

Article

Deterministic Bi-Criteria Model for Solving Stochastic Mixed Vector Variational Inequality Problems

Meiju Luo , Menghan Du and Yue Zhang

School of Mathematics and Statistics, Liaoning University, Shenyang 110036, China

* Correspondence: meijuluolu@126.com

Abstract: In this paper, we consider stochastic mixed vector variational inequality problems. Firstly, we present an equivalent form for the stochastic mixed vector variational inequality problems. Secondly, we present a deterministic bi-criteria model for giving the reasonable resolution of the stochastic mixed vector variational inequality problems and further propose the approximation problem for solving the given deterministic model by employing the smoothing technique and the sample average approximation method. Thirdly, we obtain the convergence analysis for the proposed approximation problem while the sample space is compact. Finally, we propose a compact approximation method when the sample space is not a compact set and provide the corresponding convergence results.

Keywords: stochastic; mixed vector variational inequality; deterministic bi-criteria model; sample average approximation; compact approximation

MSC: 90C33; 90C15



Citation: Luo, M.; Du, M.; Zhang, Y. Deterministic Bi-Criteria Model for Solving Stochastic Mixed Vector Variational Inequality Problems. *Mathematics* **2023**, *11*, 3376. <https://doi.org/10.3390/math11153376>

Academic Editor: Simeon Reich

Received: 21 May 2023

Revised: 26 July 2023

Accepted: 28 July 2023

Published: 2 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Let $D \subseteq \mathbb{R}^n$ be a nonempty, closed, and convex set, where $\mathbb{F}_j(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ($j = 1, 2, \dots, m$) are continuously differentiable vector-valued functions. For every $j = 1, 2, \dots, m$, $g_j(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is a continuously differentiable, proper, and convex function with closed effective domain. The mixed vector variational inequality problems (MVVIP) are defined as $x^* \in D$, for all $y \in D$, satisfying

$$(y - x^*)^T \mathbb{F}(x^*) + g(y) - g(x^*) \notin -\text{int} \mathbb{R}_+^m, \quad (1)$$

where $\mathbb{F}(x) := (\mathbb{F}_1(x), \dots, \mathbb{F}_m(x))$ and $g(\cdot) := (g_1(\cdot), \dots, g_m(\cdot))$. Obviously, for every $y \in D$, problem (1) is equivalent to finding $x^* \in D$, satisfying

$$\left((y - x^*)^T \mathbb{F}_1(x^*), \dots, (y - x^*)^T \mathbb{F}_m(x^*) \right) + (g_1(y), \dots, g_m(y)) - (g_1(x^*), \dots, g_m(x^*)) \notin -\text{int} \mathbb{R}_+^m. \quad (2)$$

The above MVVIP can be used to deal with the following multiobjective optimization problems, which originated from the study of utility theory in economics and were first proposed in the economic balance [1]:

$$\begin{aligned} \min \quad & P(x) = (p_1(x), \dots, p_m(x)) \\ \text{s.t.} \quad & q_j(x) \leq 0, \quad j = 1, \dots, t, \end{aligned} \quad (3)$$

where $p_j(j = 1, \dots, m) : \mathbb{R}^n \rightarrow \mathbb{R}$ and $q_j(j = 1, \dots, t) : \mathbb{R}^n \rightarrow \mathbb{R}$ are all continuously differentiable convex functions on $\mathbb{R}^n$. In [2], by using the exponential penalty method, the authors obtained the penalty approximation problem of (3) as follows:

$$\min \left(p_1(x), \dots, p_m(x) \right) + \left(\frac{1}{\rho_n} \sum_{j=1}^t \vartheta[\rho_n q_j(x)], \dots, \frac{1}{\rho_n} \sum_{j=1}^t \vartheta[\rho_n q_j(x)] \right), \quad (4)$$

where $\rho_n > 0$ is a penalty parameter and satisfying $\lim_{n \rightarrow \infty} \rho_n = +\infty$. $\vartheta(t) = \exp(t - 1)$ is an exponential function. It is also proved in [2] that the efficient solution of (3) can be approximated by solving the weakly efficient solution of problem (4). It is worth noting that problem (4) is an unconstrained multiobjective optimization problem and function $\frac{1}{\rho_n} \sum_{j=1}^t \vartheta[\rho_n q_j(x)]$ is a convex function.

In fact, solving the weakly efficient solution of problem (4) is equivalent to solving MVVIP (2) (the proof of which can be found in Theorem 1), which indicates that, for a sufficiently large penalty parameter, the solution of MVVIP can be used to approximate the efficient solution of problem (3). Multiobjective optimization problems can be applied to artificial intelligence, engineering, electrical engineering, and medicine [3–6]. In addition, multiobjective optimization problems also have wide applications in fields such as design optimization, manufacturing, structural health monitoring [7], and chemical engineering [8]. Therefore, MVVIP also have a wide range of applications in production and life, which have certain research significance.

Due to the existence of many uncertain factors, such as price, output, and sales, in actual production and life, in this paper, we mainly consider MVVIP with uncertain factors, that is, stochastic mixed vector variational inequality problems (SMVIP): finding $x^* \in D$ satisfies

$$(y - x^*)^T \mathbb{F}(x^*, \xi(\omega)) + g(y, \xi(\omega)) - g(x^*, \xi(\omega)) \notin -\text{int} \mathbb{R}_+^m, \quad \forall y \in D, \text{ a.s. } \xi(\omega) \in \Xi, \quad (5)$$

where $\xi : \Omega \rightarrow \Xi \subset \mathbb{R}^b$ is a stochastic vector defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and Ξ is a support set of the probability space. Under the given probability measure, “a.s.” stands the abbreviation for “almost surely”. For each $j = 1, 2, \dots, m$, $\mathbb{F}_j(x, \xi(\omega)) : \mathbb{R}^n \times \Xi \rightarrow \mathbb{R}^n$ is continuously differentiable vector-valued function at $(x, \xi(\omega))$ and $g_j(x, \xi(\omega))$ is continuously differentiable on $\mathbb{R}^n \times \Xi$. Here, $\mathbb{F}(x, \xi(\omega)) := (\mathbb{F}_1(x, \xi(\omega)), \dots, \mathbb{F}_m(x, \xi(\omega)))$ and $g(\cdot, \xi(\omega)) := (g_1(\cdot, \xi(\omega)), \dots, g_m(\cdot, \xi(\omega)))$. For simplicity, we use ξ to denote either the stochastic vector $\xi(\omega)$ or an element of $\mathbb{R}^b$ throughout this paper.

To the best of our knowledge, the SMVIP have not yet been studied. However, there are a few works on the special cases of MVVIP. For instance, Salahuddin et al. [9] and Xie et al. [10] discussed the MVVIP in topological vector spaces and established some existence Theorems. Moreover, Irfan [11] considered a new class of exponential MVVIP involving multi-valued mappings in a fuzzy environment and proved some existence theorems of solutions of exponential MVVIP. In addition, a gap function of a non-smooth mixed weak vector variational inequality problem is given by Jayswal et al. [12]. For more information about the existence of MVVIP solutions and the gap function error bounds, we refer the readers to [13–15] and the references therein.

In fact, the SMVIP is obviously a generalization of the stochastic variational inequality problems (SVIP): finding $x^* \in D$ satisfies

$$(y - x^*)^T \mathcal{F}(x^*, \xi(\omega)) \geq 0, \quad \forall y \in D, \text{ a.s. } \omega \in \Omega,$$

where $\xi : \Omega \rightarrow \Xi \subset \mathbb{R}^b$ is a stochastic vector, $\mathcal{F}(x, \xi(\omega)) : \mathbb{R}^n \times \Xi \rightarrow \mathbb{R}^n$ is a vector-valued function, and D is a nonempty closed convex set. Based on Fukushima’s work in [16], for a giving a parameter $\alpha > 0$, Luo [17] defined a regularized gap function as follows:

$$v(x, \xi(\omega)) := \max_{y \in D} \left\{ (x - y)^T \mathcal{F}(x, \xi(\omega)) - \frac{\alpha}{2} \|x - y\|^2 \right\}.$$

The SVIP have been considered in many articles. In particular, Shanbhag [18] considered the applications of SVIP in the power market and radio system design. Moreover,

Luo et al. [17] formulated SVIP as a minimization problem, and presented the expected residual minimization (ERM) model for SVIP by taking the gap function as the residual. In addition, Chen et al. [19] investigated the SVIP from a standpoint of minimization of conditional value at risk (CVaR); that is, the authors employed the gap function to define a loss function and presented a deterministic CVaR model for solving SVIP. For more research on SVIP, interested readers can refer to [20–24].

The main contributions of this paper are as follows:

- We present a reasonable deterministic bi-criteria model for solving SMVVIP, which can be regarded as a combination of the ERM model and CVaR model.
- We employ the smoothing technique and the sample average approximation method to present approximation problems for solving the given deterministic model. In addition, we consider the convergence of the proposed approximation problem when the sample space is a compact set.
- When the sample space is not compact, we also propose a compact approximation method and provide convergence results.

We adopt the following standard notations throughout this paper. For a given nonempty set $\mathbb{R}_+^m$, $\text{int}\mathbb{R}_+^m$ denotes its interior. For a given vector-valued function $\varphi : \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^t \rightarrow \mathbb{R}^m$, $\nabla_x \varphi(x, y, z) \in \mathbb{R}^{m \times n}$ denotes its Jacobian matrix at x . $\nabla_{(x,y)} \varphi(x, y, z) \in \mathbb{R}^{m \times (n+r)}$ denotes its Jacobian matrix at (x, y) . $\nabla_{(x,y,z)} \varphi(x, y, z) \in \mathbb{R}^{m \times (n+r+t)}$ denotes its Jacobian matrix at (x, y, z) . For a given nonempty set $D \subseteq \mathbb{R}^n$, $N_D(x)$ denotes its normal cone operator at $x \in D$ and $\text{conv}\{D\}$ stands for the convex hull of D . For a given real-valued function $G : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, $\partial G(x)$ denotes its subdifferential operator at x . Moreover, $\|\cdot\|$ stands for the 2-norm and $[t]_+ := \max\{t, 0\}$, $\Lambda := \{\lambda \in \mathbb{R}^m \mid \lambda_j \geq 0, \sum_{j=1}^m \lambda_j = 1\}$.

2. Preliminaries

Definition 1 ([25]). The normal cone operator for a nonempty closed convex subset $D \subseteq \mathbb{R}^n$ at $x \in D$ is defined as

$$N_D(x) := \{z \in \mathbb{R}^n \mid (y - x)^T z \leq 0, \forall y \in D\}.$$

Definition 2 ([25]). The definition of the subdifferential of a convex function $G : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ at x is

$$\partial G(x) = \{\tau \in \mathbb{R}^n \mid G(y) > G(x) + (y - x)^T \tau, \forall y \in \mathbb{R}^n\}.$$

Definition 3 ([26]). For a function $G : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ and a parameter $\alpha > 0$, the mapping $P_D^\alpha : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the proximal operator defined as

$$P_D^\alpha(z) = \arg \min_{y \in D} \left\{ G(y) + \frac{\alpha}{2} \|z - y\|^2 \right\}, \quad z \in \mathbb{R}^n,$$

where $D \subseteq \mathbb{R}^n$ is a nonempty closed convex subset.

According to reference [27], when G is proper and convex, the proximal operator has many special properties. For example, $P_D^\alpha(x)$ is single-valued [26] and the proximal operator satisfies the following non-expansive property:

$$\|P_D^\alpha(x) - P_D^\alpha(y)\| \leq \|x - y\|. \quad (6)$$

We then give the following inequalities which will be used in later.

Cauchy–Schwarz inequality Let the stochastic variables ξ_1, ξ_2 is evaluated at interval (a, b) , $-\infty \leq a < b \leq +\infty$ and $\mathbb{E}[|\xi_1|^2] < +\infty$, $\mathbb{E}[|\xi_2|^2] < +\infty$, then we have

$$|\mathbb{E}[\xi_1 \xi_2]|^2 \leq \mathbb{E}[|\xi_1|^2] \mathbb{E}[|\xi_2|^2],$$

especially, if $\xi_2 = 1$, we have

$$|\mathbb{E}[\xi_1]|^2 \leq \mathbb{E}[|\xi_1|^2].$$

We now provide the proof of the statement in the Introduction that “Solving the weakly efficient solution of the problem (4) is equivalent to solve the MVVIP (2)”.

Theorem 1. Let $\mathbb{F}_j(x^*) = \nabla_x p_j(x^*)$ and $g_j(x) = \frac{1}{\rho_n} \sum_{j=1}^t \vartheta[\rho_n q_j(x)]$. For fixed penalty parameter $\rho_n > 0$, the necessary and sufficient condition for x^* to be the solution of the MVVIP is that x^* is the weakly efficient solution of the problem (4).

Proof. (i) We first prove the necessity. Suppose x^* is not a weakly efficient solution of the problem (4). Due to $g_j(x) = \frac{1}{\rho_n} \sum_{j=1}^t \vartheta[\rho_n q_j(x)]$, for any $j = 1, \dots, m$, there exists $y \in D$ such that $p_j(y) + g_j(y) - (p_j(x^*) + g_j(x^*)) < 0$. Thus, we have

$$(p_1(y) - p_1(x^*), \dots, p_m(y) - p_m(x^*)) + (g_1(y) - g_1(x^*), \dots, g_m(y) - g_m(x^*)) \in -\text{int}\mathbb{R}_+^m.$$

Since p_j is a continuous differentiable convex function, for every $j = 1, \dots, m$, it holds that $p_j(y) - p_j(x^*) \geq (y - x^*)^T \nabla_x p_j(x^*)$. Therefore

$$((y - x^*)^T \nabla_x p_1(x^*), \dots, (y - x^*)^T \nabla_x p_m(x^*)) + (g_1(y), \dots, g_m(y)) - (g_1(x^*), \dots, g_m(x^*)) \in -\text{int}\mathbb{R}_+^m. \quad (7)$$

Since $\mathbb{F}_j(x^*) = \nabla_x p_j(x^*)$, $j = 1, \dots, m$, problem (7) is equivalent to

$$((y - x^*)^T \mathbb{F}_1(x^*), \dots, (y - x^*)^T \mathbb{F}_m(x^*)) + (g_1(y), \dots, g_m(y)) - (g_1(x^*), \dots, g_m(x^*)) \in -\text{int}\mathbb{R}_+^m.$$

This is in contradiction with the solution that x^* is the solution of MVVIP. Therefore, the necessity is established.

(ii) We then prove sufficiency. If x^* is not a solution of the MVVIP, there exists $y \in D$ such that

$$((y - x^*)^T \mathbb{F}_1(x^*), \dots, (y - x^*)^T \mathbb{F}_m(x^*)) + (g_1(y), \dots, g_m(y)) - (g_1(x^*), \dots, g_m(x^*)) \in -\text{int}\mathbb{R}_+^m.$$

That is, for any $j = 1, \dots, m$, there holds $(y - x^*)^T \mathbb{F}_j(x^*) + g_j(y) - g_j(x^*) < 0$. Because g_j is a convex function, for every $j = 1, \dots, m$, we have $g_j(y) - g_j(x^*) \geq (y - x^*)^T \nabla_x g_j(x^*)$. Thus

$$(y - x^*)^T \mathbb{F}_j(x^*) + (y - x^*)^T \nabla_x g_j(x^*) < 0, \quad \forall j = 1, \dots, m. \quad (8)$$

On the other hand, let $x_\tau = x^* + \tau(y - x^*) \in D$, for any $j = 1, \dots, m$ the following holds

$$p_j(x_\tau) - p_j(x^*) = \tau(y - x^*)^T \nabla_x p_j(x^*) + o(t) \quad (9)$$

and

$$g_j(x_\tau) - g_j(x^*) = \tau(y - x^*)^T \nabla_x g_j(x^*) + o(t), \quad (10)$$

where τ is the sufficiently small number in $(0, 1)$. Therefore, combine with (8)–(10) and $\mathbb{F}_j(x^*) = \nabla_x p_j(x^*)$ we can obtain that there exists $\omega > 0$ such that for every $j = 1, \dots, m$ and $\tau \in (0, \omega)$, we have

$$p_j(x_\tau) - p_j(x^*) + g_j(x_\tau) - g_j(x^*) < 0. \quad (11)$$

This contradicts the definition of efficient solution x^* for problem (4). Therefore, the conclusion holds. $\square$

For giving the deterministic model of SMVVIP, we first propose an equivalent formulation for MVVIP by the following Theorem, which is called the equivalent mixed scalar variational inequality problems.

Theorem 2. Solving the problem (1) is equivalent to finding $(x^*, \lambda^*) \in D \times \Lambda$ satisfying

$$(y - x^*)^T \sum_{j=1}^m \lambda_j^* \mathbb{F}_j(x^*) + \sum_{j=1}^m \lambda_j^* g_j(y) - \sum_{j=1}^m \lambda_j^* g_j(x^*) \geq 0, \quad \forall y \in D. \quad (12)$$

Proof. (i) We first prove the necessity. Obviously, problem (2) is equivalent to the following form: finding $x^* \in D$ satisfies

$$\max_{1 \leq j \leq m} (y - x^*)^T \mathbb{F}_j(x^*) + g_j(y) - g_j(x^*) \geq 0, \quad \forall y \in D.$$

Let $M(y) = \max_{1 \leq j \leq m} \{(y - x^*)^T \mathbb{F}_j(x^*) + g_j(y) - g_j(x^*)\}$. Furthermore, noting that $y = x^* \in D$ is an optimal solution of the following problem

$$\begin{aligned} \min \quad & M(y) := \max_{1 \leq j \leq m} \{(y - x^*)^T \mathbb{F}_j(x^*) + g_j(y) - g_j(x^*)\} \\ \text{s.t.} \quad & y \in D. \end{aligned}$$

In addition, we observe that the first-order necessary and sufficient condition of the above problem is $0 \in \partial M(x^*) + N_D(x^*)$. By [25], it holds that

$$0 \in \text{conv}\{\mathbb{F}_j(x^*) + \partial g_j(x^*)\} + N_D(x^*), \quad j = 1, \dots, m.$$

Thus, there exists $\lambda^* \geq 0$ with $\sum_{j=1}^m \lambda_j^* = 1$, satisfying $0 \in \sum_{j=1}^m \lambda_j^* \mathbb{F}_j(x^*) + \sum_{j=1}^m \lambda_j^* \partial g_j(x^*) + N_D(x^*)$. By the definition of normal cone, we can easily obtain that

$$(y - x^*)^T \left(\sum_{j=1}^m \lambda_j^* \mathbb{F}_j(x^*) \right) + (y - x^*)^T \left(\sum_{j=1}^m \lambda_j^* \partial g_j(x^*) \right) \geq 0. \quad (13)$$

Combining with the subdifferential operator of $g_j(x)$, $j = 1, \dots, m$, we have

$$\sum_{j=1}^m \lambda_j^* g_j(y) - \sum_{j=1}^m \lambda_j^* g_j(x^*) > (y - x^*)^T \left(\sum_{j=1}^m \lambda_j^* \partial g_j(x^*) \right). \quad (14)$$

Adding (13) and (14), we obtain

$$(y - x^*)^T \sum_{j=1}^m \lambda_j^* \mathbb{F}_j(x^*) + \sum_{j=1}^m \lambda_j^* g_j(y) - \sum_{j=1}^m \lambda_j^* g_j(x^*) \geq 0, \quad \forall y \in D.$$

Therefore, the necessity is established.

(ii) Next, we prove the sufficiency. Suppose that there exists $(x^*, \lambda^*) \in D \times \Lambda$, satisfying

$$(y - x^*)^T \sum_{j=1}^m \lambda_j^* \mathbb{F}_j(x^*) + \sum_{j=1}^m \lambda_j^* g_j(y) - \sum_{j=1}^m \lambda_j^* g_j(x^*) \geq 0, \quad \forall y \in D.$$

Since $\lambda^* \in \Lambda$, there exists at least one $\lambda_j^* > 0$, $j = 1, \dots, m$, such that

$$(y - x^*)^T \lambda_j^* \mathbb{F}_j(x^*) + \lambda_j^* g_j(y) - \lambda_j^* g_j(x^*) \geq 0, \quad \forall y \in D. \quad (15)$$

Thus, (15) is equivalent to

$$(y - x^*)^T \mathbb{F}_j(x^*) + g_j(y) - g_j(x^*) \geq 0.$$

Hence, we have

$$\left((y-x^*)^T \mathbb{F}_1(x^*), \dots, (y-x^*)^T \mathbb{F}_m(x^*) \right) + (g_1(y), \dots, g_m(y)) - (g_1(x^*), \dots, g_m(x^*)) \notin -\text{int} \mathbb{R}_+^m,$$

which means that x^* is a solution of MVVIP. $\square$

Now, in order to present the deterministic bi-criteria model for solving SMVVIP, we require introducing the following two deterministic models for solving SVIP mentioned in section one:

- The ERM model for solving SVIP:

$$\begin{aligned} & \min \mathbb{E}[v(x, \xi)] \\ & \text{s.t. } x \in D, \end{aligned}$$

where $\mathbb{E}$ stands for mathematical expectation and $v(x, \xi)$ is the regularized gap function for SVIP.

- The CVaR model for solving SVIP:

To present the CVaR model, we first introduce the value at risk (VaR) model. For any fixed x and confidence level $\beta \in (0, 1)$, the regularized gap function $v(x, \xi)$ is taken as the loss function, and the VaR for the loss associated with x is defined as

$$\text{VaR}_\beta^v(x) = \min\{u \mid \mathcal{P}[v(x, \xi) \leq u] \geq \beta\},$$

where $\mathcal{P}$ stands for the probability. However, as a decision variable function, VaR usually does not satisfy the convexity and consistency of risk measurement selection. Due to this, Rockafellar et al. [28] proposed the CVaR, which had better properties than VaR and is defined as

$$\text{CVaR}_\beta^v(x) = \frac{1}{1-\beta} \int_{v(x, \xi) \geq \text{VaR}_\beta^v(x)} v(x, \xi) d\rho(\xi),$$

where $\rho(\xi)$ represents the distribution function of ξ .

Chen et al. [19] gave the CVaR model for solving SVIP as follows:

$$\min_{x \in \mathbb{R}^n} \text{CVaR}_\beta^v(x). \quad (16)$$

By Theorem 2 of [28], problem (16) is equivalent to

$$\min_{(x, u) \in \mathbb{R}^{n+1}} \mathcal{V}(x, u) = u + \frac{1}{1-\beta} \mathbb{E}[v(x, \xi) - u]_+. \quad (17)$$

Because the above two models contain mathematical expectations that are usually difficult to calculate in practice, scholars usually apply the sample average approximation (SAA) method to dispose the expected value. That is, for an integrable function $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$, if the independent identically distributed samples of stochastic variable ξ is $\Xi_k := \{\xi_1, \dots, \xi_{N_k}\} \subseteq \Xi$, the sample average value $\lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{\xi_i \in \Xi_k} \psi(\xi_i)$ can be used to approximate the expected value $\mathbb{E}[\psi(\xi)]$, where $N_k \rightarrow +\infty$, as $k \rightarrow +\infty$. The above limit process holds with probability one (we abbreviated it by “w.p.1” below) by the strong law of large numbers, that is

$$\lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{\xi_i \in \Xi_k} \psi(\xi_i) = \mathbb{E}[\psi(\xi)], \quad \text{w.p.1.} \quad (18)$$

3. Deterministic Bi-Criteria Model for Solving SMVVIP

In this section, based on the works for solving SVIP, we will present a deterministic bi-criteria model for solving SMVVIP. By Theorem 2, similar to the equivalent form of MVVIP, we can write the SMVVIP (5) as the following equivalent stochastic mixed scalar variational inequality problems: finding $(x^*, \lambda^*) \in D \times \Lambda$ satisfies

$$(y - x^*)^T \sum_{j=1}^m \lambda_j^* \mathbb{F}_j(x^*, \xi) + \sum_{j=1}^m \lambda_j^* g_j(y, \xi) - \sum_{j=1}^m \lambda_j^* g_j(x^*, \xi) \geq 0, \quad \forall y \in D, \text{ a.s. } \xi \in \Xi. \quad (19)$$

The existence of stochastic factors lead to the fact that problem (19) does not have a common solution for almost every $\xi \in \Xi$. Therefore, in order to obtain a reasonable solution, we need to give a reasonable deterministic model for (19), and treat the solution of this deterministic model as the solution of SMVVIP. In order to give the deterministic model, motivated by the work of Fukushima [16], we define the regularized gap function $h(x, \lambda, \xi) : \mathbb{R}^n \times \Lambda \times \Xi \rightarrow \mathbb{R}$ of (19) as follows:

$$h(x, \lambda, \xi) = \max_{y \in D} \left\{ (x - y)^T \sum_{j=1}^m \lambda_j \mathbb{F}_j(x, \xi) + \sum_{j=1}^m \lambda_j g_j(x, \xi) - \sum_{j=1}^m \lambda_j g_j(y, \xi) - \frac{\alpha}{2} \|x - y\|^2 \right\}. \quad (20)$$

As stated by [29], for any $(x, \lambda) \in D \times \Lambda$ and $\xi \in \Xi$, we have

$$\begin{aligned} h(x, \lambda, \xi) &= (x - H(x, \lambda, \xi))^T \sum_{j=1}^m \lambda_j \mathbb{F}_j(x, \xi) + \sum_{j=1}^m \lambda_j g_j(x, \xi) \\ &\quad - \sum_{j=1}^m \lambda_j g_j(H(x, \lambda, \xi), \xi) - \frac{\alpha}{2} \|x - H(x, \lambda, \xi)\|^2, \end{aligned} \quad (21)$$

which can serve as a gap function of (19), while

$$H(x, \lambda, \xi) = P_D^\alpha \left(x - \alpha^{-1} \sum_{j=1}^m \lambda_j \mathbb{F}_j(x, \xi) \right).$$

Obviously, based on Fukushima's work in [30], for each $(x, \lambda) \in D \times \Lambda$, the regularized gap function $h(x, \lambda, \xi) \geq 0$, a.s. $\xi \in \Xi$. In addition, by Theorem 2.2 in [29], we have $h(x^*, \lambda^*, \xi) = 0$ for each $(x^*, \lambda^*) \in D \times \Lambda$ if and only if (x^*, λ^*) solves (19). Hence, we can easily obtain that the following optimization problem is equivalent to (19):

$$\begin{aligned} \min \quad & h(x, \lambda, \xi) \\ \text{s.t.} \quad & (x, \lambda) \in D \times \Lambda, \text{ a.s. } \xi \in \Xi. \end{aligned}$$

According to reference [30], for a given function $\gamma : \mathbb{R}^n \rightarrow \mathbb{R}^n$, when the optimal solution of $\max_{y \in D} \{\gamma(x)^T (x - y)\}$ is unique, let

$$Y(x) = \max_{y \in D} \{\gamma(x)^T (x - y)\},$$

where the gradient of Y at x is actually showed by

$$\nabla_x Y(x) = \gamma(x) - \nabla_x \gamma(x)^T (y - x). \quad (22)$$

In addition, by Theorem 3.2 in [16] and for each $j = 1, \dots, m$, $\mathbb{F}_j(x, \xi)$ and $g_j(\cdot)$ are continuously differentiable, we have $h(x, \lambda, \xi)$ is continuously differentiable at (x, λ) . Since

the optimal solution of formulation (20) is unique, based on the above results, we can obtain the gradient of $h(x, \lambda, \xi)$ at (x, λ) as follows:

$$\nabla_{(x,\lambda)} h(x, \lambda, \xi) = \begin{pmatrix} \sum_{j=1}^m \lambda_j \mathbb{F}_j(x, \xi) + \sum_{j=1}^m \lambda_j \nabla g_j(x, \xi) - \left(\sum_{j=1}^m \lambda_j \nabla_x \mathbb{F}_j(x, \xi) - \alpha I \right) (H(x, \lambda, \xi) - x) \\ (x - H(x, \lambda, \xi))^T \mathbb{F}_1(x, \xi) - g_1(H(x, \lambda, \xi), \xi) + g_1(x, \xi) \\ \vdots \\ (x - H(x, \lambda, \xi))^T \mathbb{F}_m(x, \xi) - g_m(H(x, \lambda, \xi), \xi) + g_m(x, \xi) \end{pmatrix},$$

where I denotes the unit matrix in appropriate dimension.

Next, in combination with the research of ERM model [17,22] and CVaR model [19], along with the bi-criteria model [31], for solving SVIP, we give the deterministic bi-criteria model for problem (19) as follows:

$$\min_{(x,\lambda) \in D \times \Lambda} (\mathbb{E}_{\Xi}[h(x, \lambda, \xi)], \text{CVaR}_{\beta}^h(x, \lambda)), \quad (23)$$

where $\mathbb{E}_{\Xi}[h(x, \lambda, \xi)] = \int_{\Xi} h(x, \lambda, \xi) d\xi$ and

$$\text{CVaR}_{\beta}^h(x, \lambda) = \frac{1}{1 - \beta} \int_{h(x, \lambda, \xi) \geq \text{VaR}_{\beta}^h(x, \lambda)} h(x, \lambda, \xi) d\rho(\xi).$$

We adopt the weighted sum method in [32] to deal with the multiobjective optimization problems (23), that is, given a weight coefficient $r \in [0, 1]$, which can be modeled different degrees of risk aversion to decision-makers in the real world. Using the method mentioned above, we obtain the following weighted optimization problem:

$$\min_{(x,\lambda) \in D \times \Lambda} \theta(x, \lambda) := (1 - r) \mathbb{E}_{\Xi}[h(x, \lambda, \xi)] + r \text{CVaR}_{\beta}^h(x, \lambda). \quad (24)$$

Note that if $r = 0$, (24) becomes the ERM model for solving SMVVIP. Because this model does not take into account the error between the real possible of the regularized gap function and its expected residual, the ERM model is risk neutral. If $r = 1$, (24) converts to the CVaR model for solving SMVVIP, which makes it risk averse. For some investors, though, it may be too conservative. Therefore, the above model can make a compromise between the expected residual of regularized gap function and the robust value model in the worst case. More details about the deterministic bi-criteria model can be found in reference [31]. According to Theorem 2 of [28], for a given parameter $r \in [0, 1]$, we can easily obtain that problem (24) is equivalent to

$$\min_{(x,\lambda,u) \in S} \theta(x, \lambda, u) := (1 - r) \mathbb{E}_{\Xi}[h(x, \lambda, \xi)] + ru + \frac{r}{1 - \beta} \mathbb{E}_{\Xi}[[h(x, \lambda, \xi) - u]_+], \quad (25)$$

where $S := D \times \Lambda \times \mathbb{R}$ and $\mathbb{E}_{\Xi}[[h(x, \lambda, \xi) - u]_+] = \int_{\Xi} [h(x, \lambda, \xi) - u]_+ d\xi$.

Note that due to the existence of non-smooth function $[h(x, \lambda, \xi) - u]_+$, problem (25) is a non-smooth optimization problem. In order to deal with the non-smoothness of the function, as the special case of the smoothing function proposed by Peng in [33], given a parameter $\epsilon > 0$, we can obtain the smoothing form of $[t]_+ = \max\{t, 0\}$ as follows:

$$\phi(t, \epsilon) = \epsilon \ln(e^{\frac{t}{\epsilon}} + 1).$$

In addition, $\phi(t, \epsilon)$ also has the following properties [34]:

$$\lim_{\epsilon \rightarrow 0_+} \epsilon \ln(e^{\frac{t}{\epsilon}} + 1) = [t]_+, \quad 0 \leq \epsilon \ln(e^{\frac{t}{\epsilon}} + 1) - [t]_+ \leq \epsilon \ln 2. \quad (26)$$

Thus, we give the smoothing approximation problem of problem (25) as follows:

$$\min_{(x,\lambda,u) \in S} \theta(x, \lambda, u, \epsilon) := (1-r)\mathbb{E}_{\Xi}[h(x, \lambda, \xi)] + ru + \frac{r}{1-\beta}\mathbb{E}_{\Xi}[l(x, \lambda, u, \xi, \epsilon)], \quad (27)$$

where $l(x, \lambda, u, \xi, \epsilon) = \epsilon \ln(e^{\frac{h(x, \lambda, \xi) - u}{\epsilon}} + 1)$ and let $l(x, \lambda, u, \xi) = [h(x, \lambda, \xi) - u]_+$.

Note that problem (27) contains a mathematical expectation that is difficult to compute, and we apply the SAA method introduced in Section 2 to obtain the following smoothing SAA problem of (27):

$$\min_{(x,\lambda,u) \in S} \theta^k(x, \lambda, u, \epsilon) := \frac{(1-r)}{N_k} \sum_{\xi_i \in \Xi_k} h(x, \lambda, \xi_i) + ru + \frac{r}{(1-\beta)N_k} \sum_{\xi_i \in \Xi_k} l(x, \lambda, u, \xi_i, \epsilon). \quad (28)$$

We will investigate the limiting behavior of these approximation problems in the following section.

4. Convergence Analysis

In this section, we first investigate the limiting behavior of global optimal solutions of problem (28). Then, when $\Xi \subseteq \mathbb{R}^b$ is not a compact set, we will propose the compact approximation problem and give the proof of the convergence of compact approximation problem.

4.1. Convergence Behavior of Global Optimal Solutions

We now assume that ϵ is related to k , so we let ϵ in problem (28) is ϵ^k , and $\epsilon^k \rightarrow 0$ when $k \rightarrow \infty$. From now on, $\bar{S}$ and S_k represent the sets of optimal solutions of problems (25) and (28), respectively.

Theorem 3. Suppose that $\Xi \subset \mathbb{R}^b$ is a nonempty compact set. For given parameters r, β and for each k , let $(x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k)) \in S_k$ and $(\bar{x}, \bar{\lambda}, \bar{u})$ is an accumulation point of $(x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k))$. Then, we have $(\bar{x}, \bar{\lambda}, \bar{u}) \in \bar{S}$ w.p.1.

Proof. We assume that $\lim_{k \rightarrow \infty} (x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k)) = (\bar{x}, \bar{\lambda}, \bar{u})$. Let D_1, D_2 , and D_3 be a compact convex set containing $\{(x^k(\epsilon^k), \lambda^k(\epsilon^k))\}$, $\{u^k(\epsilon^k)\}$, and $\{\epsilon^k\}$, respectively. By the continuous differentiability of $h(x, \lambda, \xi)$ on $D_1 \times \Xi$, there exists a constant $C_1 > 0$ satisfying

$$\|\nabla_{(x,\lambda)} h(x, \lambda, \xi)\| \leq C_1. \quad (29)$$

In addition, since $l(x, \lambda, u, \xi, \epsilon)$ is continuously differentiable on $D_1 \times D_2 \times \Xi \times D_3$, we have that for any $(x, \lambda, u, \xi, \epsilon) \in D_1 \times D_2 \times \Xi \times D_3$, there exists a constant $C_2 > 0$ satisfying

$$\|\nabla_{(x,\lambda,u)} l(x, \lambda, u, \xi, \epsilon)\| \leq C_2. \quad (30)$$

It follows from the mean value theorem that, for each $x^k(\epsilon^k), \lambda^k(\epsilon^k)$ and each ξ_i , there exists $(x^{k_i}, \lambda^{k_i}) = \alpha^{k_i}(x^k(\epsilon^k), \lambda^k(\epsilon^k)) + (1 - \alpha^{k_i})(\bar{x}, \bar{\lambda})$ with $\alpha^{k_i} \in (0, 1)$ satisfying

$$h(x^k(\epsilon^k), \lambda^k(\epsilon^k), \xi_i) - h(\bar{x}, \bar{\lambda}, \xi_i) = \left[\nabla_{(x,\lambda)} h(x^{k_i}, \lambda^{k_i}, \xi_i) \right]^T \left((x^k(\epsilon^k), \lambda^k(\epsilon^k)) - (\bar{x}, \bar{\lambda}) \right). \quad (31)$$

Similarly, for each k and ϵ^k , there exists $(x^{k'_i}, \lambda^{k'_i}, u^{k'_i}) = \alpha^{k'_i}(x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k)) + (1 - \alpha^{k'_i})(\bar{x}, \bar{\lambda}, \bar{u})$ with $\alpha^{k'_i} \in (0, 1)$ satisfying

$$\begin{aligned} & l(x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k), \xi_i, \epsilon^k) - l(\bar{x}, \bar{\lambda}, \bar{u}, \xi_i, \epsilon^k) \\ &= \left[\nabla_{(x,\lambda,u)} l(x^{k'_i}, \lambda^{k'_i}, u^{k'_i}, \xi_i, \epsilon^k) \right]^T \left((x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k)) - (\bar{x}, \bar{\lambda}, \bar{u}) \right). \end{aligned} \quad (32)$$

Thus, according to formulation (29)–(32) we have

$$\begin{aligned}
 & |\theta^k(x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k), \epsilon^k) - \theta^k(\bar{x}, \bar{\lambda}, \bar{u}, \epsilon^k)| \\
 & \leq \frac{(1-r)}{N_k} \sum_{\xi_i \in \Xi_k} |h(x^k(\epsilon^k), \lambda^k(\epsilon^k), \xi_i) - h(\bar{x}, \bar{\lambda}, \xi_i)| + r|u^k(\epsilon^k) - \bar{u}| \\
 & \quad + \frac{r}{(1-\beta)N_k} \sum_{\xi_i \in \Xi_k} |l(x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k), \xi_i, \epsilon^k) - l(\bar{x}, \bar{\lambda}, \bar{u}, \xi_i, \epsilon^k)| \\
 & \leq r|u^k(\epsilon^k) - \bar{u}| + \frac{C_1(1-r)}{N_k} \sum_{\xi_i \in \Xi_k} (\|x^k(\epsilon^k) - \bar{x}\| + \|\lambda^k(\epsilon^k) - \bar{\lambda}\|) \\
 & \quad + \frac{rC_2}{(1-\beta)N_k} \sum_{\xi_i \in \Xi_k} (\|x^k(\epsilon^k) - \bar{x}\| + \|\lambda^k(\epsilon^k) - \bar{\lambda}\| + |u^k(\epsilon^k) - \bar{u}|) \\
 & \rightarrow 0 \text{ as } k \rightarrow \infty, \quad \text{w.p.1.}
 \end{aligned} \tag{33}$$

On the other hand, according to (18), we know that

$$\lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{\xi_i \in \Xi_k} h(\bar{x}, \bar{\lambda}, \xi_i) = \mathbb{E}_{\Xi}[h(\bar{x}, \bar{\lambda}, \xi)] \tag{34}$$

and

$$\lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{\xi_i \in \Xi_k} l(\bar{x}, \bar{\lambda}, \bar{u}, \xi_i) = \mathbb{E}_{\Xi}[l(\bar{x}, \bar{\lambda}, \bar{u}, \xi)]. \tag{35}$$

We then have from $0 \leq l(x, \lambda, u, \xi, \epsilon^k) - l(x, \lambda, u, \xi) \leq \epsilon^k \ln 2$, $\epsilon^k \rightarrow 0$ as $k \rightarrow \infty$ and (34), (35) that

$$\begin{aligned}
 & |\theta^k(\bar{x}, \bar{\lambda}, \bar{u}, \epsilon^k) - \theta(\bar{x}, \bar{\lambda}, \bar{u})| \\
 & = \left| \frac{(1-r)}{N_k} \sum_{\xi_i \in \Xi_k} h(\bar{x}, \bar{\lambda}, \xi_i) + r\bar{u} + \frac{r}{(1-\beta)N_k} \sum_{\xi_i \in \Xi_k} l(\bar{x}, \bar{\lambda}, \bar{u}, \xi_i, \epsilon^k) \right. \\
 & \quad \left. - (1-r)\mathbb{E}_{\Xi}[h(\bar{x}, \bar{\lambda}, \xi)] - r\bar{u} - \frac{r}{1-\beta}\mathbb{E}_{\Xi}[l(\bar{x}, \bar{\lambda}, \bar{u}, \xi)] \right| \\
 & \leq (1-r) \left| \frac{1}{N_k} \sum_{\xi_i \in \Xi_k} h(\bar{x}, \bar{\lambda}, \xi_i) - \mathbb{E}_{\Xi}[h(\bar{x}, \bar{\lambda}, \xi)] \right| \\
 & \quad + \frac{r}{1-\beta} \left| \frac{1}{N_k} \sum_{\xi_i \in \Xi_k} l(\bar{x}, \bar{\lambda}, \bar{u}, \xi_i) - \mathbb{E}_{\Xi}[l(\bar{x}, \bar{\lambda}, \bar{u}, \xi)] \right| \\
 & \quad + \frac{r}{(1-\beta)N_k} \sum_{\xi_i \in \Xi_k} \left| l(\bar{x}, \bar{\lambda}, \bar{u}, \xi_i, \epsilon^k) - l(\bar{x}, \bar{\lambda}, \bar{u}, \xi_i) \right| \\
 & \leq (1-r) \left| \frac{1}{N_k} \sum_{\xi_i \in \Xi_k} h(\bar{x}, \bar{\lambda}, \xi_i) - \mathbb{E}_{\Xi}[h(\bar{x}, \bar{\lambda}, \xi)] \right| + \frac{r}{(1-\beta)N_k} \sum_{\xi_i \in \Xi_k} \epsilon^k \ln 2 \\
 & \quad + \frac{r}{1-\beta} \left| \frac{1}{N_k} \sum_{\xi_i \in \Xi_k} l(\bar{x}, \bar{\lambda}, \bar{u}, \xi_i) - \mathbb{E}_{\Xi}[l(\bar{x}, \bar{\lambda}, \bar{u}, \xi)] \right| \\
 & \rightarrow 0 \text{ as } k \rightarrow \infty, \quad \text{w.p.1.}
 \end{aligned} \tag{36}$$

Similar to (36), we have

$$\lim_{k \rightarrow \infty} \theta^k(x, \lambda, u, \epsilon^k) = \theta(x, \lambda, u). \tag{37}$$

Since

$$\begin{aligned} & |\theta^k(x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k), \epsilon^k) - \theta(\bar{x}, \bar{\lambda}, \bar{u})| \\ & \leq |\theta^k(x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k), \epsilon^k) - \theta^k(\bar{x}, \bar{\lambda}, \bar{u}, \epsilon^k)| + |\theta^k(\bar{x}, \bar{\lambda}, \bar{u}, \epsilon^k) - \theta(\bar{x}, \bar{\lambda}, \bar{u})|. \end{aligned}$$

It follows from (33) and (36) that

$$\lim_{k \rightarrow \infty} \theta^k(x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k), \epsilon^k) = \theta(\bar{x}, \bar{\lambda}, \bar{u}), \quad \text{w.p.1.} \quad (38)$$

For all k , since $(x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k)) \in \mathcal{S}_k$, the following holds

$$\theta^k(x^k(\epsilon^k), \lambda^k(\epsilon^k), u^k(\epsilon^k), \epsilon^k) \leq \theta^k(x, \lambda, u, \epsilon^k), \quad \forall (x, \lambda, u) \in S.$$

Let $k \rightarrow \infty$, by (37) and (38), for any $(x, \lambda, u) \in S$, we have

$$\theta(\bar{x}, \bar{\lambda}, \bar{u}) \leq \theta(x, \lambda, u), \quad \text{w.p.1.}$$

which means that $(\bar{x}, \bar{\lambda}, \bar{u}) \in \bar{\mathcal{S}}$ w.p.1. $\square$

Remark 1. In Theorem 3, we only show the convergence result of the global optimal solution sequence of smoothing SAA problem (28). In fact, when the function $\mathbb{E}[\epsilon \ln(e^{\frac{h(x, \lambda, \xi) - u}{\epsilon}} + 1)]$ can be evaluated, as a part of the proof of Theorem 3, we can easily obtain that the global optimal solution sequence of smoothing approximation problem (27) converges to the global optimal solution of problem (25).

Remark 2. Note that (28) is a non-convex problem, so it is difficult to obtain a global optimal solution sequence of problem (28) in general. Therefore, it is essential to consider the convergence of stationary points. According to Theorem 2 of reference [35], we can similarly prove that the stationary points of smoothing SAA problem (28) converges to the stationary points of problem (25). Here, we omit the proof process.

4.2. Convergence of the Compact Approximation Problem

It is worth noting that in Theorem 3, the sample space $\Xi \subseteq \mathbb{R}^b$ is a compact set. If $\Xi \subseteq \mathbb{R}^b$ is not a compact set, we apply the compact approximation method to solve the problem. In order to give the compact approximation problem of (25), for a sufficiently large number ν , we first define the compact approximation set of Ξ as follows:

$$\Xi_\nu := \{\xi \in \Xi \mid \|\xi\| \leq \nu\}. \quad (39)$$

Then, we present the compact approximation problem of (25) as follows:

$$\min_{(x, \lambda, u) \in S} \theta_\nu(x, \lambda, u) := (1 - r)\mathbb{E}_{\Xi_\nu}[h(x, \lambda, \xi)] + ru + \frac{r}{1 - \beta}\mathbb{E}_{\Xi_\nu}[[h(x, \lambda, \xi) - u]_+], \quad (40)$$

where $S := D \times \Lambda \times \mathbb{R}$ and D be a nonempty closed convex set, $\mathbb{E}_{\Xi_\nu}[h(x, \lambda, \xi)] = \int_{\Xi_\nu} h(x, \lambda, \xi) d\xi$ and $\mathbb{E}_{\Xi_\nu}[[h(x, \lambda, \xi) - u]_+] = \int_{\Xi_\nu} [h(x, \lambda, \xi) - u]_+ d\xi$. Ξ_ν is a compact set. Next, we make some assumptions.

(A1) For each $j = 1, \dots, m$, there exists an integrable function $\kappa_j(\xi)$ such that $\mathbb{E}_\Xi[\kappa_j^2(\xi)] < +\infty$ and

$$\|\mathbb{F}_j(y, \xi) - \mathbb{F}_j(x, \xi)\| \leq \kappa_j(\xi)\|y - x\|, \quad \forall x, y \in D.$$

Furthermore, for every $j = 1, \dots, m$, there exists $\check{x} \in D$ such that $\mathbb{E}_\Xi[\|\mathbb{F}_j(\check{x}, \xi)\|^2] < +\infty$.

(A2) There exists an integrable function $\delta(\xi)$ such that $\mathbb{E}_{\Xi}[\delta^2(\xi)] < +\infty$ and

$$\|g(y, \xi) - g(x, \xi)\| \leq \delta(\xi)\|y - x\|, \quad \forall x, y \in D.$$

In addition, there exists $\acute{x} \in D$ such that $\mathbb{E}_{\Xi}[\|g(\acute{x}, \xi)\|] < +\infty$.

Remark 3. For every $x \in D$, by assumption (A1), for every $j = 1, \dots, m$, we can easily obtain that

$$\begin{aligned} \mathbb{E}_{\Xi}[\|\mathbb{F}_j(x, \xi)\|^2] &\leq \mathbb{E}_{\Xi}[(\|\mathbb{F}_j(x, \xi) - \mathbb{F}_j(\acute{x}, \xi)\| + \|\mathbb{F}_j(\acute{x}, \xi)\|)^2] \\ &\leq 2\|x - \acute{x}\|^2 \mathbb{E}_{\Xi}[\kappa_j^2(\xi)] + 2\mathbb{E}_{\Xi}[\|\mathbb{F}_j(\acute{x}, \xi)\|^2] \\ &< +\infty. \end{aligned} \quad (41)$$

Furthermore, for every $x \in D$, we can easily obtain that

$$\mathbb{E}_{\Xi}[(\sum_{j=1}^m \kappa_j(\xi))^2] < +\infty, \quad \mathbb{E}_{\Xi}[(\sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\|)^2] < +\infty. \quad (42)$$

In addition, under conditions (A1)–(A2), by Cauchy–Schwarz inequalities and (41) and (42), for every $x \in D$ and $j = 1, \dots, m$, we can obtain

$$\mathbb{E}_{\Xi}[\|\mathbb{F}_j(x, \xi)\|] < +\infty, \quad \mathbb{E}_{\Xi}[\kappa_j(\xi)] < +\infty, \quad \mathbb{E}_{\Xi}[\delta(\xi)] < +\infty \quad (43)$$

and

$$\mathbb{E}_{\Xi}[\sum_{j=1}^m \kappa_j(\xi) \sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\|] \leq \sqrt{\mathbb{E}_{\Xi}[(\sum_{j=1}^m \kappa_j(\xi))^2] \mathbb{E}_{\Xi}[(\sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\|)^2]} < +\infty. \quad (44)$$

Similar to (44), we also have

$$\mathbb{E}_{\Xi}[\delta(\xi) \sum_{j=1}^m \kappa_j(\xi)] < +\infty, \quad \mathbb{E}_{\Xi}[\delta(\xi) \sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\|] < +\infty. \quad (45)$$

Moreover, by (A2) and (43), we have

$$\begin{aligned} \mathbb{E}_{\Xi}[\|g(x, \xi)\|] &\leq \mathbb{E}_{\Xi}[\|g(x, \xi) - g(\acute{x}, \xi)\|] + \mathbb{E}_{\Xi}[\|g(\acute{x}, \xi)\|] \\ &\leq \mathbb{E}_{\Xi}[\delta(\xi)]\|x - \acute{x}\| + \mathbb{E}_{\Xi}[\|g(\acute{x}, \xi)\|] \\ &< +\infty. \end{aligned} \quad (46)$$

Thus, for every $x \in D$ and $j = 1, \dots, m$, we have

$$\mathbb{E}_{\Xi}[\|g_j(x, \xi)\|] \leq \mathbb{E}_{\Xi}[\|g(x, \xi)\|] < +\infty. \quad (47)$$

Lemma 1. Let (A1)–(A2) hold, then we have

$$\mathbb{E}_{\Xi}[h(x, \lambda, \xi)] < +\infty, \quad \forall (x, \lambda) \in D \times \Lambda.$$

Proof. By (21), the fact that $h(x, \lambda, \xi) \geq 0$, $\sum_{j=1}^m \lambda_j \geq 0$, $\sum_{j=1}^m \lambda_j = 1$ and (A2), there holds

$$\begin{aligned} & \frac{\alpha}{2} \|x - H(x, \lambda, \xi)\|^2 \\ & \leq (x - H(x, \lambda, \xi))^T \sum_{j=1}^m \lambda_j \mathbb{F}_j(x, \xi) + \sum_{j=1}^m \lambda_j g_j(x, \xi) - \sum_{j=1}^m \lambda_j g_j(H(x, \lambda, \xi), \xi) \\ & \leq \|x - H(x, \lambda, \xi)\| \sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\| + m \|g(x, \xi) - g(H(x, \lambda, \xi), \xi)\| \\ & \leq \|x - H(x, \lambda, \xi)\| \left(\sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\| + m\delta(\xi) \right). \end{aligned} \quad (48)$$

Therefore, we have

$$\|x - H(x, \lambda, \xi)\| \leq \frac{2}{\alpha} \left(\sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\| + m\delta(\xi) \right). \quad (49)$$

In addition, we can obtain

$$\begin{aligned} & \mathbb{E}_{\Xi} [h(x, \lambda, \xi)] \\ & = \mathbb{E}_{\Xi} \left[(x - H(x, \lambda, \xi))^T \sum_{j=1}^m \lambda_j \mathbb{F}_j(x, \xi) + \sum_{j=1}^m \lambda_j g_j(x, \xi) - \sum_{j=1}^m \lambda_j g_j(H(x, \lambda, \xi), \xi) \right. \\ & \quad \left. - \frac{\alpha}{2} \|x - H(x, \lambda, \xi)\|^2 \right] \\ & \leq \mathbb{E}_{\Xi} \left[\|x - H(x, \lambda, \xi)\| \sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\| + m \|g(x, \xi) - g(H(x, \lambda, \xi), \xi)\| + \frac{\alpha}{2} \|x - H(x, \lambda, \xi)\|^2 \right] \\ & \leq \mathbb{E}_{\Xi} \left[\|x - H(x, \lambda, \xi)\| \sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\| + m\delta(\xi) \|x - H(x, \lambda, \xi)\| + \frac{\alpha}{2} \|x - H(x, \lambda, \xi)\|^2 \right] \\ & \leq \mathbb{E}_{\Xi} \left[\frac{4}{\alpha} \left(\sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\| + m\delta(\xi) \right)^2 \right] \\ & = \frac{4}{\alpha} \mathbb{E}_{\Xi} \left[\left(\sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\| \right)^2 \right] + \frac{8m}{\alpha} \mathbb{E}_{\Xi} [\delta(\xi) \sum_{j=1}^m \|\mathbb{F}_j(x, \xi)\|] + \frac{4m^2}{\alpha} \mathbb{E}_{\Xi} [\delta^2(\xi)] \\ & < +\infty, \end{aligned}$$

where the second inequality is derived from (A2) and the inequality of $\sum_{j=1}^m |x_j| \leq m\|x\|$, and the third inequality is given by (49), along with the last inequality follows from (A2), (43), (45). Therefore, the conclusion holds. $\square$

Lemma 2. Suppose that (A1)–(A2) hold and $\lim_{\nu \rightarrow \infty} (x^\nu, \lambda^\nu) = (\tilde{x}, \tilde{\lambda})$. Then, we have $\lim_{\nu \rightarrow \infty} \mathbb{E}_{\Xi_\nu} [h(x^\nu, \lambda^\nu, \xi)] = \mathbb{E}_{\Xi} [h(\tilde{x}, \tilde{\lambda}, \xi)]$ with probability one.

Proof. (i) We first prove that

$$\lim_{\nu \rightarrow \infty} |\mathbb{E}_{\Xi_\nu} [h(x^\nu, \lambda^\nu, \xi)] - \mathbb{E}_{\Xi_\nu} [h(\tilde{x}, \tilde{\lambda}, \xi)]| = 0, \quad \text{w.p.1.} \quad (50)$$

By the non-expansive property of $P_D^{\lambda, \alpha}$ (6), $\sum_{j=1}^m \lambda_j \geq 0$, $\sum_{j=1}^m \lambda_j = 1$ and (A1), we have

$$\begin{aligned} & \|H(x^\nu, \lambda^\nu, \xi) - H(\tilde{x}, \tilde{\lambda}, \xi)\| \\ & \leq \|x^\nu - \tilde{x}\| + \alpha^{-1} \left\| \sum_{j=1}^m \lambda_j^\nu \mathbb{F}_j(x^\nu, \xi) - \sum_{j=1}^m \tilde{\lambda}_j \mathbb{F}_j(\tilde{x}, \xi) \right\| \\ & \leq \|x^\nu - \tilde{x}\| + \alpha^{-1} \sum_{j=1}^m \lambda_j^\nu \|\mathbb{F}_j(x^\nu, \xi) - \mathbb{F}_j(\tilde{x}, \xi)\| + \alpha^{-1} \sum_{j=1}^m \|(\lambda_j^\nu - \tilde{\lambda}_j) \mathbb{F}_j(\tilde{x}, \xi)\| \\ & \leq \|x^\nu - \tilde{x}\| + \alpha^{-1} \|x^\nu - \tilde{x}\| \sum_{j=1}^m \kappa_j(\xi) + \alpha^{-1} \|\lambda^\nu - \tilde{\lambda}\| \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\|. \end{aligned} \quad (51)$$

In consequence, by $\sum_{j=1}^m \lambda_j \geq 0$, $\sum_{j=1}^m \lambda_j = 1$, (A1), (42)–(44) and (51), the following holds

$$\begin{aligned} & \mathbb{E}_\Xi \left[\left((x^\nu - \tilde{x} + H(\tilde{x}, \tilde{\lambda}, \xi) - H(x^\nu, \lambda^\nu, \xi))^T \sum_{j=1}^m \lambda_j^\nu \mathbb{F}_j(x^\nu, \xi) \right) \right] \\ & \leq \mathbb{E}_\Xi \left[(\|x^\nu - \tilde{x}\| + \|H(x^\nu, \lambda^\nu, \xi) - H(\tilde{x}, \tilde{\lambda}, \xi)\|) \left(\sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| + \sum_{j=1}^m \|\mathbb{F}_j(x^\nu, \xi) - \mathbb{F}_j(\tilde{x}, \xi)\| \right) \right] \\ & \leq \mathbb{E}_\Xi \left[(2\|x^\nu - \tilde{x}\| + \alpha^{-1} \|x^\nu - \tilde{x}\| \sum_{j=1}^m \kappa_j(\xi) + \alpha^{-1} \|\lambda^\nu - \tilde{\lambda}\| \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\|) \right. \\ & \quad \cdot \left(\sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| + \sum_{j=1}^m \kappa_j(\xi) \|x^\nu - \tilde{x}\| \right) \\ & = 2\|x^\nu - \tilde{x}\| \mathbb{E}_\Xi \left[\sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right] + 2\|x^\nu - \tilde{x}\|^2 \mathbb{E}_\Xi \left[\sum_{j=1}^m \kappa_j(\xi) \right] \\ & \quad + \alpha^{-1} \|x^\nu - \tilde{x}\| \mathbb{E}_\Xi \left[\sum_{j=1}^m \kappa_j(\xi) \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right] \\ & \quad + \alpha^{-1} \|x^\nu - \tilde{x}\|^2 \mathbb{E}_\Xi \left[\left(\sum_{j=1}^m \kappa_j(\xi) \right)^2 \right] + \alpha^{-1} \|\lambda^\nu - \tilde{\lambda}\| \mathbb{E}_\Xi \left[\left(\sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right)^2 \right] \\ & \quad + \alpha^{-1} \|x^\nu - \tilde{x}\| \|\lambda^\nu - \tilde{\lambda}\| \mathbb{E}_\Xi \left[\sum_{j=1}^m \kappa_j(\xi) \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right] \\ & \rightarrow 0, \text{ as } \nu \rightarrow +\infty, \text{ w.p.1.} \end{aligned} \quad (52)$$

Furthermore, by (42), (44), (45), (49), and the second part of the first inequality of (51), we have

$$\begin{aligned} & \mathbb{E}_\Xi \left[\left((\tilde{x} - H(\tilde{x}, \tilde{\lambda}, \xi))^T \left(\sum_{j=1}^m \lambda_j^\nu \mathbb{F}_j(x^\nu, \xi) - \sum_{j=1}^m \tilde{\lambda}_j \mathbb{F}_j(\tilde{x}, \xi) \right) \right) \right] \\ & \leq \frac{2}{\alpha} \mathbb{E}_\Xi \left[\left(\sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| + m\delta(\xi) \right) \left(\|x^\nu - \tilde{x}\| \sum_{j=1}^m \kappa_j(\xi) + \|\lambda^\nu - \tilde{\lambda}\| \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right) \right] \\ & = \frac{2}{\alpha} \|x^\nu - \tilde{x}\| \mathbb{E}_\Xi \left[\sum_{j=1}^m \kappa_j(\xi) \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right] + \frac{2}{\alpha} \|\lambda^\nu - \tilde{\lambda}\| \mathbb{E}_\Xi \left[\left(\sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right)^2 \right] \\ & \quad + \frac{2m}{\alpha} \|x^\nu - \tilde{x}\| \mathbb{E}_\Xi \left[\delta(\xi) \sum_{j=1}^m \kappa_j(\xi) \right] + \frac{2m}{\alpha} \|\lambda^\nu - \tilde{\lambda}\| \mathbb{E}_\Xi \left[\delta(\xi) \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right] \\ & \rightarrow 0, \text{ as } \nu \rightarrow +\infty, \text{ w.p.1.} \end{aligned} \quad (53)$$

It then follows from (A1), (42)–(45), (49), and (51) that

$$\begin{aligned}
& \frac{\alpha}{2} \mathbb{E}_{\Xi} [\|x^{\nu} - H(x^{\nu}, \lambda^{\nu}, \xi)\|^2 - \|\tilde{x} - H(\tilde{x}, \tilde{\lambda}, \xi)\|^2] \\
& \leq \frac{\alpha}{2} \mathbb{E}_{\Xi} [(\|x^{\nu} - H(x^{\nu}, \lambda^{\nu}, \xi)\| + \|\tilde{x} - H(\tilde{x}, \tilde{\lambda}, \xi)\|) \|x^{\nu} - H(x^{\nu}, \lambda^{\nu}, \xi) - \tilde{x} + H(\tilde{x}, \tilde{\lambda}, \xi)\|] \\
& \leq \mathbb{E}_{\Xi} \left[\left(\sum_{j=1}^m \|\mathbb{F}_j(x^{\nu}, \xi)\| + \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| + 2m\delta(\xi) \right) \right. \\
& \quad \cdot (2\|x^{\nu} - \tilde{x}\| + \alpha^{-1}\|x^{\nu} - \tilde{x}\| \sum_{j=1}^m \kappa_j(\xi) + \alpha^{-1}\|\lambda^{\nu} - \tilde{\lambda}\| \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\|) \Big] \\
& \leq \mathbb{E}_{\Xi} \left[\left(\sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| + \sum_{j=1}^m \|\mathbb{F}_j(x^{\nu}, \xi) - \mathbb{F}_j(\tilde{x}, \xi)\| + \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| + 2m\delta(\xi) \right) \right. \\
& \quad \cdot (2\|x^{\nu} - \tilde{x}\| + \alpha^{-1}\|x^{\nu} - \tilde{x}\| \sum_{j=1}^m \kappa_j(\xi) + \alpha^{-1}\|\lambda^{\nu} - \tilde{\lambda}\| \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\|) \Big] \\
& \leq \mathbb{E}_{\Xi} \left[\left(2 \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| + \|x^{\nu} - \tilde{x}\| \sum_{j=1}^m \kappa_j(\xi) + 2m\delta(\xi) \right) \right. \\
& \quad \cdot (2\|x^{\nu} - \tilde{x}\| + \alpha^{-1}\|x^{\nu} - \tilde{x}\| \sum_{j=1}^m \kappa_j(\xi) + \alpha^{-1}\|\lambda^{\nu} - \tilde{\lambda}\| \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\|) \Big] \\
& = 4\|x^{\nu} - \tilde{x}\| \mathbb{E}_{\Xi} [\|\mathbb{F}_j(\tilde{x}, \xi)\|] + 2\|x^{\nu} - \tilde{x}\|^2 \mathbb{E}_{\Xi} \left[\sum_{j=1}^m \kappa_j(\xi) \right] + 4m\|x^{\nu} - \tilde{x}\| \mathbb{E}_{\Xi} [\delta(\xi)] \\
& \quad + \frac{2}{\alpha} \|x^{\nu} - \tilde{x}\| \mathbb{E}_{\Xi} \left[\sum_{j=1}^m \kappa_j(\xi) \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right] + \frac{1}{\alpha} \|x^{\nu} - \tilde{x}\|^2 \mathbb{E}_{\Xi} \left[\left(\sum_{j=1}^m \kappa_j(\xi) \right)^2 \right] \\
& \quad + \frac{2m}{\alpha} \|x^{\nu} - \tilde{x}\| \mathbb{E}_{\Xi} [\delta(\xi) \sum_{j=1}^m \kappa_j(\xi)] + \frac{2}{\alpha} \|\lambda^{\nu} - \tilde{\lambda}\| \mathbb{E}_{\Xi} \left[\left(\sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right)^2 \right] \\
& \quad + \frac{1}{\alpha} \|x^{\nu} - \tilde{x}\| \|\lambda^{\nu} - \tilde{\lambda}\| \mathbb{E}_{\Xi} \left[\sum_{j=1}^m \kappa_j(\xi) \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right] \\
& \quad + \frac{2m}{\alpha} \|\lambda^{\nu} - \tilde{\lambda}\| \mathbb{E}_{\Xi} [\delta(\xi) \sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\|] \\
& \rightarrow 0, \text{ as } \nu \rightarrow +\infty, \text{ w.p.1.} \tag{54}
\end{aligned}$$

Moreover, by $\sum_{j=1}^m \lambda_j \geq 0$, $\sum_{j=1}^m \lambda_j = 1$, (A2), (43), (47), we can obtain

$$\begin{aligned}
& \mathbb{E}_{\Xi} \left[\left| \sum_{j=1}^m \lambda_j^{\nu} g_j(x^{\nu}, \xi) - \sum_{j=1}^m \tilde{\lambda}_j g_j(\tilde{x}, \xi) \right| \right] \\
& \leq \mathbb{E}_{\Xi} \left[\left| \sum_{j=1}^m \lambda_j^{\nu} g_j(x^{\nu}, \xi) - \sum_{j=1}^m \lambda_j^{\nu} g_j(\tilde{x}, \xi) \right| \right] + \mathbb{E}_{\Xi} \left[\left| \sum_{j=1}^m \lambda_j^{\nu} g_j(\tilde{x}, \xi) - \sum_{j=1}^m \tilde{\lambda}_j g_j(\tilde{x}, \xi) \right| \right] \\
& \leq \mathbb{E}_{\Xi} \left[\sum_{j=1}^m |g_j(x^{\nu}, \xi) - g_j(\tilde{x}, \xi)| \right] + \|\lambda^{\nu} - \tilde{\lambda}\| \mathbb{E}_{\Xi} \left[\sum_{j=1}^m |g_j(\tilde{x}, \xi)| \right] \\
& \leq m\|x^{\nu} - \tilde{x}\| \mathbb{E}_{\Xi} [\delta(\xi)] + \|\lambda^{\nu} - \tilde{\lambda}\| \mathbb{E}_{\Xi} \left[\sum_{j=1}^m |g_j(\tilde{x}, \xi)| \right] \\
& \rightarrow 0, \text{ as } \nu \rightarrow +\infty, \text{ w.p.1.} \tag{55}
\end{aligned}$$

Furthermore, similar to (55), we have

$$\begin{aligned}
& \mathbb{E}_{\Xi} \left[\left| \sum_{j=1}^m \lambda_j^v g_j(H(x^v, \lambda^v, \xi), \xi) - \sum_{j=1}^m \tilde{\lambda}_j g_j(H(\tilde{x}, \tilde{\lambda}, \xi), \xi) \right| \right] \\
& \leq m \mathbb{E}_{\Xi} [\delta(\xi) \|H(x^v, \lambda^v, \xi) - H(\tilde{x}, \tilde{\lambda}, \xi)\|] + \|\lambda^v - \tilde{\lambda}\| \mathbb{E}_{\Xi} \left[\sum_{j=1}^m |g_j(H(\tilde{x}, \tilde{\lambda}, \xi), \xi)| \right] \\
& \leq m \|x^v - \tilde{x}\| \mathbb{E}_{\Xi} [\delta(\xi)] + m \alpha^{-1} \|x^v - \tilde{x}\| \mathbb{E}_{\Xi} \left[\sum_{j=1}^m \kappa_j(\xi) \right] \\
& \quad + m \alpha^{-1} \|\lambda^v - \tilde{\lambda}\| \mathbb{E}_{\Xi} \left[\sum_{j=1}^m \|\mathbb{F}_j(\tilde{x}, \xi)\| \right] + \|\lambda^v - \tilde{\lambda}\| \mathbb{E}_{\Xi} \left[\sum_{j=1}^m |g_j(H(\tilde{x}, \tilde{\lambda}, \xi), \xi)| \right] \\
& \rightarrow 0, \text{ as } v \rightarrow +\infty, \text{ w.p.1,} \tag{56}
\end{aligned}$$

where the second inequality follows from (51) and the limit holds following from (43), (45), and (47). As maintained by (39) and (52)–(56), the following holds

$$\begin{aligned}
& |\mathbb{E}_{\Xi_v} [h(x^v, \lambda^v, \xi) - h(\tilde{x}, \tilde{\lambda}, \xi)]| \\
& \leq \mathbb{E}_{\Xi_v} [|h(x^v, \lambda^v, \xi) - h(\tilde{x}, \tilde{\lambda}, \xi)|] \\
& \leq \mathbb{E}_{\Xi} [|h(x^v, \lambda^v, \xi) - h(\tilde{x}, \tilde{\lambda}, \xi)|] \\
& \leq \mathbb{E}_{\Xi} [(x^v - \tilde{x} + H(\tilde{x}, \tilde{\lambda}, \xi) - H(x^v, \lambda^v, \xi))^T \sum_{j=1}^m \lambda_j^v \mathbb{F}_j(x^v, \xi)] \\
& \quad + \mathbb{E}_{\Xi} [(\tilde{x} - H(\tilde{x}, \tilde{\lambda}, \xi))^T (\sum_{j=1}^m \lambda_j^v \mathbb{F}_j(x^v, \xi) - \sum_{j=1}^m \tilde{\lambda}_j \mathbb{F}_j(\tilde{x}, \xi))] \\
& \quad + \mathbb{E}_{\Xi} [\sum_{j=1}^m \lambda_j^v g_j(x^v, \xi) - \sum_{j=1}^m \tilde{\lambda}_j g_j(\tilde{x}, \xi)] \\
& \quad + \mathbb{E}_{\Xi} [\sum_{j=1}^m \lambda_j^v g_j(H(x^v, \lambda^v, \xi), \xi) - \sum_{j=1}^m \tilde{\lambda}_j g_j(H(\tilde{x}, \tilde{\lambda}, \xi), \xi)] \\
& \quad + \frac{\alpha}{2} \mathbb{E}_{\Xi} [\|x^v - H(x^v, \lambda^v, \xi)\|^2 - \|\tilde{x} - H(\tilde{x}, \tilde{\lambda}, \xi)\|^2] \\
& \rightarrow 0, \text{ as } v \rightarrow +\infty, \text{ w.p.1.}
\end{aligned}$$

Thus, the formulation (50) holds.

(ii) We next show that $\lim_{v \rightarrow \infty} \mathbb{E}_{\Xi_v} [h(x^v, \lambda^v, \xi)] = \mathbb{E}_{\Xi} [h(\tilde{x}, \tilde{\lambda}, \xi)]$. By Lemma 1 and (39), we can easily obtain that

$$\lim_{v \rightarrow \infty} |\mathbb{E}_{\Xi} [h(\tilde{x}, \tilde{\lambda}, \xi)] - \mathbb{E}_{\Xi_v} [h(\tilde{x}, \tilde{\lambda}, \xi)]| = 0, \text{ w.p.1.} \tag{57}$$

Hence, by (50) and (57), we have

$$\begin{aligned}
& |\mathbb{E}_{\Xi_v} [h(x^v, \lambda^v, \xi)] - \mathbb{E}_{\Xi} [h(\tilde{x}, \tilde{\lambda}, \xi)]| \\
& \leq |\mathbb{E}_{\Xi_v} [h(x^v, \lambda^v, \xi)] - \mathbb{E}_{\Xi_v} [h(\tilde{x}, \tilde{\lambda}, \xi)]| + |\mathbb{E}_{\Xi_v} [h(\tilde{x}, \tilde{\lambda}, \xi)] - \mathbb{E}_{\Xi} [h(\tilde{x}, \tilde{\lambda}, \xi)]| \\
& \rightarrow 0, \text{ as } v \rightarrow +\infty, \text{ w.p.1.}
\end{aligned}$$

Therefore, the conclusion holds. $\square$

Next, we show the main convergence result and denote the sets of global optimal solutions of problems (25) and (40) by $\tilde{\mathcal{O}}$ and $\mathcal{O}_v$, respectively.

Theorem 4. Suppose that conditions (A1)–(A2) hold. For each v , let $(x^v, \lambda^v, u^v) \in \mathcal{O}_v$ and $(\tilde{x}, \tilde{\lambda}, \tilde{u})$ is an accumulation point of (x^v, λ^v, u^v) . Then, we have $(\tilde{x}, \tilde{\lambda}, \tilde{u}) \in \tilde{\mathcal{O}}$ w.p.1.

Proof. We assume that $\lim_{\nu \rightarrow \infty} x^\nu = \tilde{x}$, $\lim_{\nu \rightarrow \infty} \lambda^\nu = \tilde{\lambda}$, $\lim_{\nu \rightarrow \infty} u^\nu = \tilde{u}$. By the inequality $||s|_+ - |t|_+| \leq |s - t|$ given in reference [34] and Lemma 2, we have

$$\begin{aligned} & |\mathbb{E}_{\Xi_\nu} [h(x^\nu, \lambda^\nu, \xi) - u^\nu]_+ - \mathbb{E}_{\Xi_\nu} [h(\tilde{x}, \tilde{\lambda}, \xi) - \tilde{u}]_+| \\ & \leq \mathbb{E}_{\Xi_\nu} [|h(x^\nu, \lambda^\nu, \xi) - u^\nu|_+ - |h(\tilde{x}, \tilde{\lambda}, \xi) - \tilde{u}|_+] \\ & \leq \mathbb{E}_{\Xi_\nu} [|h(x^\nu, \lambda^\nu, \xi) - u^\nu - h(\tilde{x}, \tilde{\lambda}, \xi) + \tilde{u}|] \\ & \leq \mathbb{E}_{\Xi_\nu} [|h(x^\nu, \lambda^\nu, \xi) - h(\tilde{x}, \tilde{\lambda}, \xi)|] + |u^\nu - \tilde{u}| \\ & \rightarrow 0, \text{ as } \nu \rightarrow +\infty, \text{ w.p.1.} \end{aligned} \quad (58)$$

It follows from (58) and Lemma 2 that

$$\begin{aligned} & |\theta_\nu(x^\nu, \lambda^\nu, u^\nu) - \theta_\nu(\tilde{x}, \tilde{\lambda}, \tilde{u})| \\ & \leq (1-r) |\mathbb{E}_{\Xi_\nu} [h(x^\nu, \lambda^\nu, \xi)] - \mathbb{E}_{\Xi_\nu} [h(\tilde{x}, \tilde{\lambda}, \xi)]| + r |u^\nu - \tilde{u}| \\ & \quad + \frac{r}{1-\beta} |\mathbb{E}_{\Xi_\nu} [h(x^\nu, \lambda^\nu, \xi) - u^\nu]_+ - \mathbb{E}_{\Xi_\nu} [h(\tilde{x}, \tilde{\lambda}, \xi) - \tilde{u}]_+| \\ & \rightarrow 0, \text{ as } \nu \rightarrow +\infty, \text{ w.p.1.} \end{aligned} \quad (59)$$

On the other hand, similar to (57), we know that

$$\lim_{\nu \rightarrow +\infty} |\mathbb{E}_{\Xi} [h(\tilde{x}, \tilde{\lambda}, \xi) - \tilde{u}]_+ - \mathbb{E}_{\Xi_\nu} [h(\tilde{x}, \tilde{\lambda}, \xi) - \tilde{u}]_+| = 0, \text{ w.p.1.} \quad (60)$$

Consequently, it holds that

$$\begin{aligned} & |\theta_\nu(\tilde{x}, \tilde{\lambda}, \tilde{u}) - \theta(\tilde{x}, \tilde{\lambda}, \tilde{u})| \\ & \leq (1-r) |\mathbb{E}_{\Xi} [h(\tilde{x}, \tilde{\lambda}, \xi)] - \mathbb{E}_{\Xi_\nu} [h(\tilde{x}, \tilde{\lambda}, \xi)]| \\ & \quad + \frac{r}{1-\beta} |\mathbb{E}_{\Xi} [h(\tilde{x}, \tilde{\lambda}, \xi) - \tilde{u}]_+ - \mathbb{E}_{\Xi_\nu} [h(\tilde{x}, \tilde{\lambda}, \xi) - \tilde{u}]_+| \\ & \rightarrow 0, \text{ as } \nu \rightarrow +\infty, \text{ w.p.1.} \end{aligned} \quad (61)$$

Moreover, by (59) and (61), it holds that

$$\begin{aligned} & |\theta_\nu(x^\nu, \lambda^\nu, u^\nu) - \theta(\tilde{x}, \tilde{\lambda}, \tilde{u})| \\ & \leq |\theta_\nu(x^\nu, \lambda^\nu, u^\nu) - \theta_\nu(\tilde{x}, \tilde{\lambda}, \tilde{u})| + |\theta_\nu(\tilde{x}, \tilde{\lambda}, \tilde{u}) - \theta(\tilde{x}, \tilde{\lambda}, \tilde{u})| \\ & \rightarrow 0, \text{ as } \nu \rightarrow +\infty, \text{ w.p.1.} \end{aligned} \quad (62)$$

Notice the fact that $\forall (x, \lambda, u) \in S$, $\theta(x, \lambda, u) \geq 0$, since $(x^\nu, \lambda^\nu, u^\nu) \in \mathcal{O}_\nu$, for each ν , we have

$$\theta_\nu(x^\nu, \lambda^\nu, u^\nu) \leq \theta_\nu(x, \lambda, u) \leq \theta(x, \lambda, u). \quad (63)$$

Taking $\nu \rightarrow \infty$ to (63), according to (61), it holds that

$$\theta(\tilde{x}, \tilde{\lambda}, \tilde{u}) \leq \theta(x, \lambda, u), \quad \forall (x, \lambda, u) \in S,$$

which means that $(\tilde{x}, \tilde{\lambda}, \tilde{u}) \in \tilde{\mathcal{O}}$ w.p.1. $\square$

5. Conclusions

We established the equivalence between SMVVIP and stochastic mixed scalar variational inequality problems. Based on this result, we first proposed a deterministic bi-criteria model (25) for solving SMVVIP. Subsequently, we employed the smoothing technique to derive the smooth approximation problem (27) from (25) and utilized sample average approximation techniques to solve (27). In addition, we have discussed the convergence behavior of the global optimal solutions under the assumption that the sample space

Ξ is compact. When Ξ is not a compact set, the convergence analysis of the compact approximation problem has been considered as well.

Author Contributions: Funding acquisition, M.L.; Supervision, M.L.; Writing—original draft, M.L. and M.D.; writing—review and editing, M.L. and Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Natural Science Foundation General Project of the Department of Liaoning Province Science and Technology (No. 2021-MS-153).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Smith, A. *An Inquiry into the Nature and Causes of the Wealth of Nations*; University of Chicago Press: Chicago, IL, USA, 1977.
- Liu, S.M.; Feng, E. The exponential penalty function method for multiobjective programming problems. *Optim. Methods Softw.* **2010**, *25*, 667–675. [\[CrossRef\]](#)
- Cerda-Flores, S.C.; Rojas-Punzo, A.A.; Nápoles-Rivera, F. Applications of Multi-Objective Optimization to Industrial Processes: A Literature Review. *Processes* **2022**, *10*, 133. [\[CrossRef\]](#)
- Han, S.; Zhu, K.; Zhou, M.C.; Cai, X. Information-Utilization-Method-Assisted Multimodal Multiobjective Optimization and Application to Credit Card Fraud Detection. *IEEE Trans. Comput. Soc. Syst.* **2021**, *8*, 856–869. [\[CrossRef\]](#)
- Lyons, M.A. Pretomanid dose selection for pulmonary tuberculosis: An application of multi-objective optimization to dosage regimen design. *CPT Pharmacometrics Syst. Pharmacol.* **2021**, *10*, 211–219. [\[CrossRef\]](#) [\[PubMed\]](#)
- Li, M.; Wang, L.; Wang, Y.; Chen, Z. Sizing optimization and energy management strategy for hybrid energy storage system using multiobjective optimization and random forests. *IEEE Trans. Power Electron.* **2021**, *36*, 11421–11430. [\[CrossRef\]](#)
- Pereira, J.L.J.; Oliver, G.A.; Francisco, M.B.; Cunha, S.S.; Gomes, G.F. A Review of Multi-objective Optimization: Methods and Algorithms in Mechanical Engineering Problems. *Arch. Comput. Methods Eng.* **2022**, *29*, 2285–2308. [\[CrossRef\]](#)
- Bhaskar, V.; Gupta, S.K.; Ray, A.K. Applications of multiobjective optimization in chemical engineering. *Rev. Chem. Eng.* **2000**, *16*, 1–54. [\[CrossRef\]](#)
- Salahuddin; Ahmad, M.K.; Verma, R.U. Existence solutions to mixed vector variational-type inequalities. *Adv. Nonlinear Var.* **2013**, *16*, 115–123.
- Xie, P.D.; Salahuddin. Generalized mixed general vector variational-like inequalities in topological vector spaces. *J. Nonlinear Analysis Optim. Theory Appl.* **2013**, *4*, 163–172.
- Irfan, S.S. Exponential Type Mixed Vector Variational Inequalities with Fuzzy Mappings. *Panam. Math. J.* **2020**, *30*, 11–26.
- Jayswal, A.; Kumari, B.; Ahmad, I. A gap function and existence of solutions for a non-smooth vector variational inequality on Hadamard manifolds. *Optimization* **2021**, *70*, 1875–1889. [\[CrossRef\]](#)
- Chang, S.; Salahuddin; Wang, L.; Ma, Z.L. Error bounds for mixed set-valued vector inverse quasi-variational inequalities. *J. Inequalities Appl.* **2020**, *1*, 160. [\[CrossRef\]](#)
- Liu, D.Y.; Jiang, Y. Scalarization of Mixed Vector Variational Inequalities and Error Bounds of Gap Functions. *Appl. Math. Mech.* **2017**, *38*, 715–726.
- Salahuddin. Set valued exponential type g-mixed vector variational inequality problems. *Commun. Appl. Nonlinear Anal.* **2017**, *24*, 59–72.
- Fukushima, M. Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. *Math. Program.* **1992**, *53*, 99–110. [\[CrossRef\]](#)
- Luo, M.J.; Lin, G.H. Expected residual minimization method for stochastic variational inequality problems. *J. Optim. Theory Appl.* **2009**, *140*, 103–116. [\[CrossRef\]](#)
- Shanbhag, U.V. Stochastic Variational Inequality Problems: Applications, Analysis, and Algorithms. In *Theory Driven by Influential Applications*; Topaloglu, H., Smith, J.C., Greenberg, H.J., Eds.; Informs: Catonsville, MD, USA, 2013; pp. 71–107.
- Chen, X.J.; Lin, G.H. CVaR-based formulation and approximation method for stochastic variational inequalities. *Numer. Algebr. Control Optim.* **2011**, *1*, 35–48. [\[CrossRef\]](#)
- Bot, R.T.; Mertikopoulos, O.; Staudigl, M.; Vuong, P.T. Minibatch forward-backward-forward methods for solving stochastic variational inequalities. *Stoch. Syst.* **2021**, *11*, 112–139. [\[CrossRef\]](#)
- Jadamba, B.; Khan, A.A.; Sama, M.; Yang, Y. An iteratively regularized stochastic gradient method for estimating a random parameter in a stochastic PDE. A variational inequality approach. *J. Nonlinear Var. Anal.* **2021**, *5*, 865–880.
- Luo, M.J.; Lin, G.H. Convergence results of the ERM method for nonlinear stochastic variational inequality problems. *J. Optim. Appl.* **2009**, *142*, 569–581. [\[CrossRef\]](#)
- Wang, S.H.; Lin, R.G.; Shi, Y.T. A stochastic selective projection method for solving a constrained stochastic variational inequality problem. *J. Appl. Numer. Optim.* **2022**, *4*, 341–355.
- Yang, Z.P.; Zhang, J.; Zhu, X.D.; Lin, G.H. Infeasible interior-point algorithms based on sampling average approximations for a class of stochastic complementarity problems and their applications. *J. Comput. Appl. Math.* **2019**, *352*, 382–400. [\[CrossRef\]](#)
- Fukushima, M. *Fundamentals of Nonlinear Optimization*; Asakura Shoten: Tokyo, Japan, 2001.

26. Solodov, M.V. Merit functions and error bounds for generalized variational inequalities. *J. Math. Anal. Appl.* **2003**, *287*, 405–414. [[CrossRef](#)]
27. Bauschke, H.H.; Combettes, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*; Springer: New York, NY, USA, 2011.
28. Rockafellar, R.T.; Uryasev, S. Optimization of conditional value-at-risk. *J. Risk* **2000**, *2*, 21–42. [[CrossRef](#)]
29. Chen, L.; Lin, G.H.; Yang, X.M. Expected Residual Minimization Method for Stochastic Mixed Variational Inequality Problems. *Pac. J. Optim.* **2017**, *14*, 703–722.
30. Fukushima, M. Merit functions for variational inequality and complementarity problems. In *Nonlinear Optimization and Applications*; Di Pillo, G., Giannessi, F., Eds.; Plenum Press: New York, NY, USA, 1996; pp. 155–170.
31. Yang, X.M.; Zhao, Y.; Lin, G.H. Deterministic Bicriteria Model for Stochastic Variational Inequalities. *J. Oper. Res. Soc. China* **2018**, *6*, 507–527. [[CrossRef](#)]
32. Ehrgott, M. *Multicriteria Optimization*; Springer Science & Business Media: Heidelberg, Germany, 2005.
33. Peng, J.M. *A Smoothing Function and Its Applications. Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*; Springer: Boston, MA, USA, 1998; pp. 293–316.
34. Chen, C.H.; Mangasarian, O.L. A class of smoothing function for nonlinear and mixed complementarity problems. *Comput. Optim. Appl.* **1996**, *5*, 97–138. [[CrossRef](#)]
35. Luo, M.J.; Zhang, K. Convergence Analysis of the Approximation Problems for Solving Stochastic Vector Variational Inequality Problems. *Complexity* **2020**, *2020*, 1203627. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.