (ii), it is easy to obtain the result. □

8 Conclusions

In this paper, based on gap function, we transform stochastic multiobjective optimization problems into the stochastic optimization reformulation problems without the scalarization parameters. Then, we propose a sample average approximation method for solving the reformulation problems, and the convergence and the rates of convergence of optimal values and optimal solutions of the approximation problem are discussed. It is interesting to apply the results obtained in this paper to practical problems (such as transportation network equilibrium problem). We will investigate this topic in the near future.

References

[1] D. Banholzer, J. Fliege and R. Werner, On rates of convergence for sample average approximations in the almost sure sense and in mean, *Math. Program.* Doi.org/10.1007/s10107-019-01400-4, 2019.

[2] H. Bonnel and C. Collonge, Stochastic optimization over a Pareto set associated with a stochastic multi-objective optimization problem, *J. Optim. Theory Appl.* 162 (2014) 405–427.

[3] J. Fliege and H.F. Xu, Stochastic multiobjective optimization: sample average approximation and applications, *J. Optim. Theory Appl.* 151 (2011) 135–162.

[4] W.J. Gutjahr and A. Pichler, Stochastic multi-objective optimization: a survey on non-scalarizing methods, *Ann. Oper. Res.* 236 (2016) 475–499.

[5] S.R. Hunter, E.A. Applegate, V. Arora, B. Chong, K. Cooper, O.R. Guevara and C.V. Valencia, An introduction to multiobjective simulation optimization, *ACM Trans. Model. Comput. Simul.* 29 (2019) 1–36.

[6] S. Kim and J. Ryu, The sample average approximation method for multi-objective stochastic optimization, in: *Proceedings of the 2011 Winter Simulation Conference*, S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu (Eds.). IEEE, Piscataway, NJ, 2011, pp. 4026–4037.

[7] G.H. Lin, D.L. Zhang and Y.C. Liang, Stochastic multiobjective problems with complementarity constraints and applications in healthcare management, *Eur. J. Oper. Res.* 226 (2013) 461–470.

[8] X. Lin, H.L. Zhen, Z.H. Li, Q.F. Zhang and S. Kwong, Pareto multi-task learning, in: *Advances in Neural Information Processing Systems*, 2019, pp. 12037–12047.

[9] C.G. Liu, K.F. Ng and W.H. Yang, Merit functions in vector optimization, *Math. Program.* 119 (2009) 215–237.

[10] Y.C. Liu and Y.C. Liang, Stability analysis of stochastic multiobjective optimization problems with complementarity constraints, *Pac. J. Optim.* 12 (2016) 857–872.

[11] M. Mahdavi, T. B. Yang and R. Jin, Stochastic convex optimization with multiple objectives, in: *Advances in Neural Information Processing Systems*, 2013, pp. 1115–1123.

[12] L.P. Pang, F.Y. Meng and J.H. Wang, Asymptotic convergence of stationary points of stochastic multiobjective programs with parametric variational inequality constraint via SAA approach, *J. Ind. Mange. Optim.* 15 (2019) 1653–1675.

[13] R.T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 2009.

[14] O. Sener and V. Koltun, Multi-task learning as multi-objective optimization, in: *Advances in Neural Information Processing Systems*, 2018, pp. 527–538.

[15] A. Shapiro, D. Dentcheva and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia, 2009.

[16] H. Tanabe, Merit functions for multiobjective optimization and convergence rates analysis of multiobjective proximal gradient methods, Kyoto University, 2019.

Manuscript received 22 May 2021
revised 10 August 2021
accepted for publication 5 September 2021

Yong Zhao
College of Mathematics and Statistics
Chongqing University, Chongqing 401331, China and;
College of Mathematics and Statistics
Chongqing Jiaotong University, Chongqing 400074, China
E-mail address: zhaoyongty@126.com

Lin Chen
School of Mathematical Sciences
Chongqing Normal University, Chongqing 401331, China
E-mail address: chinaallenchen@126.com

Xinmin Yang
School of Mathematical Sciences
Chongqing Normal University, Chongqing 401331, China
E-mail address: xmyang@cqnu.edu.cn