COMMUN. MATH. SCI.
Vol. 16, No. 3, pp. 659–704

A UNIFIED SYSTEM OF FB-SDES WITH LÉVY JUMPS AND DOUBLE COMPLETELY-$\mathcal{S}$ SKEW REFLECTIONS*

WANYANG DAI†

Abstract. We study the well-posedness of a unified system of coupled forward-backward stochastic differential equations (FB-SDEs) with Lévy jumps and double completely-$\mathcal{S}$ skew reflections. Owing to the reflections, the solution to an embedded Skorohod problem may be not unique, i.e., bifurcations may occur at reflection boundaries and the well-known contraction mapping approach can not be extended directly to solve our problem. Thus, we develop a weak convergence method to prove the well-posedness of an adapted 6-tuple weak solution in the sense of distribution to the unified system. The proof heavily depends on newly established Malliavin calculus for vector-valued Lévy processes together with a generalized linear growth and Lipschitz condition that guarantees the well-posedness of the unified system even under a random environment. Nevertheless, if a stricter boundary condition is imposed, i.e., the spectral radii of each square submatrix at a corner of the reflections are strictly less than unity, a unique adapted 6-tuple strong solution (in the sense of sample paths) is considered. In addition, as applications and economic studies of our unified system, we also develop new techniques including deriving a generalized mutual information formula for signal processing over possible non-Gaussian channels with multi-input multi-output (MIMO) antennas and dynamics driven by Lévy processes.

Keywords. Stochastic differential equation, Lévy jump, completely-$\mathcal{S}$ skew reflection, Skorohod problem, weak convergence, Malliavin calculus, mutual information, queueing network, Big Data, economic modeling.

AMS subject classifications. 60H10, 60J75, 60K37, 94A17, 60K25.

1. Introduction

In this paper, we aim to prove the well-posedness of an adapted 6-tuple $((X,Y),(V,\bar{V},\tilde{V},F))$ weak solution (in the sense of distributions) to the non-Markovian system of coupled forward-backward stochastic differential equations (FB-SDEs) with Lévy jumps and double completely-$\mathcal{S}$ skew reflections under a given control rule u over time interval $[0,T]$,

$$\begin{cases} X(t) = \xi + Z(t) + RY(t), \\ V(t) = H(X(T),*) + U(t) + S(F(T) - F(t)), \end{cases} \tag{1.1}$$

where X in the forward equation is endowed with the given initial value ξ, and V in the backward equation is endowed with the known terminal value form of $H(X(T),*)$. Furthermore, the "$*$" in $H(X(T),*)$ denotes known random factors that can be explicitly expressed in terms of the driving Brownian motion and/or the Lévy process that will be defined in Section 2.

In many real-world applications (e.g., in quantum physics, queueing systems, and economics), the processes X and V in system (1.1) are referred as the state process

*Received: November 23, 2016; accepted (in revised form): October 20, 2017. Communicated by David Anderson.

†Department of Mathematics and State Key Laboratory of Novel Software Technology, Nanjing University, Nanjing 210093, P.R. China (nan5lu8@nju.edu.cn). Personal Web: http://math.nju.edu.cn/~wydai/.

This research is supported by National Natural Science Foundation of China with Grant No. 11771006 and Grant No. 11371010. The partial results in the paper were presented as invited talks at various international conferences. The author thanks the helpful comments from the participants and the editors and the anonymous reviewer.

and its associated value process. In particular, the processes Z and U are called state-dependent netput processes. For our purpose of study, we unify existing specific discussions on Z and U to the generalized form with Lévy jumps and feedback control law u in a forward and backward coupling manner as follows,

$$\begin{cases} dZ(t)= b(t^-,X,V,\bar{V},\tilde{V},u)dt+\sigma(t^-,X,V,\bar{V},\tilde{V},u)dW(t) \\ \qquad\qquad +\int_{\mathcal{Z}^h}\eta(t^-,X,V,\bar{V},\tilde{V},u,z)\tilde{N}(dt,dz), \\ dU(t)= c(t^-,X,V,\bar{V},\tilde{V},u)dt-\alpha(t^-,X,V,\bar{V},\tilde{V},u)dW(t) \\ \qquad\qquad -\int_{\mathcal{Z}^h}\zeta(t^-,X,V,\bar{V},\tilde{V},u,z)\tilde{N}(dt,dz), \end{cases} \tag{1.2}$$

where $t\in[0,T]$ and $\mathcal{Z}^h$ is the product of h number of $R-\{0\}$ (i.e., $(R-\{0\})\times\cdots\times(R-\{0\})$) or is $R_+^h=[0,\infty)^h$. The process Z in (1.2) is in a forward manner with initial condition $Z(0)=0$ while the process U is in a backward manner with terminal condition $U(T)=0$. Furthermore, $\bar{V}$ and $\tilde{V}$ in (1.2) are two unknown regulating processes that are part of our 6-tuple solution to be determined. In addition, W is a standard continuous Brownian motion and $\tilde{N}$ is a centered jump Lévy process. Note that for each functional $f\in\{b,\sigma,c,\alpha\}$,

$$f(t,X,V,\bar{V},\tilde{V},u,z)\equiv f(t,X(t),V(t),\bar{V}(t),\tilde{V}(t,\cdot),u(t,X(t)),*), \tag{1.3}$$

where the dot "$\cdot$" in $\tilde{V}(t,\cdot)$ denotes integration in terms of the Lévy measure, and the "$*$" in (1.3) represents known random factors that can be explicitly expressed in terms of W and L. However, if $f\in\{\eta,\zeta\}$, the right-hand side of (1.3) should be changed to

$$f(t,X(t),V(t),\bar{V}(t),\tilde{V}(t,z),u(t,X(t)),z,*). \tag{1.4}$$

The processes Y and F in system (1.1) are called boundary regulators (e.g., loss potential processes in practical systems), which are given by

$$\begin{cases} Y_i(t)= \int_0^t I_{D_i}(X(s))dY_i(s), \\ F_i(t)= \int_0^t I_{\bar{D}_i}(V(s))dF_i(s). \end{cases} \tag{1.5}$$

More precisely, X in system (1.1) is a p-dimensional process governed by the F-SDE with skew reflection matrix R, and V in system (1.1) is a q-dimensional process governed by the B-SDE with skew reflection matrix S. Furthermore, Y can increase only when X is on a boundary D_i, $i\in\{1,...,b\}$ and F can increase only when V is on a boundary $\bar{D}_i$, $i\in\{1,...,\bar{b}\}$, where b and $\bar{b}$ are two nonnegative integers. Both Y and F are the regulating processes with possible jumps to push X and V back into the state spaces D and $\bar{D}$ respectively.

It is not difficult to understand the existence of a solution to the system (1.1)–(1.5) when the driving processes are continuous Brownian motions (see the related discussion in [46]). However, the existence of a solution if the driving processes are general Lévy processes with jumps, in particular when the state space D or $\bar{D}$ is bounded, is not obvious. Our interpretation to this existence problem is by way of Skorohod decompositions as used in diffusion approximations for queueing networks (see [12,13], and [10]). In other words, for two given netput processes Z and U, we try to find two pairs of Skorohod decomposed processes (X,Y) and (V,F) satisfying the required properties. From the viewpoint of a sample path, this existence issue is consistent with the well-known Skorohod problem (see [12], [10]). Note that the processes Y and F are parts of the 6-tuple solution to system (1.1)–(1.5), which are frequently called Skorohod regulators (see, Figure 1.1 for an example).

The proof for the well-posedness of an adapted 6-tuple weak solution to the system of FB-SDEs in (1.1)–(1.5) is based on a general boundary reflection condition (called the *completely-$\mathcal{S}$* condition in physical queueing systems, see [12], [10], [11]). Besides the well-known mirror reflection, the condition is concerned with a general non-symmetric skew reflection over boundaries (see Figure 1.1 for such an illustration). Under this

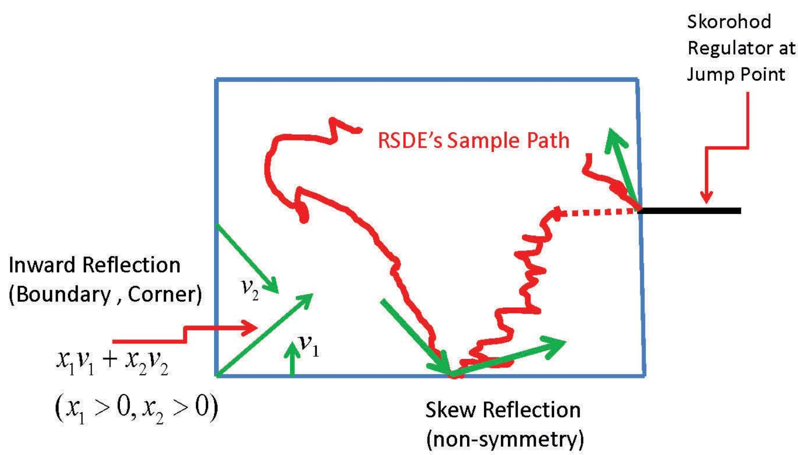

FIG. 1.1. *Skew and Inward Reflection with Skorohod Regulator under Completely-$\mathcal{S}$ Condition.*

condition, the solution to an embedded Skorohod problem may be not unique, i.e., bifurcations may occur at reflection boundaries for quantum particles moving along the solution paths. Thus, this non-uniqueness property leads to difficulties in our proof, e.g., the conventional contraction mapping approach by using the well-known Picard's iteration can not be directly extended to solve our problem. However, by proving a stricter oscillation inequality with some related property in the Skorohod topological space, and by establishing Malliavin calculus for vector-valued Lévy processes, we develop a general weak convergence method to prove the well-posedness of a 6-tuple weak solution (in the sense of distributions) to system (1.1)–(1.5).

It is worth pointing out that the coefficients of system (1.2) are functionals of the coupled forward-backward processes, which brings additional complexity to our analysis. Furthermore, in our proof, the conventional linear growth and Lipschitz constant is replaced by an adapted stochastic process that may be unbounded but is mean-square integrable (see, [14,16]), which guarantees the well-posedness of system (1.1)–(1.5) even under an external random environment. In addition, if the completely-$\mathcal{S}$ condition becomes more strict, e.g., with additional requirements that the spectral radii of each square submatrix at a corner of a reflection are strictly less than unity, a unique adapted 6-tuple strong solution in the sample pathwise sense will be considered.

Coupled FB-SDEs motivates an active area of research (see, [39] for a discussion of coupled FB-SDEs with no boundary reflection, and [29] for a study of Brownian motion driven B-SDE with reflection, and see references therein). However, to the best of our knowledge, the unified system of coupled FB-SDEs in (1.1)–(1.5) with double skew reflection matrices is new and should be the most general form of various existing SDEs. Furthermore, the study of the well-posedness in terms of an adapted 6-tuple weak solution (in the sense of distributions) with Lévy jumps and under a general completely-$\mathcal{S}$ reflection condition through the Skorohod problem are also new.

For the purpose of further illustrating the importance of our unified system in (1.1)–(1.5), we study the applications of its adapted solution in signal processing over communication channels. In a communication system, or typically in a cloud-computing associated communication system, the Lévy processes correspond to white non-Gaussian noises and can be used to model Big Data (see [18]) in its three-dimensional statistical features: high-volume (amount of data) with batch arrivals, high-velocity (speed of data in and out) corresponding to batch processing services, and high-variety (range of data types and sources). Skew reflections are due to buffer storage constraint and heavily loaded traffic, or due to system idleness and lightly loaded traffic.

To optimize service over an MIMO communication channel, the determination of its service capacity is crucial. The capacity is originally derived as the Shannon capacity for a communication channel with a white Gaussian noise corresponding to a continuous Brownian motion. The key step in obtaining the capacity is the calculation of the so-called mutual information, i.e., the information contained in the received signal about the transmitted signal over the channel (see [9]). Recently, the formula to compute the mutual information was extended to a single-input single-output (SISO) channel presented by a stochastic equation (a degenerated SDE without feedback) driven by a white non-Gaussian noise corresponding to a pure jump Lévy process (see, [19]).

Thus, we also aim to derive and prove a generalized formula of mutual information for a multi-input multi-output (MIMO) channel modeled by general nonlinear SDEs with feedback, which are driven by both continuous Brownian motions and pure jump Lévy processes.

MIMO channels are the major technology in current and likely future wireless and quantum communication systems. The dynamics governing the MIMO channels can be modeled as a queueing network consisting of arrival processes, service processes, and data buffer storages with certain service regimes and network architectures (see an example with p-users in Figure 1.2). The queue length processes in this network are the real signal processes to be transmitted over the MIMO channels. Thus, we refine two generalized queueing models under different assumptions and illustrate their well-posedness by the solution to system (1.1)–(1.5). Furthermore, based on the generalized formula of mutual information, we can derive the channel capacity region for an MIMO channel (or channels) with multiple users. Specifically, one can get the capacity region that is achievable under a coding (e.g., the dirty paper coding) technique in communication practices when the queueing signals are approximated by models driven by white Gaussian noise processes (see [23], [27]).

The remainder of the paper is organized as follows. In Section 2, we introduce our unified system of coupled FB-SDEs and state the main theorem about its well-posedness. In Section 3, we study the applications and related economical modeling of our system in signal processing over MIMO wireless channels. Particularly, we derive a generalized formula of MIMO mutual information with Lévy jumps. In Section 4, we prove our main results. In Section 5, we present the conclusion of the paper.

2. The unified FB-SDE and its well-posedness

2.1. Conditions on state spaces. We assume that the process X governed by the forward SDE in (1.1)–(1.5) lives in a state space D that is a general convex polyhedral (see [11], [12], [10]). Specific examples of a convex polyhedral include the p-dimensional positive orthant and the p-dimensional rectangle as displayed in Figure 1.2. There are b boundary faces for the polyhedral with a given integer $b \in \{0,1,2,...\}$; the i-th face is denoted by $D_i = \{x \in R^p, x \cdot n_i = b_i\}$ for $i \in \{1,...,b\}$, where b_i is some non-negative constant and n_i is the inward unit normal vector on the boundary face D_i.

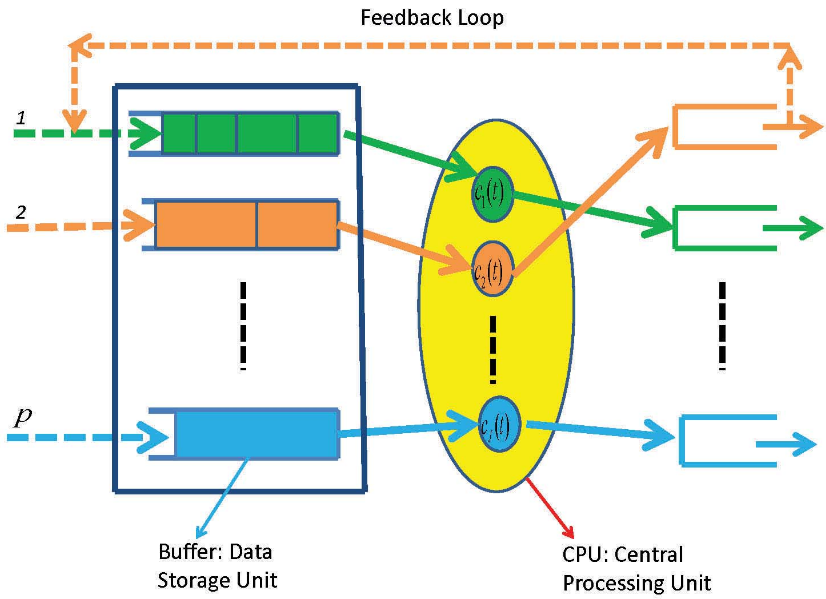

FIG. 1.2. *A queueing network system with p-job classes.*

For convenience, we define $N=(n_1,...,n_b)$ and let R in (1.1)–(1.5) be a $p\times b$ matrix, whose i-th column denoted by p-dimensional vector v_i is the reflection direction of X on D_i. The process Y in (1.1)–(1.5) is a nondecreasing predictable process with $Y(0)=0$ and boundary regulating property as explained in (1.1)–(1.5). In queueing system, this process is called the boundary idle time or blocking process.

Analogously, we suppose that V takes values in a region $\bar{D}$ with boundary face $\bar{D}_i=\{v\in R^q, v\cdot\bar{n}_i=\bar{b}_i\}$ for $i\in\{1,...,\bar{b}\}$ and a known $\bar{b}\in\{0,1,2,...\}$, where $\bar{n}_i$ is the inward unit normal vector on the boundary face $\bar{D}_i$. For convenience, we define $\bar{N}=(\bar{n}_1,...,\bar{n}_{\bar{b}})$. In finance, the given constant $\bar{b}_i$ is called the early exercise reward. Furthermore, S in (1.1) is a $q\times\bar{b}$ matrix. In addition, $F(\cdot)$ in (1.1) is a nondecreasing predictable process with $F(0)=0$ and boundary regulating property as explained in (1.1)–(1.5).

Associated with the reflection matrix R (and similarly for S), we impose the following completely-$\mathcal{S}$ condition.

DEFINITION 2.1. *A $p\times p$ square matrix R is called completely-$\mathcal{S}$ if and only if there is $x>0$ such that $\tilde{R}x>0$ for each principal submatrix $\tilde{R}$ of R (i.e., $\tilde{R}$ is a submatrix in which the set of row indices that remain is the same as the set of column indices that remain), where the vector inequalities are to be interpreted componentwise. Furthermore, a $p\times b$ matrix R is called completely-$\mathcal{S}$ if and only if each $p\times p$ square submatrix of $N'R$ and $(N'R)'$ is completely-S.*

Note that the completely-$\mathcal{S}$ condition on the reflection matrices guarantees that the coupled FB-SDEs are of inward reflection on each boundary and corner of the polyhedral (see Figure 1.1 and [12]). Furthermore, the reflection appearing here is called a skew reflection, which is a generalization of the conventional mirror (or symmetry) reflection.

2.2. Assumptions on the system and its coefficients. To talk about the existence and uniqueness of an adapted 6-tuple strong or weak solution to the coupled FB-SDEs in (1.1)–(1.5), we need to introduce the required probability and supporting topological spaces. More precisely, let $(\Omega,\mathcal{F},P)$ be a complete probability space on which we define a standard d-dimensional Brownian motion $W\equiv\{W(t),t\in[0,T]\}$ for a given $T\in[0,\infty)$ with $W(t)=(W_1(t),\ldots,W_d(t))'$ and a h-dimensional general Lévy pure jump process (or special subordinator) $L\equiv\{L(t),t\in[0,T]\}$ with $L(t)\equiv(L_1(t),\ldots,L_h(t))'$ (see [2], [4], and [43]). Note that the prime notation appearing in this paper is used to denote the transpose of a matrix or of a vector. Furthermore, W, L, and their components are assumed to be independent of each other. For each vector $\lambda=(\lambda_1,\ldots\lambda_h)'>0$, which is called a reversion rate vector in many applications, we let $L(\lambda s)=(L_1(\lambda_1 s),\ldots,L_h(\lambda_h s))'$. Then, we denote a filtration by $\{\mathcal{F}_t\}_{t\geq 0}$ with $\mathcal{F}_t\equiv\sigma\{\mathcal{G},W(s),L(\lambda s):0\leq s\leq t\}$ for each $t\in[0,T]$, where $\mathcal{G}$ is σ-algebra independent of W and L. In addition, let $I_A(\cdot)$ be the index function over the set A, and let ν_i be a Lévy measure for each $i\in\{1,\ldots,h\}$. Then, we denote by $N_i((0,t]\times A)\equiv\sum_{0<s\leq t}I_A(L_i(s)-L_i(s^-))$ a Poisson random measure with a deterministic, time-homogeneous intensity measure $ds\nu_i(dz_i)$. Thus, each L_i can be represented by

$$L_i(t)=a_i(t)+\int_{(0,t]}\int_{\mathcal{Z}}z_iN_i(ds,dz_i),\ \ t\geq 0. \tag{2.1}$$

(see Theorem 13.4 and Corollary 13.7 in [28]). For convenience, we take the constant a_i to be zero, and for later reference, we define

$$\nu(dz)=(\nu_1(dz_1),\ldots,\nu_h(dz_h))'. \tag{2.2}$$

Furthermore, for each $t\in[0,T]$ and $z\in\mathcal{Z}^h$, we let

$$\tilde{N}(\lambda dt,dz)=(\tilde{N}_1(\lambda_1 dt,dz_1),\ldots,\tilde{N}_h(\lambda_h dt,dz_h))' \tag{2.3}$$

with

$$\tilde{N}_i(\lambda_i dt,dz_i)=N_i(\lambda_i dt,dz_i)-\lambda_i dt\nu_i(dz_i) \tag{2.4}$$

for each $i\in\{1,\ldots,h\}$.

Next, based on the driven Brownian motion and Lévy process, the required supporting topological space can be defined by

$$\begin{aligned}\mathcal{Q}^2_{\mathcal{F}}([0,T])\equiv{}& D^2_{\mathcal{F}}([0,T],R^p)\times D^2_{\mathcal{F}}([0,T],R^b)\\ &\times D^2_{\mathcal{F}}([0,T],R^q)\times D^2_{\mathcal{F},p}([0,T],R^{q\times d})\\ &\times D^2_{\mathcal{F},p}([0,T]\times\mathcal{Z}^h,R^{q\times h})\times D^2_{\mathcal{F}}([0,T],R^{\bar{b}}),\end{aligned} \tag{2.5}$$

where for a $q\times d$ matrix-valued process $\bar{V}$ or a $q\times h$ matrix-valued process $\tilde{V}$, we consider it as a qd-dimensional vector-valued process or a qh-dimensional vector-valued process whenever a norm is considered. Furthermore, for a positive integer in the set $\{p,b,q,qd,qh,\bar{b}\}$, (e.g., the integer b), we let $D^2_{\mathcal{F}}([0,T],R^b)$ be the space of R^b-valued and $\{\mathcal{F}_t\}$-adapted processes with sample paths in the Skorohod topological space $D([0,T],R^b)$ (i.e., each sample path in the space is right-continuous with left-limits and the space itself is endowed with the Skorohod topology (see [20])). Furthermore, each process $Y\in D^2_{\mathcal{F}}([0,T],R^b)$ is square-integrable in the sense that

$$E\left[\int_0^T\|Y(t)\|^2dt\right]<\infty. \tag{2.6}$$

In the sequel, we use $D^2_{\mathcal{F},p}([0,T],R^b)$ to denote the corresponding predictable space. Similarly, we use $D^2_p([0,T]\times\mathcal{Z}^h,\ R^{l\times h})$ to denote the set of all $R^{l\times h}$-valued random processes $\tilde{V}(t,z)=(\tilde{V}_1(t,z_1),\ \ldots,\ \tilde{V}_h(t,z_h))$ that are predictable for each $z\in\mathcal{Z}^h$, and we endow this space with the norm

$$E\left[\sum_{i=1}^h\int_0^T\int_{\mathcal{Z}}\left\|\tilde{V}_i(t,z_i)\right\|^2\nu_i(dz_i)dt\right]<\infty. \tag{2.7}$$

In the deterministic case, we let

$$L^2_\nu(\mathcal{Z}^h,R^{q\times h})\equiv\left\{\tilde{v}:\mathcal{Z}^h\to R^{q\times h},\ \sum_{i=1}^h\int_{\mathcal{Z}}\|\tilde{v}_i(z_i)\|^2\nu_i(dz_i)<\infty\right\} \tag{2.8}$$

and we endow it with the norm

$$\|\tilde{v}\|^2_\nu\equiv\sum_{i=1}^h\int_{\mathcal{Z}}\|\tilde{v}_i(z_i)\|^2\lambda_i\nu_i(dz_i). \tag{2.9}$$

Finally, to suitably impose conditions on the initial and terminal values, we use $L^2_{\mathcal{G}_l}(\Omega,R^l)$ with $l\in\{p,q\}$ to denote the set of all R^l-valued, square-integrable, and $\mathcal{G}_l$-measurable random variables with $\mathcal{G}_p=\mathcal{G}$ and $\mathcal{G}_q=\mathcal{F}_T$.

We suppose the coefficients in (1.1)–(1.5) are $\{\mathcal{F}_t\}$-predictable and continuous in terms of W and L. More precisely, they can be presented in the forms of functionals as follows,

$$\begin{aligned}
b(t,x,u)&\equiv b(t,x,v,\bar{v},\tilde{v},u,*):[0,T]\times R^p\times R^q\times R^{q\times d}\times R^{q\times h}\times U\to R^p,\\
\sigma(t,x,u)&\equiv\sigma(t,x,v,\bar{v},\tilde{v},u,*):[0,T]\times R^p\times R^q\times R^{q\times d}\times R^{q\times h}\times U\to R^{p\times d},\\
\eta(t,x,u)&\equiv\eta(t,x,v,\bar{v},\tilde{v},u,z,*):[0,T]\times R^p\times R^q\times R^{q\times d}\times R^{q\times h}\times U\times\mathcal{Z}^h\to R^{p\times h},\\
c(t,x,u)&\equiv c(t,x,v,\bar{v},\tilde{v},u,*):[0,T]\times R^p\times R^q\times R^{q\times d}\times R^{q\times h}\times U\to R^q,\\
\alpha(t,x,u)&\equiv\sigma(t,x,v,\bar{v},\tilde{v},u,*):[0,T]\times R^p\times R^q\times R^{q\times d}\times R^{q\times h}\times U\to R^{q\times d},\\
\zeta(t,x,u)&\equiv\zeta(t,x,v,\bar{v},\tilde{v},u,z,*):[0,T]\times R^p\times R^q\times R^{q\times d}\times R^{q\times h}\times U\times\mathcal{Z}^h\to R^{q\times h}.
\end{aligned}$$

For each $f\in\{b,\sigma,c,\alpha\}$ and its corresponding values f^1,f^2 at $(x^1,\ v^1,\ \bar{v}^1,\ \tilde{v}^1,u)$ and $(x^2,\ v^2,\ \bar{v}^2,\ \tilde{v}^2,u)$, we assume that

$$\|f(u)\|\le\hat{L}(t,\omega)\left(1+\|x\|+\|v\|+\|\bar{v}\|+\|\tilde{v}(\cdot)\|\right), \tag{2.10}$$

$$\left\|f^2(u)-f^1(u)\right\|\le\hat{L}(t,\omega)\left(\left\|x^2-x^1\right\|+\left\|v^2-v^1\right\|+\left\|\bar{v}^2-\bar{v}^1\right\|+\left\|\tilde{v}^2(\cdot)-\tilde{v}^1(\cdot)\right\|\right). \tag{2.11}$$

Note that the conditions in (2.11) and the following (2.13) represent the Lipschitz condition under a given state-dependent feedback control rule u. Meanwhile, for each $f\in\{\eta,\zeta-\tilde{v}\}$ with $z\in\mathcal{Z}^h$ and and its corresponding values f^1,f^2 at $(x^1,\ v^1,\ \bar{v}^1,\ \tilde{v}^1,u)$ and $(x^2,\ v^2,\ \bar{v}^2,\ \tilde{v}^2,u)$, we suppose that

$$\sum_{i=1}^h\int_{\mathcal{Z}}\|f_i(u,z_i)\|^2\lambda_i\nu_i(dz_i)\le\hat{L}^2(t,\omega)\left(1+\|x\|^2+\|v\|^2+\|\bar{v}\|^2+\|\tilde{v}(\cdot)\|^2\right), \tag{2.12}$$

where f_i is the i-th column of f such that

$$\begin{aligned}
&\sum_{i=1}^h\int_{\mathcal{Z}}\left\|f_i^2(u,z_i)-f_i^1(u,z_i)\right\|^2\lambda_i\nu_i(dz_i)\\
&\le\hat{L}^2(t,\omega)\left(\left\|x^2-x^1\right\|^2+\left\|v^2-v^1\right\|^2+\left\|\bar{v}^2-\bar{v}^1\right\|^2+\left\|\tilde{v}^2(\cdot)-\tilde{v}^1(\cdot)\right\|^2\right).
\end{aligned} \tag{2.13}$$

We assume the known random factors in $H(X(T),\cdot)$ and all the coefficients $f\in\{b,\sigma,\eta,c,\alpha,\zeta\}$ are functionals of W and L that are continuous in W and L. Furthermore, we suppose that

$$\|H(x,*)\|\leq \hat{L}(\omega)\|x\|, \tag{2.14}$$

$$\left\|H(x^2,*)-H(x^1,*)\right\|\leq \hat{L}(\omega)\left\|x^2-x^1\right\|. \tag{2.15}$$

Note that $\hat{L}(\omega)$ in (2.14)–(2.15) is the same as $\hat{L}(T,\omega)$ in (2.10)–(2.13). Furthermore, $\hat{L}(t,\omega)$ is assumed to be a known non-negative, $\{\mathcal{F}_t\}$-adapted, and square-integrable random process, i.e.,

$$E\left[\int_0^T \hat{L}^2(t)dt\right]<\infty. \tag{2.16}$$

2.3. Main Theorem. Based on the assumptions and conditions presented in the previous subsections, we can state our main theorem as follows.

THEOREM 2.1. *Under the assumptions that*

1. *The coefficients in (1.1)–(1.5) are $\{\mathcal{F}_t\}$-predictable for each fixed $z\in\mathcal{Z}^h$ and any given $(u,v,\bar{v},\tilde{v})\in R^p\times R^q\times R^{q\times d}\times L^2_\nu(\mathcal{Z}^h,R^{q\times h})$;*
2. *$b(t,0,0,0,0,u)$, $\sigma(t,0,0,0,0,u)$, $\eta(t,0,0,0,0,u,z)$, $c(t,0,0,0,0,u)$, $\alpha(t,0,0,0,0,u)$, and $\zeta(t,0,0,0,0,u,z)$ are square-integrable;*
3. *Conditions (2.10)–(2.16) hold and the initial and terminal values satisfy*

$$(\xi,H)\in L^2_{\mathcal{G}}(\Omega,R^p)\times L^2_{\mathcal{F}_T}(\Omega,R^q), \tag{2.17}$$

then the following claims are true:

If S and R satisfy the completely-$\mathcal{S}$ condition, there exists a unique adapted 6-tuple weak solution $((X,Y),(V,\bar{V},\tilde{V},F))$ (in the sense of distributions) to system (1.1) under a given control rule u when at least one of the forward and backward SDEs has reflection boundary;

In particular, in the special cases in which all the square subprincipal matrices of $\bar{N}'S$, $(N'S)'$, $N'R$, and $(N'R)'$ are invertible, or if both of the SDEs have no reflection boundaries, there is a unique adapted 6-tuple strong solution $((X,Y)$, $(V,\bar{V},\tilde{V},F))$ (in the sample pathwise sense) to system (1.1) under a given control rule u. Furthermore, there exist two Lipschitz continuous mappings Φ and Ψ such that

$$\begin{cases} X(t)=\bar{Z}(t)+\Phi(\bar{Z})(t), & \Phi(\bar{Z})(t)=Y(t),\\ V(t)=\bar{U}(t)+\Psi(\bar{U})(t), & \Psi(\bar{U})(t)=F(T)-F(t),\end{cases} \tag{2.18}$$

where the processes $\bar{Z}$ and $\bar{U}$ in (2.18) are given by

$$\begin{cases} \bar{Z}(t) = \xi+Z(t),\\ \bar{U}(t) = H(X(T),*)+U(t).\end{cases} \tag{2.19}$$

The proof of Theorem 2.1 is given in Section 4. In the proof, the well-posedness for the system in a general weak sense is our focal point and it needs more involved work. In this case, due to the assumption of the general completely-$\mathcal{S}$ reflection condition, the solution to an embedded Skorohod problem may be not unique, i.e., bifurcations may occur at reflection boundaries. This phenomenon frequently appears in physical systems (e.g., queueing network systems) and the well-known contraction mapping approach can not be extended readily to solve our problem. Thus, a weak convergence method is developed to prove the associated result in Theorem 2.1. However, when the stricter invertibility condition on reflection matrices is imposed, the conventional contraction mapping approach can be applied. In this case, the reflection matrices can be extended to be time and/or queue-state dependent ones (see the related discussions in [36] and references therein). Furthermore, the uniqueness concerning an adapted 6-tuple strong solution $((X,Y),\,(V,\bar{V},\tilde{V},F))$ means that if there is another such solution, their difference must be zero under the product norm endowed to the space $\mathcal{Q}^2_{\mathcal{F}}([0,T])$ in (2.5). In other words, corresponding to the given forms of the state-dependent netput processes Z and U in (1.1), the related system decomposition in terms of X and V together with four regulating processes Y, $\bar{V}$, $\tilde{V}$, and F along each sample path is unique. Nevertheless, for consistency of the statements and notations in this paper, we keep the current constant reflection matrix assumption in the theorem.

Finally, before proving Theorem 2.1, we first discuss the applications of our main result in signal processing over MIMO and Possibly non-Gaussian communication channels.

3. Applications to signal processing over communication channels

In this section, we study the applications of the solution to the system of SDEs presented by (1.1)–(1.5) in signal processing over MIMO and possible non-Gaussian communication channels. E.g., we derive and prove a generalized formula of MIMO mutual information involving Lévy jumps, we discuss the well-posedness of queueing networks with applications in related economical modeling, and we determine the capacity regions of wireless MIMO channels with multiple users.

3.1. A Generalized formula of MIMO mutual information.

In this subsection we derive and prove a generalized formula of mutual information for the signal processing over an MIMO and possibly non-Gaussian Lévy communication channel based on the solution to the SDE in (1.1)–(1.5).

Mutual information is a basic concept in information theory (see [9], [33]) that originated from the well-known work of Bell Labs scientist Shannon in calculating the maximal transmission rate over a Gaussian communication channel for a single user [44]. Later on, this work was extended to multiple users and MIMO cases over Gaussian channels ([23]). More recently, to capture the jumps of data movements, this calculation was generalized the case of a non-Gaussian Lévy channel for a single user [19]. With applications to Big Data (see [18]) in mind as discussed in Section 1 of this paper, we aim to further derive a mutual information formula for multiple users and MIMO non-Gaussian Lévy channels.

More precisely, we consider a specific p-dimensional signal process S of the forward equation in (1.1)–(1.5) as the one to be transmitted over the channel, i.e.,

$$\begin{cases} dS(t) = \left(b_0(t^-)+b_1(t^-)S(t^-)+\sigma(t^-)\psi\int_{\mathcal{Z}^h}\mathrm{diag}(z)\mathrm{diag}(\lambda)\nu(dz)\right)dt \\ \qquad\qquad +\sigma(t^-)\left(\phi dW(t)+\psi\int_{\mathcal{Z}^h}\mathrm{diag}(z)\tilde{N}(\lambda dt,dz)\right), \\ S(0) \;=\; s, \end{cases} \tag{3.1}$$

where "diag(z)" in (3.1) denotes the diagonal matrix with each entry in the main diagonal given by the corresponding component of a given $z\in\mathcal{Z}^h$. The channel can be either linear or nonlinear with the corresponding received $\bar{p}$-dimensional signal process

Z given by

$$\begin{cases} dZ(t) &= \left(\bar{b}(t^-,S,Z)+\int_{\mathcal{Z}^{\bar{h}}}\bar{\psi}(t^-,S,Z,z)\mathrm{diag}(\bar{\lambda})\bar{\nu}(dz)\right)dt \\ &\quad +\bar{\phi}(t^-,S,Z)d\bar{W}(t)+\int_{\mathcal{Z}^{\bar{h}}}\bar{\psi}(t^-,S,Z,z)\tilde{\bar{N}}(\bar{\lambda}dt,dz), \\ Z(0) &= z. \end{cases} \tag{3.2}$$

In equations (3.1)–(3.2), W and $\bar{W}$ are two independent standard d-dimensional and $\bar{d}$-dimensional Brownian motions, $\tilde{N}$ and $\tilde{\bar{N}}$ are two independent centered h-dimensional and $\bar{h}$-dimensional Lévy processes with associated Lévy measures ν and $\bar{\nu}$ respectively. Furthermore, the coefficients σ and b_i with $i\in\{0,1\}$ in (3.1) are deterministic and square-integrable vector or matrix functions in t. Besides the linear growth and Lipschitz conditions for $\bar{b}$, $\bar{\phi}$, and $\bar{\psi}$ as required by Theorem 2.1, the matrices $\phi\phi'$ and $\psi\psi'$, the matrical function $\bar{\sigma}(t)\bar{\sigma}(t)'$, and the matrical processes $\bar{\phi}\bar{\phi}'(\cdot)$ and $\bar{\psi}\bar{\psi}'(\cdot)$ for all $t\in[0,T]$ are invertible.

We define the mutual information between the transmitted signal S and the received signal Z over time interval $[0,T]$ by

$$I(T,S,Z)=E\left[\ln\frac{dF_{(S,Z)}}{d(F_S\times F_Z)}(S,Z)\right], \tag{3.3}$$

where F_S, F_Z, and $F_{(S,Z)}$ denote the distributions of the processes S, Z, and (S,Z). Furthermore, the expression $dF_{(S,Z)}/d(F_S\times F_Z)$ in (3.3) is the Radon–Nikodym derivative (or density) of the joint distribution $F_{(S,Z)}$ with respect to the product distribution $F_S\times F_Z$. In other words, $I(t,S,Z)$ represents the information contained in the received signal Z about the transmitted signal S over time interval $[0,t]$ [9]. Furthermore, we define two processes by

$$\gamma(t,S,Z)\equiv\bar{b}(t,S,Z)'\left(\bar{\phi}(t,S,Z)\bar{\phi}(t,S,Z)'\right)^{-1}\bar{\phi}(t,S,Z), \tag{3.4}$$

$$\eta(t,S,Z,z)=(\eta_1(t,S,Z,z_1),\ldots,\eta_{\bar{h}}(t,S,Z,z_{\bar{h}}))' \tag{3.5}$$

with the convention that

$$\ln\eta(t,S,Z,z)\equiv(\ln\eta_1(t,S,Z,z_1),\ldots,\ln\eta_{\bar{h}}(t,S,Z,z_{\bar{h}}))'. \tag{3.6}$$

Note that for each $i\in\{1,\ldots,\bar{h}\}$ and $z\in\mathcal{Z}^{\bar{h}}$, each component of η is a positive rate process given by

$$\eta_i(t,S,Z,z_i)\equiv\sum_{j=1}^{\bar{p}}\bar{\psi}_{ij}(t,S,Z,z_i). \tag{3.7}$$

Then, we have the following proposition.

PROPOSITION 3.1. *Under the conditions in Theorem 2.1 and Equation (3.7) for the system presented by the SDEs in (3.1)–(3.2), the mutual information $I(T,S,Z)$ defined in (3.3) can be calculated by the formula*

$$\begin{aligned} E\Bigg[\int_0^T\Bigg\{&-\frac{1}{2}\Big(\left(\gamma(t^-,S,Z)\gamma(t^-,S,Z)'-\hat{\gamma}(t^-,S,Z)\hat{\gamma}(t^-,S,Z)'\right) \\ &+\int_{\mathcal{Z}^{\bar{h}}}(\eta(t,S,Z,z)-\hat{\eta}(t,S,Z,z))'\mathrm{diag}(z)\bar{\nu}(\bar{\lambda}dz)\Big) \\ &+\int_{\mathcal{Z}^{\bar{h}}}\Big(\ln\left((\eta(t,S,Z,z)+e)'\mathrm{diag}(z)\right)-\ln\left((\hat{\eta}(t,S,Z,z)+e)'\mathrm{diag}(z)\right)\Big)\bar{\nu}(\bar{\lambda}dz)\Bigg\}dt\Bigg], \end{aligned} \tag{3.8}$$

where,

$$\hat{\gamma}(t,S,Z)=\frac{E_S[M(t)\gamma(t,S,Z)]}{E_S[M(t)]}, \tag{3.9}$$

$$\hat{\eta}(t,S,Z,z)=\frac{E_S[M(t)\eta(t,S,Z,z)]}{E_S[M(t)]}, \tag{3.10}$$

and E_S is the expectation in terms of the measure dF_S. Furthermore, $M(t)$ is a stochastic exponential given by

$$\begin{aligned}&\exp\left\{-\int_0^t\left(\frac{1}{2}\gamma(s^-,S,Z)\gamma(s^-,S,Z)'+\int_{\mathcal{Z}^{\bar h}}\Big(\eta(s^-,S,Z,z)'\mathrm{diag}(z)\Big)\bar{\nu}(\bar{\lambda}dz)\right)ds\right.\\&\quad+\int_0^t\gamma(s^-,S,Z)d\bar{W}(s)\\&\quad\left.+\int_0^t\int_{\mathcal{Z}^{\bar h}}\ln\Big(\big(\eta(s^-,S,Z,z)+e\big)'\mathrm{diag}(z)\Big)\bar{N}(\bar{\lambda}ds,dz)\right\},\end{aligned} \tag{3.11}$$

where "e" denotes the $\bar{h}$-dimensional column vector of ones.

The proof of Proposition 3.1 is given in Section 4.

3.2. Queueing signal processes. In this subsection, we interpret the signal process S in (3.1) as the queue length processes which appears in many real-world network applications (see an example with p-users in Figure 1.2). The main performance measure for such a network is the queue length process denoted by $S(\cdot)=(S_1(\cdot),...,S_p(\cdot))'$, where $S_i(t)$ is the number of i-th class jobs stored in the i-th buffer for each $i\in\{1,...,p\}$ at time $t\in[0,\infty)$. Then, the queueing dynamics of the network can be modeled by

$$S(t)=S(0)+A(t)-D(t), \tag{3.12}$$

where the i-th component $A_i(t)$ of $A(t)$ for each $i\in\{1,...,p\}$ is the total number of jobs that arrive at buffer i by time t, and the i-th component $D_i(t)$ of $D(t)$ is the total number of jobs that depart from buffer i by time t. In the following studies, we use two generalized ways to characterize the arrival and departure processes.

3.2.1. Case I: Brownian networks with nominal balanced rates. In this Brownian network case, we assume that both the arrival and service processes are described by renewal processes or doubly stochastic renewal processes. In this case, the driven processes for the queueing system do not have the nice statistical properties such as memorylessness and stationarity of increments. Thus, it is usually impossible to conduct exact analysis concerning the distribution of $S(\cdot)$. However, under certain conditions (e.g., the arrival rates close to the associated service rates), one can show that the corresponding sequence of diffusion-scaled queue length processes converges in distribution to a p-dimensional reflecting Brownian motion (RBM) (see [12,13], [10], [11]), or more generally, to a reflecting diffusion with regime switching (RDRS) (see [15]). In other words, we have that

$$\hat{S}^r(\cdot)\equiv\frac{1}{r}S(r^2\cdot)\Rightarrow\hat{S}(\cdot)\ \text{ along } r\in\{1,2,...\}, \tag{3.13}$$

where "$\Rightarrow$" means "converges in distribution" and $\hat{S}(\cdot)$ is an RBM or an RDRS. For simplicity, we consider the case that the limit $\hat{S}(\cdot)$ in (3.13) is an RBM living in the state space D introduced in Section 2. Furthermore, let θ be a vector in R^p and Γ be a $p\times p$ symmetric and positive definite matrix. Then, we can introduce the definition of an RBM (see [12]) as follows.

DEFINITION 3.1. *A semimartingale RBM associated with the data* $(\mathbf{S},\theta,\Gamma,R)$ *that has initial distribution* π *is a continuous,* $\{\mathcal{F}_t\}$*-adapted, p-dimensional process* Z *defined on some filtered probability space* $(\Omega,\mathcal{F},\{\mathcal{F}_t\},\mathbf{P})$ *such that under* $\mathbf{P}$*,*

$$X(t)=Z(t)+RY(t)\ for\,all\,t\geq 0, \tag{3.14}$$

where,

(1) X *has continuous paths in* $\mathbf{S}$*,* $\mathbf{P}$*-a.s.,*
(2) under $\mathbf{P}$*,* Z *is a p-dimensional Brownian motion with drift vector* θ *and covariance matrix* Γ *such that* $\{Z(t)-\theta t,\mathcal{F}_t,t\geq 0\}$ *is a martingale and* $PZ^{-1}(0)=\pi$*,*
(3) Y *is a* $\{\mathcal{F}_t\}$*-adapted, b-dimensional process such that* $\mathbf{P}$*-a.s., for each* $i\in\{1,...,b\}$*, the i-th component* Y_i *of* Y *satisfies*
  (a) $Y_i(0)=0$*,*
  (b) Y_i *is continuous and non-decreasing,*
  (c) Y_i *can increase only when* Z *is on the face* D_i*, i.e., as given in system (1.1).*

COROLLARY 3.1. *If the reflection matrix* R *in (3.14) satisfies the conditions as required in Theorem 2.1, an RBM in Definition 3.1 is well-posed either in the weak sense or in the strong sense.*

The proof of Corollary 3.1 is a direct conclusion of Theorem 2.1. In the literature there are also some specific discussions about this particular type of SDEs. The proof for SRBMs in the weak sense over a general convex polyhedral can be found in [11] and the proof over a special positive orthant can be found in [47]. The proof in the strong sense over a special positive orthant can be found in [24].

3.2.2. Case II: Lévy networks with controllable rates. In this Lévy network case, we suppose that system arrival and service rates are designed in a dynamical controllable way. More precisely, we suppose that the job arrival rate $A_i(\cdot)$ to buffer i at time t, for $i\in\{1,...,p\}$, is a time-inhomogeneous Lévy process with intensity measure $a_i(t,S,z_i)dt\nu_i(dz_i)$ that depends on the queue state at time t. Analogously, we suppose that the assigned service rate $D_i(\cdot)$ to buffer i at time t is also a time-inhomogeneous Lévy process with intensity measure $d_i(t,S,z_i)dt\nu_i(dz_i)$. Furthermore, we assume that the routing proportion from buffer j to buffer i for jobs finishing service at buffer j is $p_{ji}(t,S,z_j)$. In addition, the service rate $d_i(t,S)$ will be set to zero when $S_i(t)=0$, i.e., system (3.15) is designed in a controllable manner. Thus, by the discussion in [2], the queue length process in (3.12) for this case can be further expressed by a forward SDE (a special form of the unified system in (1.1)), i.e.,

$$
\begin{aligned}
dS_i(t) = & \int_{\mathcal{Z}} \left(a_i(t,S,z_i) - d_i(t,S,z_i) I_{\{S_i(t)>0\}}\right) \nu_i(dz_i)dt \\
& + \sum_{j \neq i} \int_{\mathcal{Z}} p_{ji}(t,S,z_j) d_j(t,S,z_j) I_{\{S_j(t)>0\}} \nu_j(dz_j)dt \\
& + \int_{\mathcal{Z}} \left(a_i(t,S,z_i) - d_i(t,S,z_i) I_{\{S_i(t)>0\}}\right) \tilde{N}_i(dt,dz_i) \\
& + \sum_{j \neq i} \int_{\mathcal{Z}} p_{ji}(t,S,z_j) d_j(t,S,z_j) I_{\{S_j(t)>0\}} \tilde{N}_j(dt,dz_j) \\
& + \sum_{j=1}^{b} R_{ij}(t,S(t)) dY_j(t),
\end{aligned} \tag{3.15}
$$

where $\mathcal{Z} = R_+$. Furthermore, $Y_j(t)$ for each $j \in \{1,...,b\}$ in (3.15) is the Skorohod regulator process that can increase only at time t when $S_j(t) = 0$. Examples and specific formulations of system (3.15) can be found in [35], [36], and [31]. We have the following corollary.

COROLLARY 3.2. *If the reflection matrix in (3.15) is constant-valued (i.e., $R(t,S(t)) = R$) and if the conditions imposed in Theorem 3.1 hold for system (3.15) corresponding to Case II, then the system (3.15) is well-posed either in the weak sense or in the strong sense.*

The proof of Corollary 3.2 is a direct conclusion of Theorem 2.1. To be clear, we give some comments about the conditions on the reflection matrix in the corollary. As pointed out in the remark to Theorem 3.1 that the reflection matrix $R(t,S(t))$ may be time- and queue-state dependent (see the related discussions in [36] and references therein). Furthermore, the coefficients in (3.15) may be discontinuous at the queue state $S_i(t) = 0$. However, for our current purpose, and since the system in (3.15) is designed in a controllable manner, the routing probabilities in a queueing network can be designed in a stationary way and the service rate $d_i(s,S(s))$ can always be set to be zero when $S_i(t) = 0$. Thus, the constant reflection matrix can be assumed. Moreover, the generalized Lipschitz and linear growth conditions in (2.10)–(2.13) may also be imposed onto system (3.15).

3.2.3. Case III: modeling of network economics. Based on the queueing system in (3.12) (either an RBM in Definition 3.1 or a controllable form in (3.15)), we can formulate its corresponding economic model by an B-SDE with $q = p$. More precisely, in this model, each user $l \in \{1,...,p\}$ is contracted to use the system resource at each time t through a utility function $c_l(t,Q_l)$, e.g.,

$$
c_l(t,Q_l) = \begin{cases} 0 & \text{if } Q_l(t) = 0, \\ c_l Q_l(t) & \text{if } Q_l(t) \in (0,b_l), \\ c_l b_l & \text{if } Q_l(t) \geq b_l, \end{cases} \tag{3.16}
$$

where c_l is some nonnegative constant. Just as in financial option management [14,16], the system manager has an economic objective associated with a terminal condition H_l during a certain time period $[0,T]$ for a given $T \in [0,\infty)$, e.g.,

$$
H_l(Q_l(T)) = c_l Q_l(T). \tag{3.17}
$$

More exactly, the system manager wishes to maximize expected revenue while minimizing financial risk (i.e., by minimizing certain related variances). This can be done by suitably managing the value process together with controlling the arrival and service rates for each user $l\in\{1,...,p\}$, i.e., an B-SDE given by

$$\begin{aligned} V_l(t) = H_l(Q_l(T)) + \int_t^T c_l(s^-,Q_l)ds - \int_t^T \bar{V}_{l\cdot}(s^-)dW(s) \\ - \int_t^T \int_{\mathcal{Z}^h} \tilde{V}_{l\cdot}(s^-,z)\tilde{N}(ds,dz) + \Big(F_l(T) - F_l(t)\Big), \end{aligned} \tag{3.18}$$

where $\bar{V}_{l\cdot}$ and $\tilde{V}_{l\cdot}$ are, respectively, the lth rows of $\bar{V}$ and $\tilde{V}$. Furthermore, $F_l(t)$ can increase only at a time t when $V_l(t) = \bar{b}_l$. The constant $\bar{b}_l$ can be determined according to the mean and variance of the terminal target $H_l(Q_l(T))$. The process $F_l(t)$ is the payback process to the user $l\in\{1,...,p\}$, which looks like the early exercise reward process in a conventional finance system [29].

COROLLARY 3.3. *Under the conditions imposed in Corollary 3.2, the coupled system of FB-SDEs in (3.15) and (3.18) is well-posed either in a weak sense or in a strong sense.*

The proof of Corollary 3.3 is a direct conclusion of Theorem 2.1. It is worth pointing out that a system of FB-SDEs with Lévy jumps and/or reflections has the potential to be used in finance engineering and game theory. Interested readers can find such examples in [16] where an B-SDE with Lévy jumps is considered, and in [29] where an B-SDE driven by a standard Brownian motion with reflection is considered.

3.3. Multi-Users' MIMO channel capacity region. In this subsection, we illustrate how to use the mutual information derived in Proposition 3.1 and the RBM defined in Definition 3.1 or system (3.15) to determine the capacity region and schedule the capacity for wireless MIMO channels with multi-users.

More precisely, we are concerned with a channel (or channels) that can be considered as a base station having multi-antennas and p-users (p-mobiles). Each user corresponds to a queue storage buffer and may be also equipped with multi-antennas. The antennas in both base station and user's mobiles have the ability to cooperate in the sense that they can perform joint beam-forming and/or power control. Hence, the p-users can be served at the same time by a single server (called a p-parallel server) with rate allocation vector $c(t) = (c_1(t),...,c_p(t))'$ that takes values in a capacity region $\mathcal{R}$ at a time t. However, in doing so, there is a constraint on the total power that the antennas in the base station can share or a constraint on the power to each individual user.

Determining the capacity region $\mathcal{R}$ that is achievable under a coding technique is an important research area in both academia and communication practices [23], [27].

Along this direction, the most popular convention imposed is the Gaussian channel assumption. In this situation, the input signal in (3.1) (e.g., the queue length process in (3.12), (3.14), and (3.15)) reduces to a Brownian motion. Furthermore, the received signal Z in (3.2) also reduces to a Brownian motion driven process. Then, a convex capacity region that consists of the origin and L $(>p)$ boundary pieces can be derived for a constant channel (e.g., a p-user MIMO multiple access channel (MAC) or a broadcast channel (BC) in real-world wireless communication systems). In Figure 3.1, we display such an example of the capacity region for the purpose of illustration.

Note that by combining the mutual information in Proposition 3.1 and the studies in Goldsmith *et al.* [23], Jindal *et al.* [27], it is possible for one to derive the capacity

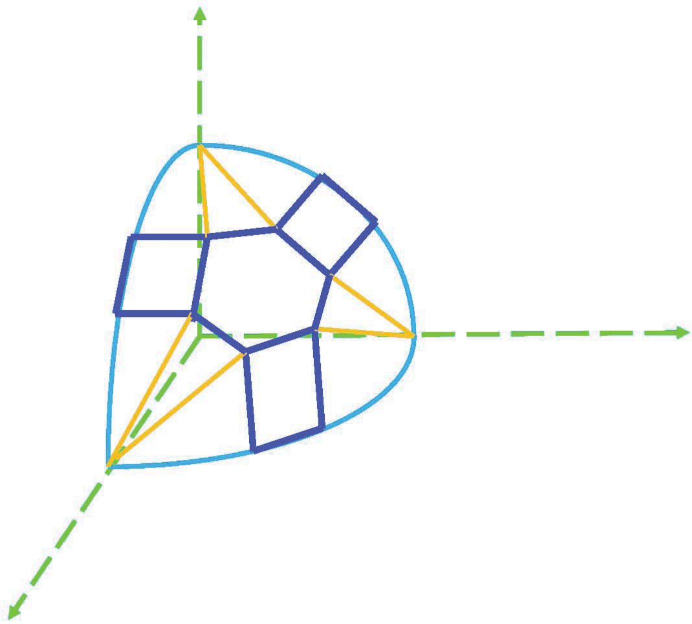

FIG. 3.1. *A capacity region for an MIMO channel.*

region for a non-Gaussian MIMO channel with multiple users and power constraint. In addition, one can also develop some techniques to handle the gap between the achievable capacity derived theoretically and the required one corresponding to real-world channels. For example, the queue based rate scheduling policy corresponding to specific service rate controls in (3.15) is designed and justified in Dai [15] to optimize the utility and performance of the achievable maximal capacity for MIMO channels with multiple users.

4. Proofs of Theorem 2.1 and Proposition 3.1

To provide the proofs for Theorem 2.1 and Proposition 3.1, we first recall the Skorohod problem with jumps and study its properties.

4.1. The Skorohod problem. Let $D([0,T],R^b)$ with $b\in\{1,2,...\}$ be the space of all functions $z:[0,T]\to R^b$ that are right-continuous with left limits and are endowed with the Skorohod topology [5], [26]. Then, we have the Skorohod problem as follows.

DEFINITION 4.1 (The Skorohod problem). *Given $z\in D([0,T],R^p)$ with $z(0)\in D$, a (D,R)-regulation of z over [0,T] is a pair $(x,y)\in D([0,T],D)\times D([0,T],R_+^b)$ such that*

$$x(t)=z(t)+Ry(t) \text{ for all } t\in[0,T],$$

where for each $i\in\{1,...,b\}$,

(1) $y_i(0)=0$,

(2) y_i is nondecreasing,

(3) y_i can increase only at a time $t\in[0,T]$ with $x(t)\in F_i$.

DEFINITION 4.2 (Maximal Set). *A set $K\subseteq\{1,...,b\}$ is called "maximal" if $K\neq\emptyset$, $D_K\neq\emptyset$, and $D_K\neq D_{\tilde{K}}$ for any $\tilde{K}\supset K$ such that $\tilde{K}\neq K$, where $D_\emptyset=D$ and*

$$D_K\equiv\cap_{i\in K}D_i. \tag{4.1}$$

Furthermore, we define the modulus of continuity with respect to a function $z(\cdot)\in D([0,T],R^b)$ and a real number $\delta>0$ by

$$w(z,\delta,T)\equiv\inf_{t_l}\max_{l}\mathrm{Osc}\left(z,[t_{l-1},t_l)\right), \tag{4.2}$$

where the infimum takes over all the finite sets $\{t_l\}$ of points satisfying $0=t_0<t_1<...<t_m=T$ and $t_l-t_{l-1}>\delta$ for $l=1,...,m$, and

$$\mathrm{Osc}(z,[t_{l-1},t_l])=\sup_{t_1\le s\le t\le t_2}\|z(t)-z(s)\|. \tag{4.3}$$

Then, we have the following lemma.

LEMMA 4.1. *Suppose that the reflection matrix R in Definition 4.1 satisfies the completely-$\mathcal{S}$ condition. Then, any (D,R)-regulation (x,y) of $z\in D([0,T],R^p)$ with $z(0)\in D$ satisfies the oscillation inequality over $[t_1,t_2]$ with $t_1,t_2\in[0,T]$*

$$Osc(x,[t_1,t_2])\le\kappa Osc(z,[t_1,t_2]), \tag{4.4}$$
$$Osc(y,[t_1,t_2])\le\kappa Osc(z,[t_1,t_2]), \tag{4.5}$$

where κ is some nonnegative constant depending only on the inward normal vector N and the reflection matrix R.

Note that the oscillation inequalities (4.4)–(4.5) for continuous paths can be found in [3], [11]. The first such inequalities for discontinuous paths with jumps can be found in [12] where the quantities in the right-hand sides of (4.4)–(4.5) are subject to an additional constraint of bounded jump sizes (see also the related discussions in [10] and [48]). Nevertheless, for the purpose of this paper, we remove such an additional constraint and directly extend the inequalities in (4.4)–(4.5) to a general discontinuous case with jumps.

Proof. **(Proof of Lemma 4.1.)** First, for a set $K\subseteq\{1,2,...,b\}$, let $d(x,D_K)$ denote the Euclidean distance between x and D_K for a point $x\in D$. Then, it follows from Lemma 3.2 in [12] or Lemma B.1 in [11] that there exist two constants $C\ge 1$ such that

$$d(x,D_K)\le C\sum_{i\in K}(n_i\cdot x-b_i). \tag{4.6}$$

Now, for each $\epsilon\ge 0$, and $K\subseteq\{1,...,b\}$ (including the empty set), we let

$$\begin{aligned} D_K^\epsilon\equiv\{x\in R^q:0\le n_i\cdot x-b_i\le C_\epsilon \text{ for all } i\in K,\\ n_i\cdot x-b_i>\epsilon \text{ for all } i\in\{1,...,b\}\setminus K\}, \end{aligned} \tag{4.7}$$

where $C_\epsilon=Cp\epsilon$. Thus, by Lemmas 4.1–4.2 in [11], we know that

$$D=\cup_{K\in\mathcal{G}}D_K^\epsilon, \tag{4.8}$$

where $\mathcal{G}$ is the collection of subsets of $\{1,...,b\}$ consisting of all maximal sets in $\{1,...,b\}$.

Second, let $(N'R)_K$ be the square matrix corresponding to a maximal set K and consider a $(D_K^\epsilon,(N'R)_K)$-regulation problem. Without loss of generality, we assume that D_K^ϵ is the p-dimensional positive orthant. Then, for each $t\in[t_1,t_2]$, we define

$$\Delta z(t)\equiv z(t)-z(t^-), \tag{4.9}$$
$$\Delta x(t)\equiv x(t)-x(t^-), \tag{4.10}$$
$$\Delta y(t)\equiv y(t)-y(t^-). \tag{4.11}$$

Since the reflection matrix $(N'R)_K = R$ satisfies the completely-$\mathcal{S}$ condition, it is easy to check that the linear complementarity problem (LCP)

$$
\begin{aligned}
\Delta x(t) &= \Delta z(t) + R\Delta y(t),\\
\Delta x(t) &\in D,\\
\Delta y(t) &\geq 0,\\
\Delta x_i(t)\Delta y_i(t) &= 0 \text{ for } i=1,\dots,p,\\
(b_i - \Delta x_i(t))\Delta y_i(t) &= 0 \text{ for } i=p+1,\dots,b,
\end{aligned}
$$

is completely solvable (see also Theorem 2.1 in [34] for a related discussion). Furthermore, we can conclude that

$$\Delta y(t) \leq C\Delta z(t) \tag{4.12}$$

for some nonnegative constant C depending only on the inward normal vector N and the reflection matrix R.

Third, the rest of the proof is the direct conclusion of the proof for Theorem 3.1 in [12] or the proof for Theorem 4.2 in [10]. □

LEMMA 4.2. *For any maximal set $K \subseteq \{1,\dots,b\}$, if $(N'R)_K$ and $(N'R)'_K$ are invertible, the Skorohod problem in Definition 4.1 is well-posed. Furthermore, there is a Lipschitz continuous mapping Φ such that*

$$x = z + R\Phi(z), \;\; y = \Phi(z). \tag{4.13}$$

Proof. Consider a maximal set $K \subseteq \{1,\dots,b\}$ and its corresponding $(D_K^\epsilon, (N'R)_K)$-regulation problem. For convenience, we let y_K denote the $|K|$-dimensional vector whose components corresponding to the indices in K and y_K^c denote the $(b-|K|)$-dimensional complement of y_K in y. Then, if y_K^c does not increase over $[t_1, t_2)$, it follows from the proof of Part (a) for Theorem 3.1 in [12] or Theorem 4.2 in [10] that

$$x(t+t_1) = x(t_1) + (z(t+t_1) - z(t_1)) + (N'R)_K (y_K(t+t_1) - y_K(t_1)). \tag{4.14}$$

Since the matrix $(N'R)_K$ is invertible, it follows from Theorem 7.2 in [6] that the corresponding $(D_K^\epsilon, (N'R)_K)$-regulation problem is well-posed. Then, by the same method as was used to prove Theorem 3.1 in [12] or Theorem 4.2 in [10], we can finish the rest proof of Lemma 4.2. □

LEMMA 4.3. *Assume that $(x^n, y^n) \to (x, y)$ along $n \in \{1,2,\dots\}$ in $D([0,T], R^p) \times D([0,T], R^b)$ and $y^n(\cdot)$ is of bounded variation for each $n \in \{1,2,\dots\}$. Furthermore, suppose that*

$$\int_0^t f(x^n(s))dy^n(s) = 0 \tag{4.15}$$

for all $n \in \{1,2,\dots\}$ and each $t \in [0,T]$, where $f \in C^b([0,T], R^b)$ is a b-dimensional bounded vector function. Then, for each $t \in [0,T]$, we have that

$$\int_0^t f(x(s))dy(s) = 0. \tag{4.16}$$

Proof. It follows from the discussion in pages 123-124 of [5] or Theorem 1.14 of [26] that there is a sequence $\{\gamma_n, n\in\{1,2,...\}\}$ of continuous and strictly increasing functions mapping from $[0,T]\to[0,T]$ with $\gamma_n(0)=0$ and $\gamma_n(T)=T$ such that

$$\sup_{t\in[0,T]}|\gamma_n(t)-t|\to 0, \tag{4.17}$$

$$\sup_{t\in[0,T]}|(x^n,y^n)(\gamma_n(t))-(x,y)(t)|\to 0. \tag{4.18}$$

Then, by the uniform convergence of (4.17)–(4.18) and condition (4.15), we know that

$$\begin{aligned}\int_0^t f(x(s))dy(s) &= \lim_{n\to\infty}\int_0^t f(x^n(\gamma_n(s)))dy^n(\gamma_n(s))\\ &= \lim_{n\to\infty}\int_0^{\gamma_n^{-1}(t)} f(x^n(u))dy^n(u)\\ &= 0,\end{aligned}$$

where $\gamma_n^{-1}(\cdot)$ is the inverse function of $\gamma_n(\cdot)$ for each $n\in\{1,2,...\}$. This completes the proof of Lemma 4.3. □

4.2. Malliavin calculus of vector-valued Lévy processes. In this subsection, we study the Malliavin calculus for vector-valued Lévy processes by generalizing the discussions for single-dimensional Lévy processes in the existing literature (see [17], [40], [45], and [32]).

First, let $m=\max\{d,h\}$ and define a vector-valued measure $\mu(t,z)=(\mu_1(t,z_1),...,\mu_m(t,z_m))'$ in terms of $(t,z)\in R_+\times\mathcal{Z}^m$, where $\mu_j(t,z_j)$ for each $j\in\{1,...,m\}$ is given by

$$\mu_j(dt,dz_j)\equiv I_{\{j\le d\}}dtd\delta_0(z_j)+I_{\{j\le h\}}\lambda_j z_j^2 dt\nu_j(dz_j). \tag{4.19}$$

Then, for each Borel set $E=E_0\times E_1\times\cdots\times E_m\in\mathcal{B}(R_+\times\mathcal{Z}^m)$ such that $\mu_j(E_j)<\infty$ for each $j\in\{1,...,m\}$, we define a vector-valued random measure $M(E)=(M_1(E_0\times E_1),...,M_m(E_0\times E_m))'$, where M_j is given by

$$\begin{aligned}M_j(E_0\times E_j)\equiv{}& I_{\{j\le d\}}\int_{E_j(0)}dW_j(t)\\ &+I_{\{j\le h\}}\lim_{n\to\infty}\int\int_{\{(s,z_j)\in E_j':(1/n)\le z_j\le n\}} z_j\tilde{N}_j(\lambda_j ds,dz_j),\end{aligned} \tag{4.20}$$

and M is centered and independently scattered. Note that in (4.20), $E_j(0)=\{s\ge 0:(s,0)\in E_0\times E_j\}$, $E_j'=E_0\times E_j-\{(s,0)\in E_0\times E_j\}$, and the limit is taken in $L^2(\Omega)$. Furthermore, μ is the control measure of M since $E[M(A)\cdot M(B)]=\mu(A\cap B)$ for any $A,B\in\mathcal{B}(R_+\times\mathcal{Z}^m)$, where the dot "·" between $M(A)$ and $M(B)$ is taken in the Euclidean sense. In the sequel, we will consider the restriction of M to $[0,T]\times\mathcal{Z}^m$ (with the same notation).

Second, let $H=L^2([0,T]\times\mathcal{Z}^m,R^m,\mu)$ denote the space of all real vector-valued functions $h(t,z)=(h_1(t,z_1),...,h_m(t,z_m))'$ with the norm

$$\|h\|_H^2=\sum_{j=1}^m\int_{[0,T]\times\mathcal{Z}} h_j^2(t,z_j)\mu_j(dt,dz_j). \tag{4.21}$$

For each $h_n \equiv (h,...,h) \in H^n \equiv H \times H \times \cdots \times H$, we let $h^2 = (h_1^2,...,h_m^2)'$ and define

$$\|h\|_{H^n}^2 = \int_{([0,T]\times\mathcal{Z}^m)^n} \prod_{k=1}^{n} \langle h^2(t^k,z^k), \mu(dt^k,dz^k)\rangle. \tag{4.22}$$

In addition, let M_n be a vector-valued multiple $(m+1)$-parameter integral in terms of the random measure M defined by

$$\begin{aligned} M_n(h_n) &\equiv \int_{([0,T]\times\mathcal{Z}^m)^n} \prod_{k=1}^{n} \langle h(t^k,z^k), M(dt^k,dz^k)\rangle \\ &= \int_{[0,T]^n} \prod_{k=1}^{n} \langle h(t^k,0), dW(t^k)\rangle \\ &\quad + \int_{((0,T]\times\mathcal{Z}^m)^n} \prod_{k=1}^{n} \left\langle h(t^k,z^k), \tilde{N}(\lambda dt^k,dz^k)\right\rangle, \end{aligned} \tag{4.23}$$

where W and $\tilde{N}$ are the corresponding m-dimensional standard Brownian motion and compensated Poisson random measure, respectively.

Third, let $\mathcal{D}_{1,2}$ be the set of all random vectors $F \in L^2_{\mathcal{F}_T}(\Omega, R^q)$ with chaos expansion

$$F_r = \sum_{n=0}^{\infty} M_n(h_{n,r}) \text{ and norm } \|F_r\|^2_{\mathcal{D}_{1,2}} = \sum_{n=0}^{\infty} nn!\|h_{n,r}\|^2_{H^n} \tag{4.24}$$

for some $h_{n,r} \in H^n$ and each $r \in \{1,...,q\}$ such that

$$\|F\|^2_{\mathcal{D}_{1,2}} \equiv \max_{r\in\{1,...,q\}} \|F_r\|^2_{\mathcal{D}_{1,2}} < \infty. \tag{4.25}$$

Then we know that $\mathcal{D}_{1,2}$ is strictly contained in $L^2_{\mathcal{F}_T}(\Omega, R^q)$ when it is endowed with the following norm. For each $F \in L^2_{\mathcal{F}_T}(\Omega, R^q)$,

$$\|F\|^2_{L^2(\Omega,P)} \equiv \max_{r\in\{1,...,q\}} \|F_r\|^2_{L^2(\Omega,P)} < \infty, \;\; \|F_r\|^2_{L^2(\Omega,P)} = \sum_{n=0}^{\infty} n!\|h_{n,r}\|^2_{H^n}. \tag{4.26}$$

DEFINITION 4.3. *For each $F \in \mathcal{D}_{1,2}$, its Malliavin derivative $\mathcal{D}F$ is defined by the matrix*

$$\mathcal{D}_{t,z}F \equiv (\mathcal{D}_{t,z}F_1,...,\mathcal{D}_{t,z}F_q)' \tag{4.27}$$

with $(t,z) \in [0,T]\times\mathcal{Z}^m$, where

$$\mathcal{D}_{t,z}F_r = (D_{t,z_1}F_r,...,D_{t,z_m}F_r), \;\; r \in \{1,...,q\}, \tag{4.28}$$

$$\mathcal{D}_{t,z_j}F_r = \sum_{n=1}^{\infty} nM_{n-1}(h_{n,r}(\cdot,t,z_j)), \;\; j \in \{1,...,m\}. \tag{4.29}$$

In (4.29), $M_{n-1}(h_{n,r}(\cdot,t,z_j))$ means that the $(n-1)$-fold iterated integral of $h_{n,r}$ is regarded as a function of its $(n-1)$ first pairs of variables $(t^1,z^1),...,(t^{n-1},z^{n-1})$ while the final pair (t,z_j) is kept as a parameter.

LEMMA 4.4. *If a random vector $F\in\mathcal{D}_{1,2}$, we can conclude that*

$$\mathcal{D}F\in L^2_{\mathcal{F}_T}(\mu\times P)=L^2_{\mathcal{F}_T}(\Omega\times[0,T]\times\mathcal{Z}^m,R^{q\times m}).$$

Proof. For an $F\in\mathcal{D}_{1,2}$, the claim in the lemma follows from the following calculation,

$$\begin{aligned}\|\mathcal{D}F\|^2_{L^2_{\mathcal{F}_T}(\mu\times P)}&=\max_{r\in\{1,\dots,q\}}\left(\sum_{j=1}^{m}\int_{[0,T]\times\mathcal{Z}}E\left[\left(\mathcal{D}_{t,z_j}F_r\right)^2\right]\mu_j(dt,dz_j)\right)\\&=\max_{r\in\{1,\dots,q\}}\left(\sum_{j=1}^{m}\left(\int_{[0,T]\times\mathcal{Z}}\sum_{n=0}^{\infty}n^2(n-1)!\,\|h_r\|^2_{H^{n-1}}\mu_j(dt,dz_j)\right)\right)\\&=\max_{r\in\{1,\dots,q\}}\left(\sum_{n=0}^{\infty}nn!\,\|h_r\|^2_{H^n}\right)\\&=\|F\|^2_{\mathcal{D}_{1,2}}\\&<\infty,\end{aligned}\tag{4.30}$$

where the second equality in (4.30) follows from Definition 4.3. This completes the proof of the lemma. □

LEMMA 4.5. *Let $L^2(\Omega,R^q)$ be the space of square-integrable R^q-valued random vectors and $L^2(\Omega,H^q)$ be the space of H^q-valued processes, which is endowed with the norm (4.25). Then, the unbounded operator in Definition 4.3 is closable from $L^2(\Omega,R^q)$ to $L^2(\Omega,H^q)$.*

Proof. Let $\{F^i: i\in\{1,2,\dots\}\}$ be a sequence of smooth random vectors with chaos expansions, which converges to zero along $i\in\{1,2,\dots\}$ in $L^2(\Omega,R^q)$. Then, we have that

$$F^i_r\to 0 \text{ along } i\in\{1,2,\dots\} \text{ in } L^2(\Omega,P) \text{ for each } r\in\{1,\dots,q\}.\tag{4.31}$$

We suppose that the corresponding sequence related to the Malliavin derivatives converges to some η in $L^2(\Omega,H^q)$. Then, by (4.31), we know that

$$\sum_{j=1}^{m}\int_{[0,T]\times\mathcal{Z}}E\left[\left(\mathcal{D}_{t,z_j}F^i_r\right)^2\right]\mu_j(dt,dz_j)=\sum_{n=0}^{\infty}nn!\left\|h^i_r\right\|^2_{H^n}\to 0\tag{4.32}$$

along $i\in\{1,2,\dots\}$ for each $r\in\{1,\dots,q\}$, which implies that $\eta_r=0$ and hence $\eta=0$. Therefore, by the definition of the closable operator [49], we conclude that the claim in the lemma is true. □

In the sequel, we will use $\mathcal{D}^{\infty}_{1,2}$ to denote the domain of the unbounded operator

$$\mathcal{D}\colon L^2(\Omega,R^q)\to L^2(\Omega,H^q).\tag{4.33}$$

Owing to Lemma 4.5, this domain is the closure of the class of smooth random variables $\mathcal{D}_{1,2}$ with the norm (4.25). Furthermore, we will use $L^2_{1,2}(\Omega,R^q)$ to denote the space of product measurable and $\mathcal{F}_t$-adapted processes, which is endowed with the norm,

$$\|F\|^2_{1,2}=\|F\|^2_{\mathcal{D}_{1,2}}+\|\mathcal{D}F\|^2_{L^2(\Omega,H^q)}.\tag{4.34}$$

LEMMA 4.6. *For a matrix-valued process $Z\in\mathcal{D}^2_{\mathcal{F},p}([t,T]\times\mathcal{Z}^m,R^{q\times m})$ with each $t\in[0,T]$ and corresponding space H over $[t,T]$, we have*

$$Z\in L^2_{1,2}(\Omega,H^q) \text{ if and only if } F\equiv\int_{[t,T]\times\mathcal{Z}^m} Z(s,z)M(ds,dz)\in\mathcal{D}^{\infty}_{1,2}. \tag{4.35}$$

Furthermore, for each $\theta\in[0,T]$,

$$\mathcal{D}_{\theta,z}F(t)=\begin{cases}\int_{[t,T]\times\mathcal{Z}^m}\mathcal{D}_{\theta,z}Z(s,y)\tilde{M}(ds,dy) & \text{if } \theta\le t,\\ Z_r'(\theta,z)+\int_{[t,T]\times\mathcal{Z}^m}\mathcal{D}_{\theta,z}Z(s,y)\tilde{M}(ds,dy) & \text{if } \theta>t,\end{cases} \tag{4.36}$$

where $\tilde{M}$ is a $mm\times m$ matrix measure given by $\tilde{M}=(M',...,M')'$.

Proof. Without loss of the generality, we let $t=0$ and assume that the (rj)-th entry of Z for each $r\in\{1,...,q\}$ and $j\in\{1,...,m\}$ has the chaos expansion

$$Z_{rj}(s,z_j)=M_n(h_{n,rj}(\cdot,s,z_j)), \tag{4.37}$$

where $h_{n,rj}(t^1,z^1,...,t^n,z^n,s,z_j)\in H^n$ is defined as in (4.24). Since $Z\in\mathcal{D}^2_{\mathcal{F},p}([0,T]\times\mathcal{Z}^m,R^{q\times m})$, the expression in (4.37) is unique for each $r\in\{1,...,q\}$ and $j\in\{1,...,m\}$. Then, the remaining proof of the lemma reduces to the one for the single-dimensional Lévy processes, see the proof of Lemma 3.3 in [17] and the associated proof of Theorem 6.1 of [45]. This completes the proof of the lemma. □

LEMMA 4.7. *Let $F\in\mathcal{D}_{1,2}$ and ϕ be a real continuous vector function on R^q. If $\phi(F)\in L^2(\Omega,R)$ and $\phi(F+(\mathcal{D}_{t,z}F)e)\in L^2(\Omega\times[0,T]\times\mathcal{Z}^m,R)$, then we have $\phi(F)\in\mathcal{D}_{1,2}$, which satisfies*

$$\mathcal{D}_{t,z}\phi(F)=\phi(F+(\mathcal{D}_{t,z}F)e)-\phi(F), \tag{4.38}$$

where "e" is the m-dimensional column vector of ones.

Proof. First, for given $h^1,...,h^q\in H$, we define each component of their corresponding exponential random vector $R(T)=(R_1(T),...,R_q(T))'$ by

$$R_r(T)=\exp\left\{\int_{[0,T]\times\mathcal{Z}^m}\left\langle h^r(s,z),M(ds,dz)\right\rangle\right\} \text{ for each } r\in\{1,...,q\}. \tag{4.39}$$

Then, by replacing the terminal time T by a time $t\in[0,T]$, it follows from Itô's formula [38] that

$$dR_r(t)=R_r(t^-)\int_{\mathcal{Z}^m}\left\langle e^r(t,z),M(dt,dz)\right\rangle, \tag{4.40}$$

where $e^r(t,z)$ is a row vector function given by

$$e^r(t,z)=\left(e^{h^r(t,z_1)}-1,...,e^{h^r(t,z_m)}-1\right).$$

Now, for any $n\in\{1,2,...\}$ and $(t^1,...,t^n)\in[0,T]^n$, let e^r_n be the function corresponding to $e^r(t,z)$ as in (4.24). Then, we can extend the discussion given for Example 12.5 of [40] to our vector case with the combination of Gaussian and Pure Jump Lévy noises. More precisely, for $t^1 \leq \ldots \leq t^n$,

$$\begin{aligned}
R_r(T) &= 1 + \int_{[0,T]\times\mathcal{Z}^m} R_r(t^{1,-})\Big\langle e^r(t^1,z^1), M(dt^1,dz^1)\Big\rangle \\
&= 1 + \int_{[0,T]\times\mathcal{Z}^m} \Big\langle e^r(t^{1,-},z^1), M(dt^1,dz^1)\Big\rangle \\
&\quad + \int_{[0,T]\times\mathcal{Z}^m}\int_{[0,t^1]\times\mathcal{Z}^m} R_r(t^{2,-})\Big\langle e^r(t^1,z^1), M(dt^1,dz^1)\Big\rangle\Big\langle e^r(t^2,z^2), M(dt^2,dz^2)\Big\rangle \\
&\quad . \\
&\quad . \\
&\quad . \\
&= \sum_{k=0}^{n-1}\frac{1}{k!}M_k(e_k^r) + \int_{[0,T]\times\mathcal{Z}^m}\int_{[0,t^1]\times\mathcal{Z}^m}\cdots\int_{[0,t^n]\times\mathcal{Z}^m} R_r(t^{n,-}) \\
&\qquad\qquad \Big\langle e^r(t^1,z^1), M(dt^1,dz^1)\Big\rangle\cdots\Big\langle e^r(t^n,z^n), M(dt^n,dz^n)\Big\rangle \\
&\quad . \\
&\quad . \\
&\quad . \\
&= \sum_{n=0}^{\infty}\frac{1}{n!}M_n(e_n^r),
\end{aligned} \tag{4.41}$$

where the third equality follows from the fact that the set $S_n = \{(t^1,\ldots,t^n)\in[0,T]^n : 0 \leq t^1 \leq \ldots \leq t^n \leq T\}$ occupies the fraction $1/n!$ of the whole n-dimensional box $[0,T]^n$, and the last equality follows from the mean-square convergence.

Next, let $F_r = R_r(T)$ for each $r \in \{1,\ldots,q\}$. Then, it follows from (4.24), (4.41), and Definition 4.3 that

$$\begin{aligned}
\mathcal{D}_{t,z}F_r &= \sum_{n=1}^{\infty}\frac{n}{n!}M_{n-1}(e_n^r(\cdot,t,z)) \\
&= e^r(t,z)\sum_{n=1}^{\infty}\frac{1}{(n-1)!}M_{n-1}(e_{n-1}^r) \\
&= e^r(t,z)F_r.
\end{aligned} \tag{4.42}$$

Thus, we can conclude that

$$\mathcal{D}_{t,z}(y_1F_1+\ldots+y_qF_q) = e(t,y,z)(y_1F_1+\ldots+y_qF_q), \tag{4.43}$$

for any given real number point $y \in R^q$, where

$$e(t,y,z) = \left(e^{y_1h^1(t,z_1)+\ldots+y_qh^q(t,z_1)}-1,\ldots,e^{y_1h^1(t,z_m)+\ldots+y_qh^q(t,z_m)}-1\right).$$

Furthermore, let $\mathcal{D}_{1,2}^{\mathcal{E}}$ be the set of linear combinations of all exponential random vectors $F = R(T)$ as in (4.39). Then, for each $F \in \mathcal{D}_{1,2}^{\mathcal{E}}$ and a given $y \in R^q$, it follows from induction with respect to $n \in \{1,2,\ldots\}$ that

$$\begin{aligned}
\mathcal{D}_{t,z}((y_1F_1+\ldots+y_qF_q)^n) &= \Big((y_1F_1+\ldots+y_qF_q) + (\mathcal{D}_{t,z}(y_1F_1+\ldots+y_qF_q))e\Big)^n \\
&\quad - \Big(y_1F_1+\ldots+y_qF_q\Big)^n.
\end{aligned} \tag{4.44}$$

For example, when $n=2$, it follows from (4.42) that

$$\begin{aligned}\mathcal{D}_{t,z}((y_1F_1+...+y_qF_q)^2) &= e^2(t,y,z)(y_1F_1+...+y_qF_q)^2\\ &= \Big((y_1F_1+...+y_qF_q)+(\mathcal{D}_{t,z}(y_1F_1+...+y_qF_q))e\Big)^2\\ &\quad -\Big(y_1F_1+...+y_qF_q\Big)^2,\end{aligned}$$

where

$$e^2(t,y,z)=\Big(e^{2(y_1h^1(t,z_1)+...+y_qh^q(t,z_1))}-1,...,e^{2(y_1h^1(t,z_m)+...+y_qh^q(t,z_m))}-1\Big).$$

Now, assuming that ϕ has compact support and $F\in\mathcal{D}^{\mathcal{E}}_{1,2}$, then

$$\phi(F)=\left(\frac{1}{\sqrt{2\pi}}\right)^{-q}\int_{y\in R^q}e^{i(y_1F_1+...+y_qF_q)}\hat{\phi}(y)dy,$$
$$\hat{\phi}(y)=\left(\frac{1}{\sqrt{2\pi}}\right)^{-q}\int_{x\in R^q}e^{-i\langle x,y\rangle}\phi(x)dx,$$

where $\hat{\phi}$ is the Fourier transform of ϕ. Furthermore, it follows from (4.44) and Lemma 4.5 that

$$\begin{aligned}\mathcal{D}_{t,z}\phi(F) &= \left(\frac{1}{\sqrt{2\pi}}\right)^{-q}\int_{y\in R^q}\sum_{n=0}^{\infty}\frac{1}{n!}i^n\Big(\big((y_1F_1+...+y_qF_q)+\mathcal{D}_{t,z}(y_1F_1+...+y_qF_q)\big)^n\\ &\quad -(y_1F_1+...+y_qF_q)\Big)\hat{\phi}(y)dy\\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^{-q}\int_{y\in R^q}\Big(e^{i((y_1F_1+...+y_qF_q)+\mathcal{D}_{t,z}(y_1F_1+...y_qF_q)e)}-e^{i(y_1F_1+...y_qF_q)}\Big)\hat{\phi}(y)dy\\ &= \phi(F+(\mathcal{D}_{t,z}F)e)-\phi(F),\end{aligned}$$

where i is the unit imaginary number. Next, for a general $F\in\mathcal{D}_{1,2}$, the lemma is proved using an approximation by a sequence $F^n\in\mathcal{D}^{\mathcal{E}}_{1,2}$ (see Lemma 9.8 and the proof for Theorem 12.8 [40]).This completes the proof of the lemma. ∎

LEMMA 4.8. *Any $F\in L^2_{\mathcal{F}_T}(\Omega,R^q)$ can be approximated by a sequence $\{F^n,n\in\{1,2,...\}\}\subset\mathcal{D}_{1,2}$ in the sense that*

$$\|F^n-F\|^2_{L^2(\Omega,P)}\to 0 \text{ as } n\to\infty, \tag{4.45}$$

where the norm (4.45) is defined in (4.26).

Proof. By combining the discussion for Theorem 4.3.3 [37] and for Theorem 9.10 [40], we have the following Itô's representation formula

$$F_r=E[F_r]+\int_{[0,T]\times\mathcal{Z}^m}\Big\langle\Psi^r(t,z),M(dt,dz)\Big\rangle$$

for each $F\in L^2_{\mathcal{F}_T}(\Omega,R^q)$ and $r\in\{1,...,q\}$, where $\Psi^r(t,z)$ is an m-dimensional predictable process for $(t,z)\in[0,T]\times\mathcal{Z}^m$. Then, as in (4.41) for proving Lemma 4.7, we can obtain

$$\begin{aligned}F_r &= E[F_r]+\int_{[0,T]\times\mathcal{Z}^m}\Big\langle E\left[\Psi^r_1(t^1,z^1)\right],M(dt^1,dz^1)\Big\rangle\\ &\quad+\int_{[0,T]\times\mathcal{Z}^m}\int_{[0,t^1]\times\mathcal{Z}^m}\Big\langle\Psi^r_2(t^2,z^2)M(dt^2,dz^2),M(dt^1,dz^1)\Big\rangle,\end{aligned}$$

where $\Psi_2^r(t^2,z^2)$ is an $m\times q$ predictable matrix process. Along this line of iterative calculations, we can conclude that

$$F=\sum_{k=0}^{\infty}M_k(h_k)$$

for some $h_n\in H^n$ (see the proof for Theorem 10.2 in [40]). Finally, for each $n\in\{1,2,...\}$, we define

$$F^n=\sum_{k=0}^{n}M_k(h_k).$$

Then, we know that $F^n\in\mathcal{D}_{1,2}$ and $F^n\to F$ as $n\to\infty$ in the sense as stated in (4.45). This completes the proof of the lemma. □

4.3. Proof of Theorem 2.1. We divide the proof of the theorem into four parts which correspond to the different boundary reflection conditions.

Part A (Existence). We consider the case that $\hat{L}(t,\omega)$ as it appeared in (2.10)–(2.15) is a constant and both the forward and the backward SDEs have reflection boundaries. In this case, we need to prove the claim that there is an adapted weak solution $((X,Y),(V,\bar{V},\tilde{V},F))$ to the system in (1.1)–(1.5).

In fact, for a positive integer b, let $D^2_{\mathcal{F}}([0,T],R^b)$ be the space of R^b-valued and $\{\mathcal{F}_t\}$-adapted process Y with sample paths in the Skorohod topological space $D([0,T],R^b)$, which is square-integrable, i.e.,

$$E\left[\int_0^T\|Y(t)\|^2dt\right]<\infty. \tag{4.46}$$

Furthermore, we use $D^2_{\mathcal{F},p}([0,T],R^b)$ to denote the corresponding predictable space. Then, for a given $n\in\{1,2,...\}$ and a 4-tuple

$$\begin{aligned}(X^n,V^n,\bar{V}^n,\tilde{V}^n)\in D^2_{\mathcal{F}}([0,T],R^p)\times D^2_{\mathcal{F}}([0,T],R^q)\times D^2_{\mathcal{F},p}([0,T],R^{q\times d})\\ \times D^2_{\mathcal{F},p}([0,T]\times\mathcal{Z}^h,R^{q\times h})\end{aligned} \tag{4.47}$$

with $X^n(0)\in D$ and $V^n(T)\in\bar{D}$, and by the study of the continuous dynamic complementarity problem (DCP) [3], [12], [10], [34], [42], we know that there is a 2-tuple process $(X^{n+1},Y^{n+1})\in D^2_{\mathcal{F}}([0,T],R^p)\times D^2_{\mathcal{F}}([0,T],R^b)$ such that

$$X^{n+1}(t)=X(0)+Z^n(t)+RY^{n+1}(t)\in D \tag{4.48}$$

along each sample path. Furthermore, the process Z in (4.48) has the decomposition

$$Z^n(t)=Z_1^n(t)+Z_2^n(t),$$

where,

$$\begin{aligned}Z_1^n(t)&=\int_0^t b(s^-,X^n,V^n,\bar{V}^n,\tilde{V}^n,u)ds,\\ Z_2^n(t)&=\int_0^t\sigma(s^-,X^n,V^n,\bar{V}^n,\tilde{V}^n,u)dW(s)\\ &\quad+\int_0^t\int_{\mathcal{Z}^h}\eta(s^-,X^n,V^n,\bar{V}^n,\tilde{V}^n,u,z)\tilde{N}(ds,dz).\end{aligned}$$

In addition, (X^{n+1},Y^{n+1}) satisfies property (3) in Definition 3.1, i.e., for all $t\geq 0$,

$$\int_0^t I_{D_i}(X^{n+1}(s))dY_i^{n+1}(s)=Y_i^{n+1}(t). \tag{4.49}$$

Since $Y^{n+1}(t)$ is finite a.s. along each sample path, it can be approximated by a sequence of non-decreasing continuous processes. Thus, without loss of generality, we assume that it is non-decreasing and continuous. Therefore, we have the expression

$$Y_i^{n+1}(t)=\int_0^t \alpha_i^{n+1}(s)ds \tag{4.50}$$

where $\alpha_i^{n+1}(t)\geq 0$ is the corresponding derivative of $Y_i^{n+1}(t)$ along each sample path a.s. However, α^{n+1} may be unbounded. Hence, we approximate it by a sequence of b-dimensional bounded vector processes $\alpha^{c,n+1}(t)$ for each $c\in\{1,2,...\}$ whose component associated with each $i\in\{1,...,b\}$ is given by

$$\alpha_i^{c,n+1}(t)\equiv\alpha_i^{n+1}(t)I\left\{\alpha_i^{n+1}(t)\leq c\right\}. \tag{4.51}$$

Then, by the monotone convergence theorem, we have that

$$\begin{aligned}\left\|Y^{n+1}-Y^{c,n+1}\right\|_{[0,T]} &= \sup_{t\in[0,T]}\left\|Y^{n+1}(t)-Y^{c,n+1}(t)\right\| \\ &\leq \max_{i\in\{1,...,b\}}\int_0^T \alpha_i^{n+1}(s)I\left\{\alpha_i^{n+1}(s))\geq c\right\}ds \\ &\to 0 \text{ a.s. as } c\to\infty. \end{aligned}\tag{4.52}$$

Furthermore,

$$E\left[\left\|Y^{n+1}-Y^{c,n+1}\right\|^2_{[0,T]}\right]\to 0 \text{ as } c\to\infty. \tag{4.53}$$

In addition, for each given c, it follows from Lemma 2.4 of [30] that there exists a sequence $\left\{\alpha^{r,c,n+1},r\in\{1,2,...\}\right\}$ of simple processes such that

$$E\left[\int_0^T \left\|\alpha^{r,c,n+1}(s)-\alpha^{c,n+1}(s)\right\|^2 ds\right]\to 0 \text{ as } r\to\infty. \tag{4.54}$$

Next, for each sufficiently large c and r, let t_i denote a dissection point of $[0,T]$ with $t_0=0$ such that $\cup_{i\in\{0,1,...,I-1\}}[t_i,t_{i+1})=[0,T]$ for some integer $I\in\{1,2,...\}$. Due to Lemma 4.8, each random vector $\alpha^{r,c,n+1}(t_i)$ can be approximated by a sequence $\{\alpha^{v,r,c,n+1}(t_i),v\in\{1,2,...\}\}$ of random vectors that are Malliavin differentiable such that

$$E\left[\left\|\{\alpha^{v,r,c,n+1}(t_i)-\alpha^{r,c,n+1}(t_i)\right\|^2\right]\to 0 \text{ as } v\to\infty. \tag{4.55}$$

Then, it follows from (4.55) that the corresponding simple process $\alpha^{v,r,c,n+1}(t)$ constructed by $\{\alpha^{v,r,c,n+1}(t_i)\}_{\{0\leq i\leq I-1\}}$ with $t\in[0,T]$ for each $v\in\{1,2,...\}$ is Malliavin differentiable. Hence, we know that

$$Y^{v,r,c,n+1}(t)\equiv\int_0^t \alpha^{v,r,c,n+1}(s)ds \tag{4.56}$$

is also Malliavin differentiable for each $t\in[0,T]$. Furthermore, we can write the equation corresponding to (4.48) and $Y^{v,r,c,n+1}(t)$ along $n\in\{1,2,...\}$ as

$$X^{v,r,c,n+1}(t)=X(0)+Z^{v,r,c,n}(t)+RY^{v,r,c,n+1}(t). \tag{4.57}$$

Then, iterating in terms of $n\in\{1,2,...\}$, and owing to Lemma 4.8, we know that $X^{v,r,c,n+1}(t)$ is Malliavin differentiable. More precisely, for each $t,\theta\in[0,T]$, we have

$$\mathcal{D}_{\theta,z}X^{v,r,c,n+1}(t)=\mathcal{D}_{\theta,z}X(0)+\mathcal{D}_{\theta,z}Z^{v,r,c,n}(t)+\mathcal{D}_{\theta,z}\left(RY^{v,r,c,n+1}(t)\right). \tag{4.58}$$

Note that if any component of z in (4.58) is zero, the corresponding Malliavin derivatives are in terms of the associated Brownian motion component. Furthermore, by Lemma 4.6 and Lemma 4.7, we have that

$$\mathcal{D}_{\theta,z}Z^{v,r,c,n}(t)=\Theta^{v,r,c,n}(t,z)+\mathcal{D}_{\theta,z}Z_1^{v,r,c,n}(t)+\mathcal{D}_{\theta,z}Z_2^{v,r,c,n}(t), \tag{4.59}$$

where if we use $\bar{h}$ to denote the number of zero components of z, the initial value can be calculated as

$$\Theta^{v,r,c,n}(t,z)=\begin{cases}0 & \text{if } \theta\in[t,T],\\ \sigma(t,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u) & \text{if } z=0,\theta<t,\\ \eta(t,X^{v,r,,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u,z) & \text{if } z\in\mathcal{Z}^{h-\bar{h}}\times\{0\}^{\bar{h}},\theta<t,\end{cases} \tag{4.60}$$

where $\bar{h}<h$ in the second equation of (4.60) and the notion $\{0\}^{\bar{h}}$ denotes the product of $\bar{h}$ number of sets $\{0\}$, i.e., $\{0\}^{\bar{h}}=\{0\}\times...\times\{0\}$. In addition, we have

$$\begin{aligned}\mathcal{D}_{\theta,z}Z_1^{v,r,c,n}(t)&=\int_0^t\mathcal{D}_{\theta,z}b(s^-,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u)ds,\\ \mathcal{D}_{\theta,z}Z_2^{v,r,c,n}(t)&=\int_0^t\mathcal{D}_{\theta,z}\sigma(s^-,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u)d\tilde{W}(s)\\ &\quad+\int_0^t\int_{\mathcal{Z}^h}\mathcal{D}_{\theta,z}\eta(s^-,X^{v,r,,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u,y)\tilde{\tilde{N}}(ds,dy),\end{aligned}$$

where $\tilde{W}$ is a $hd\times d$ matrix process given by $\tilde{W}=(W',...,W')'$ and $\tilde{\tilde{N}}$ is a $hh\times h$ measure given by $\tilde{\tilde{N}}=(\tilde{N}',...,\tilde{N}')'$. Note that for each $f\in\{b,\sigma,\eta\}$ (e.g., $f=b$) and by the chain rule in Lemma 4.7, we can interpret the Malliavin derivative of f as follows,

$$\begin{aligned}&\mathcal{D}_{\theta,z}b(s^-,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u)\\ &=b(s^-,X^{v,r,c,n}+\mathcal{D}_{\theta,z}X^{v,r,c,n}e_h,V^{v,r,c,n}+\mathcal{D}_{\theta,z}V^{v,r,c,n}e_h,\\ &\qquad\bar{V}^{v,r,c,n}+\mathcal{D}_{\theta,z}\bar{V}^{v,r,c,n}e_{dh},\tilde{V}^{v,r,c,n}+\mathcal{D}_{\theta,z}\tilde{V}^{v,r,c,n}e_{hh},*+\mathcal{D}_{\theta,z}*,u)\\ &\quad-b(s^-,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u),\end{aligned} \tag{4.61}$$

where e_m for each $m\in\{p,dh,hh\}$ is the m-dimensional vector of ones and $\mathcal{D}_{\theta,z}*$ denotes the Malliavin derivative in terms of the known random factors as explained in (1.3).

Now, it follows from the martingale representation theorem (see Theorem 5.3.5 in [2]), the discussion about BSDEs with Lévy jumps (see Proposition 18 in [16]), and the above mentioned study on DCP, we know that there is a 4-tuple process

$$\begin{aligned}&(V^{n+1},\bar{V}^{n+1},\tilde{V}^{n+1},F^{n+1}))\\ &\in D^2_{\mathcal{F}}([0,T],R^q)\times D^2_{\mathcal{F},p}([0,T],R^{q\times d})\\ &\quad\times D^2_{\mathcal{F},p}([0,T]\times\mathcal{Z}^h,R^{q\times h})\times D^2_{\mathcal{F}}([0,T],R^{q\times\bar{b}})\end{aligned} \tag{4.62}$$

such that

$$V^{n+1}(t)=H(X^n(T),\cdot)+SF^n(T)+U^n(t)-SF^{n+1}(t)\in\bar{D} \tag{4.63}$$

along each sample path. Furthermore, (V^{n+1},F^{n+1}) satisfies property (3) in Definition 3.1. More precisely, F^{n+1} is a q-dimensional $\{\mathcal{F}_t\}$-adapted process such that the i-th component F_i^{n+1} of F^{n+1} for each $i\in\{1,...,\bar{b}\}$ $\mathbf{P}$-a.s. has the properties that $F_i^{n+1}(0)=0$, F_i^{n+1} is non-decreasing and F_i^{n+1} can increase only when V^{n+1} is on the corresponding boundary face $\bar{D}_i$, i.e.,

$$\int_0^t I_{\bar{D}_i}(V^{n+1}(s))dF_i^{n+1}(s)=F_i^{n+1}(t) \text{ for all } t\geq 0. \tag{4.64}$$

In addition, the process $U^n(t)$ in (4.63) has the decomposition

$$U^n(t)=U_1^n(t)-U_2^n(t)-U_3^n(t), \tag{4.65}$$

where

$$\begin{aligned}
U_1^n(t)=&\int_t^T c(s^-,X^n,V^n,\bar{V}^n,\tilde{V}^n,u)ds,\\
U_2^n(t)=&\int_t^T \left(\alpha(s^-,X^n,V^n,\bar{V}^n,\tilde{V}^n,u)-\bar{V}^n(s^-)\right)dW(s)\\
&+\int_t^T\int_{\mathcal{Z}^h}\left(\zeta(s^-,X^n,V^n,\bar{V}^n,\tilde{V}^n,u,z)-\tilde{V}^n(s^-,z)\right)\tilde{N}(ds,dz),\\
U_3^n(t)=&\int_t^T \bar{V}^{n+1}(s^-)dW(s)+\int_t^T\int_{\mathcal{Z}^h}\tilde{V}^{n+1}(s^-,z)\tilde{N}(ds,dz).
\end{aligned}$$

Just as in the case for (4.57), there is a sequence $\{\beta^{v,r,c,n+1}(t_i),v\in\{1,2,...\}\}$ of random vectors that are Malliavin differentiable such that

$$F^{v,r,c,n+1}(t)\equiv\int_0^t\beta^{v,r,c,n+1}(s)ds \tag{4.66}$$

is also Malliavin differentiable for each $n\in\{1,2,...\}$ and $v\in\{1,2,...\}$. Thus, we can express the equation corresponding to (4.63) as

$$V^{v,r,c,n+1}(t)=H(X^{v,r,c,n}(T),\cdot)+SF^{v,r,c,n}(T)+U^{v,r,c,n}(t)-SF^{v,r,c,n+1}(t), \tag{4.67}$$

where

$$U^{v,r,c,n}(t)=U_1^{v,r,c,n}(t)-U_2^{v,r,c,n}(t)-U_3^{v,r,c,n}(t). \tag{4.68}$$

Furthermore, we have the following expressions for all $U_i^{v,r,c,n}(t)$ with $i\in\{1,2,3\}$:

$$\begin{aligned}
U_1^{v,r,c,n}(t)=&\int_t^T c(s^-,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u)ds,\\
U_2^{v,r,c,n}(t)=&\int_t^T\left(\alpha(s^-,X^{v,r,c,n},V^n,\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u)-\bar{V}^{v,r,c,n}(s^-)\right)dW(s)\\
&+\int_t^T\int_{\mathcal{Z}^h}\Big(\zeta(s^-,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u,z)\\
&\qquad-\tilde{V}^{v,r,c,n}(s^-,z)\Big)\tilde{N}(ds,dz),\\
U_3^{v,r,c,n}(t)=&\int_t^T\bar{V}^{v,r,c,n+1}(s^-)dW(s)+\int_t^T\int_{\mathcal{Z}^h}\tilde{V}^{v,r,c,n+1}(s^-,z)\tilde{N}(ds,dz).
\end{aligned}$$

Then, for every $t,\theta\in[0,T]$ and by taking Malliavin derivatives on both sides of (4.68), we have

$$\begin{aligned}
\mathcal{D}_{\theta,z}V^{v,r,c,n+1}(t)=&\mathcal{D}_{\theta,z}H(X^{v,r,c,n}(T),\cdot)+\mathcal{D}_{\theta,z}(SF^{v,r,c,n}(T))\\
&+\mathcal{D}_{\theta,z}U^{v,r,c,n}(t)-\mathcal{D}_{\theta,z}(SF^{v,r,c,n+1}(t)).
\end{aligned} \tag{4.69}$$

Furthermore, it follows from Lemma 4.6 and Lemma 4.7 that

$$\mathcal{D}_{\theta,z}U^{v,r,c,n}(t)=\Upsilon^{v,r,c,n}(t,z)+\mathcal{D}_{\theta,z}U_1^{v,r,c,n}(t)-\mathcal{D}_{\theta,z}U_2^{v,r,c,n}(t)-\mathcal{D}_{\theta,z}U_3^{v,r,c,n}(t), \tag{4.70}$$

where if $\theta\leq t$, we have

$$\Upsilon^{v,r,c,n}(t,z)=0. \tag{4.71}$$

Otherwise, for $z=0$, we have

$$\begin{aligned}\Upsilon^{v,r,c,n}(t,z)&\equiv\Upsilon^{v,r,c,n}(t)\\&=-\left(\alpha(t,X^{v,r,c,n},V^n,\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u)-\bar{V}^{v,r,c,n}(t)\right)\\&\quad-\bar{V}^{v,r,c,n+1}(t).\end{aligned} \tag{4.72}$$

Moreover, as explained in (4.60), for each $z\in\mathcal{Z}^{h-\bar{h}}\times\{0\}^{\bar{h}}$ with $\bar{h}<h$, we have

$$\begin{aligned}\Upsilon^{v,r,c,n}(t,z)=&-\left(\zeta(t,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u,z)-\tilde{V}^{v,r,c,n}(t,z)\right)\\&-\tilde{V}^{v,r,c,n+1}(t,z).\end{aligned} \tag{4.73}$$

In addition, for each $t,\theta\in[0,T]$, we have that

$$\begin{aligned}\mathcal{D}_{\theta,z}U_1^{v,r,c,n}(t)=&\int_t^T\mathcal{D}_{\theta,z}c(s^-,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u)ds,\\ \mathcal{D}_{\theta,z}U_2^{v,r,c,n}(t)=&\int_t^T\mathcal{D}_{\theta,z}\left(\alpha(s^-,X^{v,r,c,n},V^n,\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u)-\bar{V}^{v,r,c,n}(s^-)\right)d\tilde{W}(s)\\&+\int_t^T\int_{\mathcal{Z}^h}\mathcal{D}_{\theta,z}\Big(\zeta(s^-,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u,y)\\&\qquad-\tilde{V}^{v,r,c,n}(s^-,y)\Big)\tilde{\tilde{N}}(ds,dy),\\ \mathcal{D}_{\theta,z}U_3^{v,r,c,n}(t)=&\int_t^T\mathcal{D}_{\theta,z}\bar{V}^{v,r,c,n+1}(s^-)d\tilde{W}(s)+\int_t^T\int_{\mathcal{Z}^h}\mathcal{D}_{\theta,z}\tilde{V}^{v,r,c,n+1}(s^-,y)\tilde{\tilde{N}}(ds,dy),\end{aligned}$$

where $\tilde{W}$ and $\tilde{\tilde{N}}$ are defined in (4.58). Furthermore, for each $f\in\{c,\alpha,\zeta\}$, its Malliavin derivative $\mathcal{D}_{\theta,z}f$ in the corresponding integral is interpreted as in (4.61). Note that if we take $\theta=t$ and replace T by a time $s\in[\theta,T]$ in the associated integrals of (4.70), the Itô integral corresponding to the Brownian motion is continuous a.s. with respect to $s\in[\theta,T]$, while the integral corresponding to the pure Lévy jump process is cádlág a.s. in terms of $s\in[\theta,T]$ (see Theorem 4.2.12 and 4.2.14 [2] and the proof for Corollary 4.1 in [17]). Then, by taking $s\downarrow t$, it follows from (4.70) and (4.72) that

$$\begin{aligned}\bar{V}^{v,r,c,n+1}(t)=&\alpha(t,X^{v,r,c,n},V^n,\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u)-\bar{V}^{v,r,c,n}(t)\\&-\mathcal{D}_{t,0}U^{v,r,c,n}(t)\end{aligned} \tag{4.74}$$

for almost all $(t,\omega)\in[0,T]\times\Omega$. Furthermore, it follows from (4.70) and (4.73) that

$$\begin{aligned}\tilde{V}^{v,r,c,n+1}(t,z)=&\zeta(t,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u,z)-\tilde{V}^{v,r,c,n}(t,z)\\&-\mathcal{D}_{t,z}U^{v,r,c,n}(t)\end{aligned} \tag{4.75}$$

for any $z\in\mathcal{Z}^{h-\bar{h}}\times\{0\}^{\bar{h}}$ with the corresponding $\bar{h}<h$ as explained in (4.60).

Thus, by summarizing the discussions for (4.48) and (4.63), we have the conclusion that there is a 6-tuple process

$$((X^{n+1},Y^{n+1}),(V^{n+1},\bar{V}^{n+1},\tilde{V}^{n+1},F^{n+1}))\in\mathcal{Q}_{\mathcal{F}}^2([0,T]) \tag{4.76}$$

with the properties as required by (1.1)–(1.5), where $\mathcal{Q}^2_{\mathcal{F}}([0,T])$ is defined in (2.5). Then, we can prove that the following sequence of stochastic processes along $n\in\{1,2,...\}$,

$$\Xi^n=((X^{n+1},Y^{n+1}),(V^{n+1},\bar{V}^{n+1},\tilde{V}^{n+1}(\cdot),F^{n+1})),\,(X^1,V^1,\bar{V}^1,\tilde{V}^1)=(0,0,0,0) \tag{4.77}$$

is relatively compact in the Skorohod topology over the space

$$\begin{aligned}\mathcal{P}[0,T]\equiv\, & D^2_{\mathcal{F}}([0,T],R^p)\times D^2_{\mathcal{F}}([0,T],R^b)\\ & \times D^2_{\mathcal{F}}([0,T],R^q)\times D^2_{\mathcal{F},p}([0,T],R^{q\times d})\\ & \times D^2_{\mathcal{F},p}([0,T],R^{q\times h})\times D^2_{\mathcal{F}}([0,T],R^{\bar{b}}).\end{aligned} \tag{4.78}$$

Along the lines of [12, 15], [10], and by Corollary 7.4 in [20], it suffices to prove the following two conditions are true. First, for each $\epsilon>0$ and rational $t>0$, there is a constant $C(\epsilon,t)$ such that

$$\liminf_{n\to\infty}P\left\{\|\Xi^n\|^2\leq C(\epsilon,t)\right\}\geq 1-\epsilon; \tag{4.79}$$

Second, for each $\epsilon>0$ and $T>0$, there is a constant $\delta>0$ such that

$$\limsup_{n\to\infty}P\left\{w(\Xi^n,\delta,T)\geq\epsilon\right\}\leq\epsilon, \tag{4.80}$$

where the definition of w is stated in (4.2).

To prove the two conditions in (4.79) and (4.80) to be true, we first define the norm along each sample path

$$\|f\|_{[a,b]}=\sup_{a\leq t\leq b}\|f(t)\|$$

for every $f\in\{X^n,Z^n,U^n,V^n,\bar{V}^n,\tilde{V}^n\}$ with $a,b\in[0,T]$. Then, for some constant $\gamma>0$ that will be chosen and explained in the following proof, we introduce the space

$$\begin{aligned}\mathcal{Q}_\gamma[0,T]\equiv\, & D^2_{\mathcal{F}}([0,T],R^p)\times D^2_{\mathcal{F}}([0,T],R^q)\times D^2_{\mathcal{F},p}([0,T],R^{q\times d})\\ & \times D^2_{\mathcal{F},p}([0,T]\times\mathcal{Z}^h,R^{q\times h}).\end{aligned} \tag{4.81}$$

Note that for each 4-tuple process $(X,V,\bar{V},\tilde{V})$ in this space, the norm is defined by

$$\begin{aligned}\left\|(X,V,\bar{V},\tilde{V})\right\|^2_{\mathcal{Q}_\gamma[0,T]}\equiv\, & E\left[\sup_{t\in[0,T]}\left(\|X(t)\|^2+\|V(t)\|^2\right)e^{2\gamma t}\right]\\ & +E\left[\int_0^T\left\|\bar{V}(t)\right\|^2e^{2\gamma t}dt\right]+E\left[\int_0^T\left\|\tilde{V}(t,\cdot)\right\|^2_\nu e^{2\gamma t}dt\right].\end{aligned} \tag{4.82}$$

Thus, by Lemma 4.1, there is a positive constant C_1 such that

$$\begin{aligned}&\left\|(X^{n+1},Y^{n+1})(t)\right\|\\ &\leq\left\|(X^{n+1},Y^{n+1})(0)\right\|+\kappa\mathrm{Osc}(Z^n,[0,T])\\ &\leq C_1\left(\|X(0)\|+\|Z^n\|_{[0,T]}\right).\end{aligned} \tag{4.83}$$

Furthermore, there are two nonnegative constants $\bar{C}_1$ and $\bar{C}_2$ such that

$$
\begin{aligned}
&\left\|(V^{n+1},\bar{V}^{n+1},\tilde{V}^{n+1}(\cdot),F^{n+1})(t)\right\| \\
\leq &\left\|(V^{n+1},\bar{V}^{n+1},\tilde{V}^{n+1}(\cdot))(t)\right\| + \left\|F^{n+1}(t)\right\| \\
\leq &\left\|V^{n+1}(T)\right\| + \left\|F^{n+1}(0)\right\| + 2\kappa \mathrm{Osc}(U^n,[0,T]) \\
&+ \left\|\Delta\bar{V}^{v,r,c,n+1}(t)\right\| + \left\|\Delta\tilde{V}^{v,r,c,n+1}(\cdot)(t)\right\| + \|\mathcal{D}_{t,0}U^{v,r,c,n}(t)\| + \|\mathcal{D}_{t,\cdot}U^{v,r,c,n}(t)\|\Big) \\
\leq &\bar{C}_1\Big(\|V^n(T)\| + \|F^n(T)\| + \|U^n\|_{[0,T]} \\
&+ \left\|\Delta\bar{V}^{v,r,c,n+1}(t)\right\| + \left\|\Delta\tilde{V}^{v,r,c,n+1}(\cdot)(t)\right\| + \|\mathcal{D}_{t,0}U^{v,r,c,n}(t)\| + \|\mathcal{D}_{t,\cdot}U^{v,r,c,n}(t)\|\Big) \\
\leq &\bar{C}_2\Big(1 + \|X^n(T)\| + \left\|U^{n-1}\right\|_{[0,T]} + \|U^n\|_{[0,T]} \\
&+ \left\|\Delta\bar{V}^{v,r,c,n+1}(t)\right\| + \left\|\Delta\tilde{V}^{v,r,c,n+1}(\cdot)(t)\right\| + \|\mathcal{D}_{t,0}U^{v,r,c,n}(t)\| + \|\mathcal{D}_{t,\cdot}U^{v,r,c,n}(t)\|\Big) \\
\leq &C_1\Big(1 + \|X(0)\| + \left\|Z^{n-1}\right\|_{[0,T]} + \left\|U^{n-1}\right\|_{[0,T]} + \|U^n\|_{[0,T]} \\
&+ \left\|\Delta\bar{V}^{v,r,c,n+1}(t)\right\| + \left\|\Delta\tilde{V}^{v,r,c,n+1}(\cdot)(t)\right\| + \|\mathcal{D}_{t,0}U^{v,r,c,n}(t)\| + \|\mathcal{D}_{t,\cdot}U^{v,r,c,n}(t)\|\Big),
\end{aligned} \tag{4.84}
$$

where the second inequality follows from (4.74) and (4.75), and moreover,

$$\Delta\bar{V}^{v,r,c,n+1}(t) = \bar{V}^{n+1}(t) - \bar{V}^{v,r,c,n+1}(t), \tag{4.85}$$

$$\Delta\tilde{V}^{v,r,c,n+1}(\cdot)(t) = \tilde{V}^{n+1}(\cdot)(t) - \tilde{V}^{v,r,c,n+1}(\cdot)(t), \tag{4.86}$$

$$\mathcal{D}_{t,\cdot}U^{v,r,c,n}(t) = \int_{\mathcal{Z}^h} \mathcal{D}_{t,z}U^{v,r,c,n}(t)\lambda\nu(dz) \tag{4.87}$$

with $\lambda\nu(dz) = (\lambda_1\nu_1(dz_1),\ldots,\lambda_h\nu_h(dz_h))'$. Therefore, it follows from Markov's inequality and the linear growth condition that

$$P\Big\{\|Z_1^n\|_{[0,T]} \geq K\Big\} \leq \frac{2\hat{L}^2T}{K-\hat{L}T}\left\|(X^n,V^n,\bar{V}^n,\tilde{V}^n)\right\|_{\mathcal{Q}_\gamma[0,T]} \tag{4.88}$$

for each $n\in\{1,2,\ldots\}$ and any constant $K>\hat{L}T$. Furthermore, by Lemma 4.2.8 in [2] (or the related theorem in [22]) and the linear growth condition, we know that

$$P\Big\{\|Z_2^n\|_{[0,T]} \geq K\Big\} \leq \frac{\bar{K}}{K^2} + \frac{\hat{L}^2T}{\bar{K}-\hat{L}^2T}\left\|(X^n,V^n,\bar{V}^n,\tilde{V}^n(\cdot))\right\|^2_{\mathcal{Q}_\gamma[0,T]} \tag{4.89}$$

for all constant $\bar{K}>\hat{L}^2T$. Using similar reasoning as was applied to the inequalities in (4.84)–(4.83), we know that

$$P\Big\{\|U_1^n\|_{[0,T]} \geq K\Big\} \leq \frac{2\hat{L}^2T}{(K-\hat{L}T)^2}\left\|(X^n,V^n,\bar{V}^n,\tilde{V}^n)\right\|^2_{\mathcal{Q}_\gamma[0,T]}, \tag{4.90}$$

$$P\Big\{\|U_2^n\|_{[0,T]} \geq K\Big\} \leq \frac{\bar{K}}{K^2} + \frac{\hat{L}^2T}{\bar{K}-\hat{L}^2T}\left\|(X^n,V^n,\bar{V}^n,\tilde{V}^n)\right\|^2_{\mathcal{Q}_\gamma[0,T]}, \tag{4.91}$$

$$P\Big\{\|U_3^n\|_{[0,T]} \geq K\Big\} \leq \frac{\bar{K}}{K^2} + \frac{1}{\bar{K}}\left\|(X^{n+1},V^{n+1},\bar{V}^{n+1},\tilde{V}^{n+1})\right\|^2_{\mathcal{Q}_\gamma[0,T]}. \tag{4.92}$$

Note that for any $t\in[0,T]$, it follows from the proof of Proposition 18 for an BSDE with jumps in [16] and from Lemma 4.1 that

$$\begin{aligned}
&\left\|(U^n,\bar{V}^n,\tilde{V}^n)\right\|^2_{\mathcal{Q}_\gamma[t,T]}\\
&\le K_\gamma\left(2\hat{L}^2(T-t)+\left\|(X^{n-1},V^{n-1},\bar{V}^{n-1},\tilde{V}^{n-1})\right\|^2_{\mathcal{Q}_\gamma[t,T]}\right)\\
&\le K_\gamma\left(2\hat{L}^2(T-t)+e^{2\gamma T}E\left[\|V^{n-1}\|^2_{[t,T]}\right]+e^{2\gamma T}\int_t^T E\left[\|X^{n-1}\|^2_{[0,s]}\right]ds\right)\\
&\quad+K_\gamma\left\|(U^{n-1},\bar{V}^{n-1},\tilde{V}^{n-1})\right\|^2_{\mathcal{Q}_\gamma[t,T]},
\end{aligned}\tag{4.93}$$

where there exists a nonnegative constant $\tilde{K}$ depending only on $\hat{L}$, T, d, and h such that

$$K_\gamma=\frac{\tilde{K}}{\gamma}<1\tag{4.94}$$

for some suitable chosen $\gamma>0$. Therefore, by Lemma 4.1, the It$\hat{o}$'s isometry formula, and (4.93), we know that

$$\begin{aligned}
&E\left[\|V^n\|^2_{[t,T]}\right]\\
&\le\bar{K}_1\left(E\left[\|V^n(T)\|^2\right]+E\left[\|F^{n-1}(T)\|^2\right]+\kappa^2E\left[\mathrm{Osc}(U^{n-1},[t,T])^2\right]\right)\\
&\le K_1\left(1+E\left[\|X^n\|^2_{[0,T]}\right]+\kappa^2E\left[\mathrm{Osc}(U^{n-2},[0,T])^2\right]+\kappa^2E\left[\mathrm{Osc}(U^{n-1},[t,T])^2\right]\right)\\
&\le K_1\left(1+24\kappa^2\hat{L}^2T^2+24\kappa^2L^2(T-t)^2\right)+K_1E\left[\|X^n\|^2_{[0,T]}\right]\\
&\quad+24K_1\kappa^2\hat{L}^2T\left(\int_0^T E\left[\|X^{n-2}\|^2_{[0,s]}\right]ds+E\left[\|V^{n-2}\|^2_{[0,T]}\right]\right)\\
&\quad+24K_1\kappa^2\hat{L}^2(T-t)\left(\int_t^T E\left[\|X^{n-1}\|^2_{[0,s]}\right]ds+E\left[\|V^{n-1}\|^2_{[t,T]}\right]\right)\\
&\quad+24K_1\kappa^2\hat{L}^2T\left\|(U^{n-2},\bar{V}^{n-2},\tilde{V}^{n-2})\right\|^2_{\mathcal{Q}_\gamma[0,T]}\\
&\quad+4K_1\kappa^2\left\|(U^{n-1},\bar{V}^{n-1},\tilde{V}^{n-1})\right\|^2_{\mathcal{Q}_\gamma[0,T]}\\
&\quad+24K_1\kappa^2\hat{L}^2(T-t)\left\|(U^{n-1},\bar{V}^{n-1},\tilde{V}^{n-1})\right\|^2_{\mathcal{Q}_\gamma[t,T]}\\
&\quad+4K_1\kappa^2\left\|(U^n,\bar{V}^n,\tilde{V}^n)\right\|^2_{\mathcal{Q}_\gamma[t,T]}\\
&\le K_3+K_2\left(\int_0^T E\left[\|X^{n-1}\|^2_{[0,s]}\right]ds+\left\|(U^{n-2},\bar{V}^{n-2},\tilde{V}^{n-2})\right\|^2_{\mathcal{Q}_\gamma[0,T]}\right.\\
&\quad\left.+\left\|(U^{n-1},\bar{V}^{n-1},\tilde{V}^{n-1})\right\|^2_{\mathcal{Q}_\gamma[0,T]}+\left\|(U^n,\bar{V}^n,\tilde{V}^n)\right\|^2_{\mathcal{Q}_\gamma[0,T]}\right),
\end{aligned}\tag{4.95}$$

where K_i for $i\in\{1,2,3\}$ are some nonnegative constants depending only on T, $\hat{L}$, κ, and

$E\left[\|V(T)\|^2\right]$. Furthermore, for any $t\in[0,T]$, we have

$$
\begin{aligned}
E\left[\|X^n\|_{[0,t]}^2\right] &\leq 2E\left[\|X(0)\|^2\right]+2\kappa^2E\left[\mathrm{Osc}(Z^{n-1},[0,t])^2\right]\\
&\leq 2E\left[\|X(0)\|^2\right]+6\kappa^2\hat{L}^2t^2\\
&\quad +6\kappa^2\hat{L}^2t\left(\int_0^t E\left[\|X^{n-1}\|_{[0,s]}^2\right]ds+E\left[\|V^{n-1}\|_{[0,T]}^2\right]\right)\\
&\quad +6\kappa^2\hat{L}^2t\left\|(U^{n-1},\bar{V}^{n-1},\tilde{V}^{n-1})\right\|_{\mathcal{Q}_\gamma[0,T]}^2\\
&\leq 2E\left[\|X(0)\|^2\right]+12\kappa^4\hat{L}^2t^2+6\kappa^2\hat{L}^2tE\left[\|V^2(T)\|\right]\\
&\quad +6\kappa^2\hat{L}^2t\int_0^t E\left[\|X^{n-1}\|_{[0,s]}^2\right]ds\\
&\quad +6\kappa^2\hat{L}^2t(1+2\kappa^2)\left\|(U^{n-1},\bar{V}^{n-1},\tilde{V}^{n-1})\right\|_{\mathcal{Q}_\gamma[0,T]}^2. \qquad (4.96)
\end{aligned}
$$

Thus, by repeating the calculations in (4.90)–(4.96) for all $n\in\{1,2,\ldots\}$, we have that

$$
\left\|(X^n,V^n,\bar{V}^n,\tilde{V}^n)\right\|_{\mathcal{Q}_\gamma[0,T]}^2\leq A_0+\sum_{k=1}^n\frac{A_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\gamma^k\right)+A_2\sum_{k=1}^n K_\gamma^k, \qquad (4.97)
$$

where A_0, A_1, and A_2 are some constants depending only on $\hat{L}$, T, d, and h. Summarizing the inequalities in (4.90)–(4.92), we know that

$$
P\left\{\|U^n\|_{[0,T]}\geq K\right\}\leq C_2\max\left\{\frac{\bar{K}}{K^2},\frac{1}{\bar{K}},\frac{1}{\bar{K}-\hat{L}^2T},\frac{1}{(\bar{K}-\hat{L}^2T)^2}\right\} \qquad (4.98)
$$

for some nonnegative constant C_2.

Now, for each Malliavin differentiable 4-tuple process $(X,V,\bar{V},\tilde{V})\in\mathcal{Q}_\gamma[0,T]$, we define

$$
\mathcal{D}_{\theta,z}(X,V,\bar{V},\tilde{V})\equiv(\mathcal{D}_{\theta,z}X,\mathcal{D}_{\theta,z}V,\mathcal{D}_{\theta,z}\bar{V},\mathcal{D}_{\theta,z}\tilde{V}) \qquad (4.99)
$$

for $t\in[0,T]$, $\theta\in[t,T]$, $z=0$ or $z\in\mathcal{Z}^{h-\bar{h}}\times\{0\}^{\bar{h}}$ with $\bar{h}<h$. To discuss the solutions of the corresponding Malliavin derivative based systems of FB-SDEs, we introduce some support spaces. More precisely, if $z=0$, this support space for a given $t\in[0,T]$ and a constant $\beta>0$ (which can be similarly chosen as in (4.81)) can be constructed as follows:

$$
\begin{aligned}
\mathcal{Q}_\beta[t,T]\equiv{}& D_{\mathcal{F}}^2([t,T],R^{p\times h})\times D_{\mathcal{F}}^2([t,T],R^{q\times h})\times D_{\mathcal{F},p}^2([t,T],R^{q\times d\times h})\\
&\times D_{\mathcal{F},p}^2([t,T]\times\mathcal{Z}^h,R^{q\times h\times h}), \qquad (4.100)
\end{aligned}
$$

where the norm endowed to this space for a process in (4.99) and $\theta\in[t,T]$ is given by

$$
\begin{aligned}
\left\|\mathcal{D}_0(X,V,\bar{V},\tilde{V})\right\|_{\mathcal{Q}_\beta[t,T]}^2\equiv{}& E\left[\sup_{\theta\in[t,T]}\left(\|\mathcal{D}_{\theta,0}X(t)\|^2+\|\mathcal{D}_{\theta,0}V(t)\|^2\right)e^{2\beta\theta}\right]\\
&+E\left[\int_t^T\left\|\mathcal{D}_{\theta,0}\bar{V}(t)\right\|^2e^{2\beta\theta}d\theta\right]\\
&+E\left[\int_t^T\left\|\mathcal{D}_{\theta,0}\tilde{V}(t,\cdot)\right\|_\nu^2e^{2\beta\theta}d\theta\right]. \qquad (4.101)
\end{aligned}
$$

On the other hand, corresponding to $z\neq0$, we need to define a support space with a norm related to double integral with respect to Lévy measure in $z\in\mathcal{Z}^h$. Specifically, this space, for a given $t\in[0,T]$ and a constant $\zeta>0$ is defined as

$$
\begin{aligned}
\mathcal{Q}_\zeta[t,T]\equiv{}& D_{\mathcal{F}}^2([t,T],R^{p\times h})\times D_{\mathcal{F}}^2([t,T],R^{q\times h})\times D_{\mathcal{F},p}^2([t,T],R^{q\times d\times h})\\
&\times D_{\mathcal{F},p}^2([t,T]\times\mathcal{Z}^h,R^{q\times h\times h}). \qquad (4.102)
\end{aligned}
$$

where the norm endowed to this space for $\theta\in[t,T]$ is given by

$$\begin{aligned}\left\|\mathcal{D}(X,V,\bar{V},\tilde{V})\right\|^2_{\mathcal{Q}_\zeta[t,T]} \equiv & E\left[\sup_{\theta\in[t,T]}\left(\|\mathcal{D}_{\theta,\cdot}X(t)\|^2_\nu+\|\mathcal{D}_{\theta,\cdot}V(t)\|^2_\nu\right)e^{2\zeta\theta}\right]\\ &+E\left[\int_t^T\|\mathcal{D}_{\theta,\cdot}\bar{V}(t)\|^2_\nu e^{2\zeta\theta}d\theta\right]\\ &+E\left[\int_t^T\left\|\left\|\mathcal{D}_{\theta,\cdot}\tilde{V}(t,\cdot)\right\|^2_\nu\right\|^2_\nu e^{2\zeta\theta}d\theta\right].\end{aligned}\tag{4.103}$$

Next, corresponding to the coupled Malliavin derivative based system of FB-SDEs in (4.58) and (4.69) with a possible solution of the form in (4.99), we let

$$\mathcal{D}_{\theta,z}(X,V,\bar{V},\tilde{V})^{v,r,c,n}\equiv(\mathcal{D}_{\theta,z}X^{v,r,c,n},\mathcal{D}_{\theta,z}V^{v,r,c,n},\mathcal{D}_{\theta,z}\bar{V}^{v,r,c,n},\mathcal{D}_{\theta,z}\tilde{V}^{v,r,c,n}).\tag{4.104}$$

In this case, the process $(X,V,\bar{V},\tilde{V})^{v,r,c,n}$ can be considered as a given random environment to the Malliavin derivative based system in (4.58) and (4.69). Furthermore, this coupled system is endowed with initial value $\Theta^{v,r,c,n}(t,z)$ in (4.60) and terminal condition $\Upsilon^{v,r,c,n}(t,z)$ in (4.72) and (4.73) respectively. However, the Malliavin derivatives $\mathcal{D}_{\theta,z}(RY^{v,r,c,n+1}(t))$ and $\mathcal{D}_{\theta,z}(SF^{v,r,c,n+1}(t))$ in (4.58) and (4.69) may be unbounded. Therefore, we truncate them with each $g\in\{1,2,...\}$ as follows,

$$\mathcal{D}_{\theta,z}(RY^{g,v,r,c,n+1}(t))\equiv g\bigwedge\mathcal{D}_{\theta,z}(RY^{v,r,c,n+1}(t)),\tag{4.105}$$

$$\mathcal{D}_{\theta,z}(SF^{g,v,r,c,n+1}(t))\equiv g\bigwedge\mathcal{D}_{\theta,z}(SF^{v,r,c,n+1}(t)),\tag{4.106}$$

$$\Theta^{g,v,r,c,n}(t,z)\equiv g\bigwedge\Theta^{v,r,c,n}(t,z),\tag{4.107}$$

$$\Upsilon^{g,v,r,c,n}(t,z)\equiv g\bigwedge\Upsilon^{v,r,c,n}(t,z),\tag{4.108}$$

$$(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}\equiv g\bigwedge(X,V,\bar{V},\tilde{V})^{v,r,c,n},\tag{4.109}$$

where the operator $\bigwedge$ denotes the smaller of two numbers and is interpreted in a componentwise way. Then, along $g\in\{1,2,...\}$, we have the following convergence either in the a.s. sample pathwise sense or in the mean-square sense due to the monotone convergence theorem: as $g\to\infty$,

$$\mathcal{D}_{\theta,z}(RY^{g,v,r,c,n+1}(t))\to\mathcal{D}_{\theta,z}(RY^{v,r,c,n+1}(t)),\tag{4.110}$$

$$\mathcal{D}_{\theta,z}(SF^{g,v,r,c,n+1}(t))\to\mathcal{D}_{\theta,z}(SF^{v,r,c,n+1}(t)),\tag{4.111}$$

$$\Theta^{g,v,r,c,n}(t,z)\to\Theta^{v,r,c,n}(t,z),\tag{4.112}$$

$$\Upsilon^{g,v,r,c,n}(t,z)\to\Upsilon^{v,r,c,n}(t,z),\tag{4.113}$$

$$(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}\to(X,V,\bar{V},\tilde{V})^{v,r,c,n}.\tag{4.114}$$

Thus, associated with (4.105)–(4.109), the counterpart of system (4.58) and (4.69) can be written as

$$\left\{\begin{aligned}\mathcal{D}_{\theta,z}X^{g,v,r,c,n+1}(t)&=\mathcal{D}_{\theta,z}X(0)+\mathcal{D}_{\theta,z}Z^{g,v,r,c,n}(t)+\mathcal{D}_{\theta,z}\left(RY^{g,v,r,c,n+1}(t)\right),\\ \mathcal{D}_{\theta,z}V^{g,v,r,c,n+1}(t)&=\mathcal{D}_{\theta,z}H(X^{g,v,r,c,n}(T),\cdot)+\mathcal{D}_{\theta,z}(SF^{g,v,r,c,n}(T))\\ &\quad+\mathcal{D}_{\theta,z}U^{g,v,r,c,n}(t)-\mathcal{D}_{\theta,z}(SF^{g,v,r,c,n+1}(t)).\end{aligned}\right.\tag{4.115}$$

Due to the truncation in (4.105)–(4.109), the system in (4.115) is not a real Malliavin derivative process oriented partial differential system. In other words, the original Malliavin derivative-based sequence indexed by (v,r,c,n) is replaced by a sequence indexed by (g,v,r,c,n), i.e.,

$$\mathcal{D}_{\theta,z}(X,V,\bar{V},\tilde{V})^{g,v,r,c,n},\tag{4.116}$$

which corresponds to the truncated values defined in (4.105)–(4.109). For example, by the corresponding discussion in (4.61), the coefficient in (4.115) has the form

$$\begin{aligned}&b(s^-,X^{g,v,r,c,n}+\mathcal{D}_{\theta,z}X^{g,v,r,c,n}e_h,V^{g,v,r,c,n}+\mathcal{D}_{\theta,z}V^{g,v,r,c,n}e_h,\\&\bar{V}^{g,v,r,c,n}+\mathcal{D}_{\theta,z}\bar{V}^{g,v,r,c,n}e_{dh},\tilde{V}^{g,v,r,c,n}+\mathcal{D}_{\theta,z}\tilde{V}^{g,v,r,c,n}e_{hh},*+\mathcal{D}_{\theta,z}*,u)\\&-b(s^-,X^{v,r,c,n},V^{v,r,c,n},\bar{V}^{v,r,c,n},\tilde{V}^{v,r,c,n},u).\end{aligned}\tag{4.117}$$

Thus, if we consider the sequence (4.116) as a solution to system (4.115) for each $n\in\{1,2,\ldots\}$, it follows from (4.117) that the sequence $(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}$ along $n\in\{1,2,\ldots\}$ is the given random environment. Therefore, for system (4.115), the binding Lipschitz process $\hat{L}^{g,v,r,c,n}(\omega)$ corresponding to its counterpart $\hat{L}(\omega)$ in conditions (2.10)–(2.16) is a functional of $(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}$, so $\hat{L}^{g,v,r,c,n}(\omega)=\hat{L}((X,V,\bar{V},\tilde{V})^{g,v,r,c,n},\omega)$. Due to the truncation in (4.109), it can be assumed to be a constant for each $g\in\{1,2,\ldots\}$.

In the subsequent discussion, we first fix the 4-tuple number (g,v,r,c) and consider system (4.115) with the initial value $\Theta^{g,v,r,c,n}(t,z)$ and terminal condition $\Upsilon^{g,v,r,c,n}(t,z)$ as a conventional system of FB-SDEs without reflection boundaries. Then, for the case $z=0$, it follows from Lemma 4.7 and the conditions in (2.10)–(2.16) that we can get inequalities analogous to those in (4.83)–(4.92) by suitably handling constants K and $\bar{K}$:

$$P\Big\{\|\mathcal{D}_0Z_1^{g,v,r,c,n}\|_{[t,T]}\geq K\Big\}\leq\frac{2\hat{L}^2T}{K-\hat{L}T}\Big\|\mathcal{D}_0(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}\Big\|^2_{\mathcal{Q}_\beta[t,T]},\tag{4.118}$$

$$P\Big\{\|\mathcal{D}_0Z_2^{g,v,r,c,n}\|_{[t,T]}\geq K\Big\}\leq\frac{\bar{K}}{K^2}+\frac{\hat{L}^2T}{\bar{K}-\hat{L}^2T}\Big\|\mathcal{D}_0(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}\Big\|^2_{\mathcal{Q}_\beta[t,T]},\tag{4.119}$$

$$P\Big\{\|\mathcal{D}_0U_1^{g,v,r,c,n}\|_{[t,T]}\geq K\Big\}\leq\frac{2\hat{L}^2T}{(K-\hat{L}T)^2}\Big\|\mathcal{D}_0(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}\Big\|^2_{\mathcal{Q}_\beta[t,T]},\tag{4.120}$$

$$P\Big\{\|\mathcal{D}_0U_2^{g,v,r,c,n}\|_{[t,T]}\geq K\Big\}\leq\frac{\bar{K}}{K^2}+\frac{\hat{L}^2T}{\bar{K}-\hat{L}^2T}\Big\|\mathcal{D}_0(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}\Big\|^2_{\mathcal{Q}_\beta[t,T]},\tag{4.121}$$

$$P\Big\{\|\mathcal{D}_0U_3^{g,v,r,c,n}\|_{[t,T]}\geq K\Big\}\leq\frac{\bar{K}}{K^2}+\frac{1}{\bar{K}}\Big\|\mathcal{D}_0(X,V,\bar{V},\tilde{V})^{g,v,r,c,n+1}\Big\|^2_{\mathcal{Q}_\beta[t,T]}.\tag{4.122}$$

Furthermore, it follows from the facts in (4.118)–(4.122) and a similar argument as was used for proving inequality (4.98) that

$$P\Big\{\|\mathcal{D}_0U^{g,v,r,c,n}(t)\|_{[t,T]}\geq K\Big\}\leq C_3\max\left\{\frac{\bar{K}}{K^2},\frac{1}{\bar{K}},\frac{1}{\bar{K}-\hat{L}^2T},\frac{1}{(\bar{K}-\hat{L}^2T)^2}\right\}.\tag{4.123}$$

Similarly, for the case in which $z\neq0$, it follows from Lemma 4.7 and the conditions in (2.10)–(2.16) that

$$P\Big\{(\|\mathcal{D}Z_1^{g,v,r,c,n}\|_\nu)_{[t,T]}\geq K\Big\}\leq\frac{2\hat{L}^2T}{K-\hat{L}T}\Big\|\mathcal{D}(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}\Big\|^2_{\mathcal{Q}_\varsigma[t,T]},\tag{4.124}$$

$$P\Big\{(\|\mathcal{D}Z_2^{g,v,r,c,n}\|_\nu)_{[t,T]}\geq K\Big\}\leq\frac{\bar{K}}{K^2}+\frac{\hat{L}^2T}{\bar{K}-\hat{L}^2T}\Big\|\mathcal{D}(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}\Big\|^2_{\mathcal{Q}_\varsigma[t,T]},\tag{4.125}$$

$$P\Big\{(\|\mathcal{D}U_1^{g,v,r,c,n}\|_\nu)_{[t,T]}\geq K\Big\}\leq\frac{2\hat{L}^2T}{(K-\hat{L}T)^2}\Big\|\mathcal{D}(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}\Big\|^2_{\mathcal{Q}_\varsigma[t,T]},\tag{4.126}$$

$$P\Big\{(\|\mathcal{D}U_2^{g,v,r,c,n}\|_\nu)_{[t,T]}\geq K\Big\}\leq\frac{\bar{K}}{K^2}+\frac{\hat{L}^2T}{\bar{K}-\hat{L}^2T}\Big\|\mathcal{D}(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}\Big\|^2_{\mathcal{Q}_\varsigma[t,T]},\tag{4.127}$$

$$P\Big\{(\|\mathcal{D}U_3^{g,v,r,c,n}\|_\nu)_{[t,T]}\geq K\Big\}\leq\frac{\bar{K}}{K^2}+\frac{1}{\bar{K}}\Big\|\mathcal{D}(X,V,\bar{V},\tilde{V})^{g,v,r,c,n+1}\Big\|^2_{\mathcal{Q}_\varsigma[t,T]}.\tag{4.128}$$

Therefore, by (4.87) and (4.124)–(4.128), we have that

$$
\begin{aligned}
&P\Big\{\|\mathcal{D}.U^{g,v,r,c,n}(t)\|_{[t,T]}\geq K\Big\}\\
&\leq P\Big\{(\|\mathcal{D}U^{g,v,r,c,n}(t)\|_{\nu})_{[t,T]}\geq K\Big\}\\
&\leq C_4\max\left\{\frac{\bar{K}}{K^2},\frac{1}{\bar{K}},\frac{1}{\bar{K}-\hat{L}^2T},\frac{1}{(\bar{K}-\hat{L}^2T)^2}\right\}.
\end{aligned}
\tag{4.129}
$$

Here, we remark that the nonnegative constants C_3 and C_4 in (4.123) and (4.129) may depend on (g,v,r,c). However, owing to the convergence in (4.52)–(4.54), (4.110)–(4.114) and the similar computational procedure for a conventional system of FB-SDEs (see in (4.115) with $z=0$), we can take (g,v,r,c) as a 4-tuple integer function satisfying $(g(v),v(r),r(c),c(n))\to\infty$ when $n\to\infty$ such that

$$
\left\|\Delta(\mathcal{D}_0(X,V,\bar{V},\tilde{V})^{g,v,r,c,n})\right\|^2_{\mathcal{Q}_\beta[t,T]}\to 0 \text{ as } n\to\infty,
\tag{4.130}
$$

where the difference Δ is defined by

$$
\Delta(\mathcal{D}_{\theta,z}(X,V,\bar{V},\tilde{V})^{g,v,r,c,n})=\mathcal{D}_{\theta,z}(X,V,\bar{V},\tilde{V})^{g,v,r,c,n}-\mathcal{D}_{\theta,z}(X,V,\bar{V},\tilde{V})^{v,r,c,n}.
\tag{4.131}
$$

Moreover, once again by the convergence in (4.52)-(4.54) and (4.110)-(4.114), we have that

$$
\left\|((\Delta X^{v,r,c,n},\Delta V^{v,r,c,n}),(\Delta\bar{V}^{v,r,c,n},\Delta\tilde{V}^{v,r,c,n}(\cdot)))\right\|^2_{\mathcal{Q}_\gamma[0,T]}\to 0 \text{ as } n\to\infty,
\tag{4.132}
$$

where the 3-tuple function satisfies $(v(r),r(c),c(n))\to\infty$ as $n\to\infty$. In addition, the notation Δ in (4.132) denotes the difference between the corresponding processes of the system in (1.1) and the one in (4.63) and (4.70) (see those as defined in (4.85)-(4.86)).

Hence, for each given $\epsilon>0$, it follows from the initial condition in (4.77) and the facts in equations (4.98), (4.123), (4.129), (4.130), and (4.132) with suitably chosen constants K and $\bar{K}$, that

$$
\begin{aligned}
&\inf_n P\Big\{\|\Xi^n(t)\|\leq C, 0\leq t\leq T\Big\}\\
&\geq\inf_n\min\Big\{P\Big\{\big\|(X^{n+1},Y^{n+1})(t)\big\|\leq C, 0\leq t\leq T\Big\},\\
&\qquad P\Big\{\Big\|(V^{n+1},\bar{V}^{n+1},\tilde{V}^{n+1}(\cdot),F^{n+1})(t)\Big\|\leq C, 0\leq t\leq T\Big\}\Big\}\\
&\geq 1-\epsilon
\end{aligned}
\tag{4.133}
$$

for some nonnegative constant C. Thus, the sequence $\{\Xi^n\}$ along $n\in\{1,2,\ldots\}$ satisfies condition (4.79).

Now, we show this sequence $\{\Xi^n\}$ satisfies condition (4.80). In doing so, for any $\epsilon>0$ and a constant $\delta>0$, we consider a finite set $\{t_l\}$ of points satisfying $0=t_0<t_1<\ldots<t_p=T$ and $t_l-t_{l-1}=\delta<\epsilon/\hat{L}$ with $l\in\{1,\ldots p\}$ and $p\in\{1,2,\ldots\}$. Furthermore, for all $0\leq s\leq t\leq T$, it follows from (4.83) and (4.84) that

$$
\big\|(X^{n+1},Y^{n+1})(t)-(X^{n+1},Y^{n+1})(s)\big\|\leq 2C_1\left(\|X(0)\|+\|Z^n\|_{[0,T]}\right),
\tag{4.134}
$$

$$
\begin{aligned}
&\Big\|(V^{n+1},\bar{V}^{n+1},\tilde{V}^{n+1}(\cdot),F^{n+1})(t)-(V^{n+1},\bar{V}^{n+1},\tilde{V}^{n+1}(\cdot),F^{n+1})(s)\Big\|\\
&\leq 2C_1\Big(1+\|X(0)\|+\big\|Z^{n-1}\big\|_{[0,T]}+\big\|U^{n-1}\big\|_{[0,T]}+\|U^n\|_{[0,T]}\\
&\qquad+\big\|\Delta\bar{V}^{v,r,c,n+1}(t)\big\|+\Big\|\Delta\tilde{V}^{v,r,c,n+1}(\cdot)(t)\Big\|+\|\mathcal{D}_{t,0}U^{v,r,c,n}(t)\|+\|\mathcal{D}_{t,\cdot}U^{v,r,c,n}(t)\|\Big).
\end{aligned}
\tag{4.135}
$$

Then, by equations (4.77), (4.93)–(4.96), (4.134)–(4.135), and the explanation given for (4.83), we know that

$$\begin{aligned}
&P\Big\{w(Z_1^n,\delta,T)\geq\epsilon\Big\}\\
&\leq\frac{3\hat{L}^2\delta}{(\epsilon-\hat{L}\delta)^2}\left(E\left[\|X^n\|^2_{[0,T]}+\|V^n\|^2_{[0,T]}\right]+\left\|(U^n,\bar{V}^n,\tilde{V}^n)\right\|^2_{\mathcal{Q}_\gamma[0,T]}\right)\\
&\leq\frac{3\hat{L}^2\delta}{(\epsilon-\hat{L}\delta)^2}\left(A_0+\sum_{k=1}^{n}\frac{A_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\gamma^k\right)+A_2\sum_{k=1}^{n}K_\gamma^k\right),
\end{aligned}\tag{4.136}$$

where as stated for the inequality in (4.97), A_0, A_1, and A_2 are nonnegative constants depending only on $\hat{L}$, T, d, and h. By Lemma 4.2.8 [2] (or the related theorem in [22]) and the linear growth condition, we know that

$$\begin{aligned}
&P\Big\{w(Z_2^n,\delta,T)\geq\epsilon\Big\}\\
&\leq\frac{\bar{\epsilon}}{\epsilon^2}+\frac{3\hat{L}^2}{\bar{\epsilon}-3\hat{L}^2\delta}\left(\delta E\left[\|X^n\|^2_T\right]+\delta E\left[\|V^n\|^2_T\right]+E\left[\left\|(U^n,\bar{V}^n,\tilde{V}^n)\right\|^2_{\mathcal{Q}_\gamma[0,T]}\right]\right)\\
&\leq\frac{\bar{\epsilon}}{\epsilon^2}+\frac{3\hat{L}^2}{\bar{\epsilon}-3\hat{L}^2\delta}\left(\delta\left(A_0+\sum_{k=1}^{n}\frac{A_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\gamma^k\right)+A_2\sum_{k=1}^{n}K_\gamma^k\right)+A_3\sum_{k=1}^{n}K_\gamma^k\right)
\end{aligned}\tag{4.137}$$

for each nonnegative constant $\bar{\epsilon}>3\hat{L}^2\delta$, where A_3 is some nonnegative constant depending only on $\hat{L}$, T, d, and h. Similarly, there are some constants B_0, B_1, B_2, and B_3 depending only on $\hat{L}$, T, d, and h such that

$$\begin{aligned}
&P\Big\{w(U_1^n,\delta,T)\geq\epsilon\Big\}\\
&\leq\frac{3\hat{L}^2\delta}{(\epsilon-\hat{L}\delta)^2}\left(B_0+\sum_{k=1}^{n}\frac{B_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\gamma^k\right)+B_2\sum_{k=1}^{n}K_\gamma^k\right),
\end{aligned}\tag{4.138}$$

$$\begin{aligned}
&P\Big\{w(U_2^n,\delta,T)\geq\epsilon\Big\}\\
&\leq\frac{\bar{\epsilon}}{\epsilon^2}+\frac{3\hat{L}^2}{\bar{\epsilon}-3\hat{L}^2\delta}\left(\delta\left(B_0+\sum_{k=1}^{n}\frac{B_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\gamma^k\right)+B_2\sum_{k=1}^{n}K_\gamma^k\right)+B_3\sum_{k=1}^{n}K_\gamma^k\right),
\end{aligned}\tag{4.139}$$

$$\begin{aligned}
&P\Big\{w(U_3^n,\delta,T)\geq\epsilon\Big\}\\
&\leq\frac{\bar{\epsilon}}{\epsilon^2}+\frac{3\hat{L}^2}{\bar{\epsilon}}\left(\delta\left(B_0+\sum_{k=1}^{n+1}\frac{B_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\gamma^k\right)+B_2\sum_{k=1}^{n}K_\gamma^k\right)+B_3\sum_{k=1}^{n+1}K_\gamma^k\right).
\end{aligned}\tag{4.140}$$

Next, instead of the interval $[0,T]$ consider the interval $[t,T]$ in the left-hand side of (4.2), and just as in the discussion of (4.136)–(4.140) there are nonnegative constants $\bar{A}_0$, $\bar{A}_1$, $\bar{A}_2$,

$\bar{B}_0$, $\bar{B}_1$, $\bar{B}_2$, and $\bar{B}_3$ depending only on $\hat{L}$, T, d, and h such that

$$
\begin{aligned}
&P\Big\{w(\mathcal{D}_0 Z_1^{g,v,r,c,n},\delta,[t,T])\geq\epsilon\Big\} \\
&\leq \frac{3\hat{L}^2\delta}{(\epsilon-\hat{L}\delta)^2}\left(\bar{A}_0+\sum_{k=1}^{n}\frac{\bar{A}_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\beta^k\right)+\bar{A}_2\sum_{k=1}^{n}K_\beta^k\right), \qquad (4.141)\\
&P\Big\{w(\mathcal{D}_0 Z_2^{g,v,r,c,n},\delta,[t,T])\geq\epsilon\Big\} \\
&\leq \frac{\bar{\epsilon}}{\epsilon^2}+\frac{3\hat{L}^2}{\bar{\epsilon}-3\hat{L}^2\delta}\left(\delta\left(\bar{A}_0+\sum_{k=1}^{n}\frac{\bar{A}_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\beta^k\right)+\bar{A}_2\sum_{k=1}^{n}K_\beta^k\right)+\bar{A}_3\sum_{k=1}^{n}K_\beta^k\right), \qquad (4.142)\\
&P\Big\{w(\mathcal{D}_0 U_1^{g,v,r,c,n},\delta,[t,T])\geq\epsilon\Big\} \\
&\leq \frac{3\hat{L}^2\delta}{(\epsilon-\hat{L}\delta)^2}\left(\bar{B}_0+\sum_{k=1}^{n}\frac{\bar{B}_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\beta^k\right)+\bar{B}_2\sum_{k=1}^{n}K_\beta^k\right), \qquad (4.143)\\
&P\Big\{w(\mathcal{D}_0 U_2^{g,v,r,c,n},\delta,[t,T])\geq\epsilon\Big\} \\
&\leq \frac{\bar{\epsilon}}{\epsilon^2}+\frac{3\hat{L}^2}{\bar{\epsilon}-3\hat{L}^2\delta}\left(\delta\left(\bar{B}_0+\sum_{k=1}^{n}\frac{\bar{B}_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\beta^k\right)+\bar{B}_2\sum_{k=1}^{n}K_\beta^k\right)+\bar{B}_3\sum_{k=1}^{n}K_\beta^k\right), \qquad (4.144)\\
&P\Big\{w(\mathcal{D}_0 U_3^{g,v,r,c,n},\delta,[t,T])\geq\epsilon\Big\} \\
&\leq \frac{\bar{\epsilon}}{\epsilon^2}+\frac{3\hat{L}^2}{\bar{\epsilon}}\left(\delta\left(\bar{B}_0+\sum_{k=1}^{n+1}\frac{\bar{B}_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\beta^k\right)+\bar{B}_2\sum_{k=1}^{n}K_\beta^k\right)+\bar{B}_3\sum_{k=1}^{n+1}K_\beta^k\right), \qquad (4.145)
\end{aligned}
$$

where as in (4.94), there exists a nonnegative constant $\tilde{K}$ depending only on $\hat{L}$, T, d, and h, such that, for a suitable chosen $\beta>0$,

$$K_\beta=\frac{\tilde{K}}{\beta}<1. \qquad (4.146)$$

Furthermore, there are nonnegative constants $\tilde{A}_0$, $\tilde{A}_1$, $\tilde{A}_2$, $\tilde{B}_0$, $\tilde{B}_1$, $\tilde{B}_2$, and $\tilde{B}_3$ depending only on $\hat{L}$, T, d, and h such that

$$
\begin{aligned}
&P\Big\{\|w(\mathcal{D}Z_1^{g,v,r,c,n},\delta,[t,T])\|_\nu \geq \epsilon\Big\}\\
&\leq \frac{3\hat{L}^2\delta}{(\epsilon-\hat{L}\delta)^2}\left(\tilde{A}_0+\sum_{k=1}^{n}\frac{\tilde{A}_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\zeta^k\right)+\tilde{A}_2\sum_{k=1}^{n}K_\zeta^k\right), \qquad (4.147)\\
&P\Big\{(\|w(\mathcal{D}Z_2^{g,v,r,c,n},\delta,[t,T])\|_\nu)\geq \epsilon\Big\}\\
&\leq \frac{\bar{\epsilon}}{\epsilon^2}+\frac{3\hat{L}^2}{\bar{\epsilon}-3\hat{L}^2\delta}\left(\delta\left(\tilde{A}_0+\sum_{k=1}^{n}\frac{\tilde{A}_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\zeta^k\right)+\tilde{A}_2\sum_{k=1}^{n}K_\zeta^k\right)+\tilde{A}_3\sum_{k=1}^{n}K_\zeta^k\right), \qquad (4.148)\\
&P\Big\{(\|w(\mathcal{D}U_1^{g,v,r,c,n},\delta,[t,T])\|_\nu \geq \epsilon\Big\}\\
&\leq \frac{3\hat{L}^2\delta}{(\epsilon-\hat{L}\delta)^2}\left(\tilde{B}_0+\sum_{k=1}^{n}\frac{\tilde{B}_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\zeta^k\right)+\tilde{B}_2\sum_{k=1}^{n}K_\zeta^k\right), \qquad (4.149)\\
&P\Big\{(\|w(\mathcal{D}U_2^{g,v,r,c,n},\delta,[t,T])\|_\nu \geq \epsilon\Big\}\\
&\leq \frac{\bar{\epsilon}}{\epsilon^2}+\frac{3\hat{L}^2}{\bar{\epsilon}-3\hat{L}^2\delta}\left(\delta\left(\tilde{B}_0+\sum_{k=1}^{n}\frac{\tilde{B}_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\zeta^k\right)+\tilde{B}_2\sum_{k=1}^{n}K_\zeta^k\right)+\tilde{B}_3\sum_{k=1}^{n}K_\zeta^k\right), \qquad (4.150)\\
&P\Big\{(\|w(\mathcal{D}U_3^{g,v,r,c,n},\delta,[t,T])\|_\nu)\geq \epsilon\Big\}\\
&\leq \frac{\bar{\epsilon}}{\epsilon^2}+\frac{3\hat{L}^2}{\bar{\epsilon}}\left(\delta\left(\tilde{B}_0+\sum_{k=1}^{n+1}\frac{\tilde{B}_1^{k+1}T^{k+1}}{(k+1)!}\left(1+K_\zeta^k\right)+\tilde{B}_2\sum_{k=1}^{n}K_\zeta^k\right)+\tilde{B}_3\sum_{k=1}^{n+1}K_\zeta^k\right), \qquad (4.151)
\end{aligned}
$$

where as in (4.146) and for a suitable chosen $\zeta>0$,

$$K_\zeta=\frac{\tilde{K}}{\zeta}<1. \qquad (4.152)$$

Hence, for each given $\epsilon>0$, it follows from the convergence in (4.130) and (4.132), the facts in (4.136)–(4.145) and (4.147)–(4.151), there exist suitably chosen constants $\bar{\epsilon}$, δ, sufficiently small numbers of γ, β, and ζ (via the expressions of K_γ, K_β, and K_ζ in (4.94), (4.146), and (4.152)) that

$$\limsup_{n\to\infty}P\Big\{w(\Xi^n),\delta,T)\geq\epsilon\Big\}\leq\epsilon. \qquad (4.153)$$

Thus condition (4.80) is true for the sequence of $\{\Xi^n\}$. Furthermore, by (4.84), (4.153), and Corollary 7.4 in [20], this sequence is relatively compact. Therefore, there is a subsequence of $\{\Xi^n\}$ that converges weakly in Skorohod topology to

$$\Xi\equiv((X,Y),(V,\bar{V},\hat{\tilde{V}},F)). \qquad (4.154)$$

For convenience, we suppose that the subsequence is the sequence itself, i.e.,

$$\Xi^n\Rightarrow\Xi. \qquad (4.155)$$

Due to the Skorohod representation theorem (see Theorem 1.8 in [20]), we can assume that the convergence in (4.155) is a.s. in the Skorohod topology. Now, for each $n\in\{1,2,\ldots\}$, we define

$$
\begin{aligned}
&\Upsilon_{ji}(t^-,X^n,V^n,\bar{V}^n,\tilde{V}^n,u,z_i)\\
&\equiv\zeta_{ji}(t^-,X^n,V^n,\bar{V}^n,\tilde{V}^n,u,z_i)+\tilde{V}_{ji}^{n+1}(t^-,z_i)-\tilde{V}_{ji}^n(t^-,z_i). \qquad (4.156)
\end{aligned}
$$

It follows from (4.63) and the Itô-Lévy isometry that

$$E\Big[\int_0^T\int_{\mathcal{Z}}\sum_{j=1}^{q}\sum_{i=1}^{h}\hat{K}\bigwedge\Big(\Upsilon_{ji}(t^-,X^m,V^m,\bar{V}^m,\tilde{V}^n,u,z_i)\\ -\Upsilon_{ji}(t^-,X^n,V^n,\bar{V}^n,\tilde{V}^n,u,z_i)\Big)^2\lambda_i\nu_i(dz_i)dt\Big]\to 0 \tag{4.157}$$

as $m,n\to\infty$ for each given constant $\hat{K}>0$. Thus, subject to the constraint of $\hat{K}$, we see that $\Upsilon(t^-,X^n,V^n,\bar{V}^n,\tilde{V}^n,u,z)$ is a Cauchy sequence along $n\in\{1,2,...\}$ in the mean-square sense. Furthermore, as $m,n\to\infty$, it follows from the Lipschitz condition in (2.13) that

$$E\Big[\int_0^T\int_{\mathcal{Z}}\sum_{j=1}^{q}\sum_{i=1}^{h}\hat{K}\bigwedge\Big(\tilde{V}^{m+1}(t^-,z_i)-\tilde{V}^{n+1}(t^-,z_i)\Big)^2\lambda_i\nu_i(dz_i)dt\Big]\to 0. \tag{4.158}$$

In other words, $\tilde{V}^{n+1}(t^-,z)$ is also a Cauchy sequence along $n\in\{1,2,...\}$ in the mean-square sense. Therefore, it has a convergent sequence in the mean-square sense, which further implies that it has an a.s. convergent subsequence. Hence, we can conclude that $\tilde{V}^{n+1}$ corresponding to the subsequence converges a.s. along $n\in\{1,2,...\}$ to a limit $\tilde{V}$ such that $\hat{\tilde{V}}$ in (4.154) can be explicitly expressed as

$$\hat{\tilde{V}}(t)=\tilde{V}(t,\cdot)=\int_{\mathcal{Z}^m}\tilde{V}(t,z)\lambda\nu(dz)$$

under the constraint of $\hat{K}$. Then, let $\hat{K}\to\infty$, it follows from the monotone convergence theorem that the required $\tilde{V}(t,z)$ can be derived.

Finally, by claim (a) in Theorem 1.14 (or claim (a) in Proposition 2.1) of [26] and the fact that $Y^{n+1}(0)=0$ and Y^{n+1} is nondecreasing, we can conclude that $Y(0)=0$ and Y is nondecreasing. Furthermore, by Lemma 4.3 and (4.50)

$$\int_0^t I_{D_i}(X(s))dY_i(s)=Y_i(t)\ \text{ for all } t\geq 0, i\in\{1,...,b\}. \tag{4.159}$$

Analogously, we know that $F(0)=0$, F is non-decreasing, and

$$\int_0^t I_{\bar{D}_i}(V(s))dF_i(s)=F_i(t)\ \text{ for all } t\geq 0, i\in\{1,...,\bar{b}\}. \tag{4.160}$$

Therefore, by the Lipschitz condition in (2.11), we know that $((X,Y),(V,\bar{V},\tilde{V},F))$ satisfies the FB-SDEs with Lévy jumps in (1.1) a.s. Thus, by the Skorohod representation theorem again, it is a weak solution to the FB-SDEs in (1.1)–(1.5).

Part A (Uniqueness). Assume that $((X^j,Y^j),(V^j,\bar{V}^j,\tilde{V}^j,F^j))$ for $j\in\{1,2\}$ are two weak solutions to the FB-SDEs in (1.1). Since Y_i^j for each $i\in\{1,...,b\}$ and $j\in\{1,2\}$ is non-decreasing and finite a.s. along each sample path, it can be approximated by a sequence of non-decreasing continuous processes. Therefore, without loss of generality, we suppose that Y_i^j for each $i\in\{1,...,b\}$ and $j\in\{1,2\}$ is non-decreasing and continuous. Furthermore, it follows from the discussion in (4.50) that

$$Y_i^j(t)=\int_0^t\alpha_i^j(s)ds \tag{4.161}$$

for some process $\alpha_i^j(\cdot)\geq 0$. Nevertheless as in (4.50), it may be unbounded. Thus, for each $c\in\{1,2,...\}$, we let $\alpha^{c,j}(t)$ be a b-dimensional vector whose component associated with each $i\in\{1,...,b\}$ is given by

$$\alpha_i^{c,j}(t)\equiv\alpha_i^j(t)I\Big\{\alpha_i^j(t)\leq c\Big\}. \tag{4.162}$$

Then, by the monotone convergence theorem we have that

$$\left\|Y^j - Y^{c,j}\right\|_{[0,T]} \to 0 \text{ a.s. as } c \to \infty. \tag{4.163}$$

In addition, for each given c, there exists a sequence $\{\alpha^{r,c,j}, r \in \{1,2,\ldots\}\}$ of simple processes such that

$$E\left[\int_0^T \left\|\alpha^{r,c,j}(s) - \alpha^{c,j}(s)\right\|^2 ds\right] \to 0 \text{ as } r \to \infty. \tag{4.164}$$

For convenience define

$$Y^{r,c,j}(t) \equiv \int_0^t \alpha^{r,c,j}(s)ds. \tag{4.165}$$

Similarly, we can find such a sequence $\{\beta^{r,c,j}\}$ corresponding to $F(t)$. such that

$$F^{r,c,j}(t) \equiv \int_0^t \beta^{r,c,j}(s)ds. \tag{4.166}$$

Therefore, for each $j \in \{1,2\}$, the corresponding FB-SDEs in (1.1) can be rewritten as

$$\begin{cases} X^j(t) = \xi + Z^j(t) + R(Y^j(t) - Y^{r,c,j}(t)) + RY^{r,c,j}(t), \\ V^j(t) = H(X^j(T),\cdot) + U^j(t) + S\left((F^j(T) - F^j(t)) - (F^{r,c,j}(T) - F^{r,c,j}(t))\right) \\ \qquad +S(F^{r,c,j}(T) - F^{r,c,j}(t)). \end{cases} \tag{4.167}$$

Next, for each $f^j \in \{X^j, Y^{r,c,j}, V^j, \bar{V}^j, \tilde{V}^j, F^{r,c,j}\}$ with $j \in \{1,2\}$, we define

$$\Delta f = f^1 - f^2 \tag{4.168}$$

and construct the following quadratic function

$$\zeta(t) \equiv (\mathrm{Tr}(\Delta X(t)) + \mathrm{Tr}(\Delta V(t)))\, e^{2\gamma t} \tag{4.169}$$

where $\gamma > 0$ is some constant, and $\mathrm{Tr}(A)$ denotes the trace of the matrix $A'A$ for a given matrix A. Then, by the expressions in (4.161) and (4.165)–(4.166), the Itô's formula, and the discussion for Proposition 18 in [16] and Lemma 4.1, we have that

$$\begin{aligned} &\left\|((\Delta X, \Delta\alpha^{r,c}), (\Delta V, \Delta\bar{V}, \Delta\tilde{V}, \Delta\beta^{r,c}))\right\|^2_{\mathcal{Q}_\gamma[0,T]} \\ &\le \frac{\bar{C}}{1-\bar{K}_\gamma}\left(E\left[\sup_{t\in[0,T]}\|\Delta(Y(t) - Y^{r,c}(t))\|^2\right] + E\left[\sup_{t\in[0,T]}\|\Delta(F(t) - F^{r,c}(t))\|^2\right]\right) \\ &\to 0 \text{ as } c \to \infty, \end{aligned} \tag{4.170}$$

where $\bar{C}$ and $\bar{K}_\gamma < 1$ are nonnegative constants only depending on $\hat{L}$, T, d, and h. Furthermore, r in (4.170) is a function of c, satisfying $r(c) \to \infty$ as $c \to \infty$. In addition, the norm (4.170) is defined as

$$\begin{aligned} &\left\|((\Delta X, \Delta\alpha^{r,c}), (\Delta V, \Delta\bar{V}, \Delta\tilde{V}, \Delta\beta^{r,c}))\right\|^2_{\mathcal{Q}_\gamma[0,T]} \\ &\equiv E\left[\sup_{t\in[0,T]}\left(\|\Delta X(t)\|^2 + \|\Delta V(t)\|^2\right)e^{2\gamma t}\right] \\ &\quad + E\left[\int_0^T \|\Delta\alpha^{r,c}(t)\|^2 e^{2\gamma t}dt\right] + E\left[\int_0^T \|\Delta\beta^{r,c}(t)\|^2 e^{2\gamma t}dt\right] \\ &\quad + E\left[\int_0^T \|\Delta\bar{V}(t)\|^2 e^{2\gamma t}dt\right] + E\left[\int_0^T \left\|\Delta\tilde{V}(t,\cdot)\right\|^2_\nu e^{2\gamma t}dt\right]. \end{aligned} \tag{4.171}$$

Thus, it follows from (4.170) and (4.171) that

$$(\Delta X,\Delta\alpha^{r,c},\Delta V,\Delta\bar{V},\Delta\tilde{V},\Delta\beta^{r,c})\to 0 \text{ as } r\to\infty \tag{4.172}$$

for almost all $(t,z,\omega)\in[0,T]\times\mathcal{Z}^m\times\Omega$. By the same argument used in proving relatively compactness for the sequence in (4.77), we know that the sequence $\{(\Delta X,\Delta\alpha^{r,c},\ \Delta V,\Delta\bar{V},\ \Delta\tilde{V},\ \Delta\beta^{r,c}),c\in\{1,2,...\}\}$ is relatively compact in the Skorohod topology. It follows from (4.172) that the zero process is the unique limit process for all of its convergent subsequences. Thus, we can conclude that the whole sequence itself converges to the zero process weakly, which implies that

$$(\Delta X,\Delta Y^{r,c},\Delta V,\Delta\bar{V},\Delta\tilde{V},\Delta F^{r,c})\Rightarrow 0 \text{ as } r\to\infty, \tag{4.173}$$

where "$\Rightarrow$" denotes "convergence in distribution". Therefore, it follows from (4.163)–(4.166) and Theorem 4.4.6 in [7] that

$$\begin{aligned}
&((X^1,Y^1),(V^1,\bar{V}^1,\tilde{V}^1,F^1))\\
&=((X^1,Y^1),(V^1,\bar{V}^1,\tilde{V}^1,F^1))-((X^1,Y^{r,c,1}),(V^1,\bar{V}^1,\tilde{V}^1,F^{r,c,1}))\\
&\quad+((X^2,Y^{r,c,2}),(V^2,\bar{V}^2,\tilde{V}^2,F^{r,c,2}))-((X^2,Y^2),(V^2,\bar{V}^2,\tilde{V}^2,F^2))\\
&\quad+((\Delta X,\Delta Y^{r,c}),(\Delta V,\Delta\bar{V},\Delta\tilde{V},\Delta F^{r,c}))+((X^2,Y^2),(V^2,\bar{V}^2,\tilde{V}^2,F^2))\\
&\Rightarrow((X^2,Y^2),(V^2,\bar{V}^2,\tilde{V}^2,F^2)) \text{ as } r\to\infty.
\end{aligned} \tag{4.174}$$

Thus, we know that $(X^j,Y^j,\ V^j,\bar{V}^j,\ \tilde{V}^j,F^j)$ with $j\in\{1,2\}$ have the same distribution. In other words, the weak uniqueness of solution to the system of FB-SDEs in (1.1) holds. This completes the proof of Part A.

Part B. We consider the case that $\hat{L}(t,\omega)$ appearing in (2.10)–(2.15) is a constant and the spectral radii of S and each $p\times p$ sub-principal matrix of $N'R$ are strictly less than one. In this case, we need to prove that there is a unique strong adapted solution $((X,Y),(V,\bar{V},\tilde{V},F))$ to system (1.1)–(1.5).

In fact, it follows from [24], [15], Lemma 7.1 and Theorem 7.2 in [6] that there exist two Lipschitz continuous mappings Φ_1 and Ψ_1 such that

$$(X^{n+1},Y^{n+1})(t)=(\bar{Z}^n(t)+\Phi_1(\bar{Z}^n)(t),\Phi_1(\bar{Z}^n)(t)), \tag{4.175}$$
$$(V^{n+1},F^{n+1})(t)=(\bar{U}^n(t)+\Psi_1(\bar{U}^n)(t),\Psi_1(\bar{U}^n)(t)) \tag{4.176}$$

for each $n\in\{1,2,...\}$ and $t\in[0,T]$, where the processes $\bar{Z}^n$ and $\bar{U}^n$ are defined by

$$\begin{aligned}
&\bar{Z}^n(t)=\xi+Z^n(t),\\
&\bar{U}^n(t)=H(X^n(T),*)+SF^n(T)+U^n(t).
\end{aligned}$$

Then, it follows from (4.175)–(4.176), the related estimates in Part A, and the conventional Picard's iterative method, that we can prove the claim in terms of the unique existence of a strong solution to (1.1)–(1.5) in Part B. Furthermore, we know that there are two Lipschitz continuous mappings Φ and Ψ such that

$$\begin{aligned}
&\Phi(\bar{Z})(t)=\Phi_1(\bar{Z})(t),\\
&\Psi(\bar{U})(t)=F(T)-F(t).
\end{aligned}$$

Hence, we finish the proof of Part B.

Part C. We consider the case that $\hat{L}(t,\omega)$ appearing in (2.10)–(2.15) is a constant and both of the SDEs have no reflection boundaries. In this case, we need to prove that there is a unique strong adapted solution $((X,Y),(V,\bar{V},\tilde{V},F))$ to system (1.1)–(1.5). In fact, by the related estimates in Part A, this case can be proved by directly generalizing the conventional Picard's iterative method.

Part D. We consider the case that $\hat{L}(t,\omega)$ appearing in (2.10)–(2.15) is a general adapted and mean-square integrable stochastic process. The proofs corresponding to the cases stated in Part A, Part B, and Part C can be accomplished along the lines of proofs for Lemma 4.1 in [14] associated with a forward SDE under random environment and Proposition 18 in [16] for a backward SDE under random environment. The key in the proofs is to introduce the following sequence of $\{\mathcal{F}_t\}$-stopping times, i.e.,

$$\tau_n \equiv \inf\{t>0, \|\hat{L}(t)\|>n\} \text{ for each } n\in\{1,2,\ldots\}.$$

By the assumption in (2.16), τ_n is nondecreasing and a.s. tends to infinity as $n\to\infty$.

Finally, by summarizing the cases presented in Part A to Part D, we finish the proof of Theorem 2.1.

4.4. Proof of Proposition 3.1. First, we claim that the stochastic exponential M defined in Proposition 3.1 is an $\{\mathcal{F}_t\}$- and P-local martingale. In fact, it follows from Theorem 2.1 that the system described by the SDEs in (3.1)–(3.2) is well-posed, i.e., the joint distribution $F_{(S,Z)}$ of the processes S and Z is uniquely determined. Thus, we can conclude that

$$\begin{aligned} E\Bigg[\int_0^T \Bigg\{&\gamma(s^-,S,Z)\gamma(s^-,S,Z)'+\sum_{i=1}^{\bar{h}}\int_{\mathcal{Z}}\Bigg(\Big(\eta_i(s^-,S,Z,z_i)z_i\Big)^2 \\ &+\Big(\ln\left(\left(\eta(t,S,Z,z_i)+1\right)z_i\right)\Big)^2\Bigg)\bar{\nu}_i(\bar{\lambda}_i dz_i)\Bigg\}ds\Bigg]<\infty. \end{aligned} \tag{4.177}$$

Therefore, the following process, denoted by $G(t)$ for each $t\in[0,T]$, is well-defined:

$$\begin{aligned} G(t)=&-\int_0^t\left(\frac{1}{2}\gamma(s^-,S,Z)\gamma(s^-,S,Z)'+\int_{\mathcal{Z}^{\bar{h}}}\Big(\eta(s^-,S,Z,z)'\ \mathrm{diag}(z)\Big)\bar{\nu}(\bar{\lambda}dz)\right)ds \\ &+\int_0^t\gamma(s^-,S,Z)d\bar{W}(s) \\ &+\int_0^t\int_{\mathcal{Z}^{\bar{h}}}\Big(\ln\left(\eta(s^-,S,Z,z)+e\right)'\ \mathrm{diag}(z)\Big)\bar{N}(\bar{\lambda}ds,dz). \end{aligned} \tag{4.178}$$

Furthermore, it follows from the Itô's formula that the stochastic exponential $M(t)=\exp\{G(t)\}$ defined in Proposition 3.1 is an $\{\mathcal{F}_t\}$- and P-local martingale.

Second, let $\Gamma(t)$ denote the unique strong solution of the following SDE with Lévy jumps,

$$d\Gamma(t)=\bar{\phi}(t^-,S,\Gamma)d\bar{W}(t)+\int_{\mathcal{Z}^{\bar{h}}}\bar{\psi}(t^-,S,\Gamma,z)\tilde{\bar{N}}(\bar{\lambda}dt,dz). \tag{4.179}$$

Then, by the fact that the process $M(t)$ is an $\{\mathcal{F}_t\}$-local martingale, and by the Girsanov–Meyer Theorem (see [41] and [33]), we know that

$$\frac{dF_{(S,Z)}}{dF_{(S,\Gamma)}}((S,Z)(t))=M(t). \tag{4.180}$$

Furthermore, we define a process X by

$$\begin{aligned} dX(t)=&\int_0^t\gamma(s^-,S,Z)d\bar{W}(s) \\ &+\int_0^t\int_{\mathcal{Z}^{\bar{h}}}\eta(s^-,S,Z,z)diag(z)\tilde{\bar{N}}(\bar{\lambda}ds,dz). \end{aligned} \tag{4.181}$$

Then it follows from the proof of Theorem 37 in [41] that the Doléans-Dade exponential of X can be calculated by

$$\begin{aligned}\mathcal{E}(X)_t &= \exp\left\{X(t) - \frac{1}{2}[X,X]^c(t)\right\}\prod_{s\leq t}\Big(1+\Delta X(s)\Big)\exp\Big\{-\Delta X(s)\Big\}\\ &= M(t),\end{aligned} \tag{4.182}$$

where $[X,X]^c$ is the continuous part of $[X,X]$. Thus, it follows that

$$M(t) = 1 + \int_0^t M(s^-)dX(s). \tag{4.183}$$

Hence, by the assumption of independence among different driving noises in (3.1)–(3.2), we know that

$$\begin{aligned}E_S[M(t)] = 1 &+ \int_0^t E_S\left[M(s^-)\gamma(s^-,S,Z)\right]d\bar{W}(s)\\ &+ \int_0^t\int_{\mathcal{Z}^{\bar{h}}} E_S\left[M(s^-)\eta(s^-,S,Z,z)'\right]diag(z)\tilde{\bar{N}}(\bar{\lambda}ds,dz).\end{aligned} \tag{4.184}$$

In addition, let

$$H(t) = \ln E_S[M(t)]. \tag{4.185}$$

Then, by applying the Itô's formula to the SDE in (4.184), we have that

$$\begin{aligned}dH(t) = &-\frac{1}{2(E_S[M(t^-)])^2}E_S\left[M(t^-)\gamma(t^-,S,Z)\right]E_S\left[M(t^-)\gamma(t^-,S,Z)\right]'dt\\ &-\int_{\mathcal{Z}^{\bar{h}}}\frac{1}{E_S[M(t^-)]}E_S\left[M(t^-)\eta(t^-,S,Z,z)'diag(z)\right]\bar{\nu}(\bar{\lambda}dz)dt\\ &+\frac{1}{E_S[M(t^-)]}E_S\left[M(t^-)\gamma(t^-,S,Z)\right]d\bar{W}(t)\\ &+\int_{\mathcal{Z}^{\bar{h}}}\ln\left(\left(\frac{E_S\left[M(t^-)\eta(t^-,S,Z,z)\right]}{E_S[M(t^-)]}+e\right)'diag(z)\right)\bar{N}(\bar{\lambda}dt,dz)\\ = &-\left(\frac{1}{2}\hat{\gamma}(t^-,S,Z)\hat{\gamma}(t^-,S,Z)' + \int_{\mathcal{Z}^{\bar{h}}}\hat{\eta}(t^-,S,Z,z)diag(z)\bar{\nu}(\bar{\lambda}dz)\right)dt\\ &+\hat{\gamma}(t^-,S,Z)d\bar{W}(t)\\ &+\int_{\mathcal{Z}^{\bar{h}}}\ln\Big(\left(\hat{\eta}(t^-,S,Z,z)+e\right)'diag(z)\Big)\bar{N}(\bar{\lambda}dt,dz).\end{aligned} \tag{4.186}$$

Note that by an explanation similar to the one for $M(t)$, we know that the following process $\tilde{M}(t)$ is also an $\{\mathcal{F}_t\}$- and P- local martingale,

$$\tilde{M}(t) \equiv E_S[M(t)] = \exp\{H(t)\}. \tag{4.187}$$

Furthermore, by a similar argument given for (4.180), we know that

$$\frac{dF_Z}{dF_\Gamma}(Z(t)) = \tilde{M}(t). \tag{4.188}$$

Next, it follows from (4.180) and (4.188) that the absolute continuity among the corresponding measures of the distributions F_Z, F_Γ, $F_{(S,Z)}$, and $F_{(S,\Gamma)}$ is true, i.e.,

$$F_Z \ll F_\Gamma,\;\; F_{(S,Z)} \ll F_{(Z,\Gamma)}. \tag{4.189}$$

Thus, it follows from the chain rule for Radon–Nikodym derivatives, and (4.189), (4.180), and (4.184), that

$$\frac{dF_{(S,Z)}}{d(F_S\times F_Z)}((S,Z)(t))=\frac{\frac{dF_{(S,Z)}}{d(F_S\times F_\Gamma)}((S,Z)(t))}{\frac{dF_Z}{dF_\Gamma}(Z(t))}=M(t)\tilde{M}^{-1}(t). \tag{4.190}$$

Furthermore, by the relationship in (4.190) and the definition of the mutual information in (3.3), we know that

$$I(T,S,Z)=\int\left(\ln M(T)-\ln\tilde{M}(T)\right)dF_{(S,Z)}. \tag{4.191}$$

Hence, it follows from equations (3.11), (4.184), (4.191), and Corollary 8.7 in [26] that the formula given by (3.8) in Proposition 3.1 is true.

5. Conclusion

In this paper, we are concerned with the well-posedness and applications of a unified system of coupled FB-SDEs with completely-$\mathcal{S}$ skew reflections and Lévy jumps. Owing to the reflections, the solution to an embedded Skorohod problem may be not unique, i.e., bifurcations may occur at reflection boundaries, the well-known contraction mapping approach can not be extended directly to solve our problem. Thus, we develop a weak convergence method to prove the well-posedness of an adapted 6-tuple weak solution (in the sense of distributions) to the unified system. Furthermore, in our proof we adopt a generalized linear growth and Lipschitz condition that guarantees the well-posedness of the unified system even under a random environment. In addition, if the spectral radii for the reflections are strictly less than unity, a unique adapted 6-tuple strong solution is considered. As applications of our unified system, we also develop new techniques including deriving a generalized mutual information formula for signal processing over possible non-Gaussian MIMO channels with dynamics driven by Lévy processes. Finally, since our unified system is formulated possibly in the most general form with feedback control concerning various SDEs, we predict that our main results can be applied to more areas.

REFERENCES

[1] A.S. Acampora, S. Bhardwaj, and R.M. Tamari, *On best-case throughput of cellular data networks with cooperating base stations*, Proc. of the Allerton Conference on Communication, Control, and Computing, Monticello, IL, Step., 2006.

[2] D. Applebaum, *Lévy Processes and Stochastic Calculus*, Cambridge University Press, Cambridge, 2004.

[3] A. Bernard and A. El Kharroubi, *Régulations déterministes et stochastiques dans le premier "orthant" de R^n*, Stochastics Stochastics Rep., 34:149–167, 1991.

[4] J. Bertoin, *Lévy Processes*, Cambridge University Press, Cambridge, 1998.

[5] P. Billingsley, *Convergence of Probability Measures*, Second Edition, New York: John Wiley & Sons, 1999.

[6] H. Chen and D.D. Yao, *Fundamental of Queueing Networks*, Springer-Verlag, New York, 2001.

[7] K.L. Chung, *A Course in Probability*, Wiley, New York, 1974.

[8] M.H.M. Costa, *Writing on dirty paper*, IEEE Transactions on Information Theory, 29(3):439–441, 1983.

[9] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., New York, 1991.

[10] J.G. Dai and W. Dai, *A heavy traffic limit theorem for a class of open queueing networks with finite buffers*, Queueing Systems, 32(1-3):5–40, 1999.

[11] J.G. Dai and R.J. Williams, *Existence and uniqueness of semimartingale reflecting Brownian motions in convex polyhedrons*, Theory of Probability and its Applications, 40(1):1–40, 1995.

[12] W. Dai, *Brownian approximations for queueing networks with finite buffers: modeling, heavy traffic analysis and numerical implementations*, Ph.D Thesis, The Georgia Institute of Technology, 1996. Also published in UMI Dissertation Services, A Bell & Howell Company, Michigan, U.S.A., 1997.

[13] W. Dai, *A Brownian model for multiclass queueing networks with finite buffers*, J. Comput. Appl. Math., 144(1-2):145–160, 2002.

[14] W. Dai, *Mean-variance portfolio selection based on a generalized BNS stochastic volatility model*, Inte. J. Comput. Math., 88(16):3521–3534, 2011.

[15] W. Dai, *Optimal rate scheduling via utility-maximization for J-user MIMO Markov fading wireless channels with cooperation*, Operations Research, 61(6):1450–1462, 2013.

[16] W. Dai, *Mean-variance hedging based on an incomplete market with external risk factors of non-Gaussian OU processes*, Mathematical Problems in Engineering, Volume 2015 (Regular Issue), Article ID 625289, 2015.

[17] L. Delong and P. Imkeller, *On Malliavin's differentiability of BSDE with time delayed generators driven by Brownian motions and Poisson random measures*, Stochastic Processes Appl., 120(9):1748–1775, 2010.

[18] A. De Mauro, M. Greco, and M. Grimaldi, *A formal definition of big data based on its essential features*, Library Review, 65:122–135, 2016.

[19] T.E. Duncan, *Mutual information for stochastic signals and Lévy processes*, IEEE Transactions on Information Theory, 56(1):18–24, 2010.

[20] S.N. Ethier and T.G. Kurtz, *Markov Processes: Characterization and Convergence*, New York: John Wiley & Sons Inc., 1986.

[21] A.E. Gamal and T.M. Cover, *Multiple user information theory*, Proceedings of the IEEE, 68(12):1466–1483, 1980.

[22] I.I. Gihman and A.V. Skorohod, *Stochastic Differential Equations*, Springer-Verlag, Berlin, 1972.

[23] A. Goldsmith, S.A. Jafar, N. Jindal, and N. Vishwanath, *Capacity limits of MIMO channels*, IEEE Journal on Selected Areas in Communications, 21(5):684–702, 2003.

[24] J.M. Harrison and M.I. Reiman, *Reflected Brownian motion on an orthant*, Annals of Probability, 9(2):302–308, 1981.

[25] J.B. Folland, *Real Analysis: Modern Techniques and Their Applications*, Second Edition, Wiley, 2007.

[26] J. Jacod and A.N. Shiryaev, *Limit Theorems for Stochastic Processes*, Second Edition, Springer, Berlin, 2002.

[27] N. Jindal, S. Vishwanath, and A. Goldsmith, *On the duality of Gaussian multiple-access and broadcast channels*, IEEE Transactions on Information Theory, 50(5):768–783, 2004.

[28] O. Kallenberg, *Foundations of Modern Probability*, Springer-Verlag, Berlin, 1997.

[29] I. Karatzas and Q. Li, *BSDE approach to non-zero-sum stochastic differential games of control and stopping*, in Stochastic Processes, Finance and Control, World Scientific Publishers, 105–153, 2012.

[30] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Second Edition, Springer, Berlin, 1991.

[31] T. Konstantopoulos, G. Last, and S.J. Lin, *On a class of Lévy stochastic networks*, Queueing Systems, 46:409–437, 2004.

[32] J.A. León, J.L. Solé, F. Utzet, and J. Vives, *Local Malliavin calculus for Lévy processes and applications*, Stochastics, 86(4):551-572, 2014.

[33] R.S. Liptser and A.N. Shiryaev, *Statistics of Random Process*, Second Edition, Springer, Berlin, 2001.

[34] A. Mandelbaum, *The dynamic complementarity problem*, preprint, 1989.

[35] A. Mandelbaum and W. A. Massey, *Strong approximations for time-dependent queues*, Mathematics of Operations Research, 20(1):33–64, 1995.

[36] A. Mandelbaum and G. Pats, *State-dependent stochastic networks Part I: approximations and applications with continuous diffusion limits*, Ann. Appl. Probab., 8(2):569–646, 1998.

[37] B. ∅ksendal, *Stochastic Differential Equations*, Sixth Edition, Springer, New York, 2005.

[38] B. ∅ksendal and A. Sulem, *Applied Stochastic Control of Jump Diffusions*, Springer-Verlag, Berlin, 2005.

[39] B. ∅ksendal, A. Sulem, and T. Zhang, *A stochastic HJB equation for optimal control of forward-backward SDEs*, The Fascination of Probability, Statistics and their Applications, 435–446, 2015.

[40] G.D. Nunno, B. Øksendal, and F. Proske, *Malliavin Calculus for Lévy Processes with Applications to Finance*, Springer, Berlin, 2009.

[41] P.E. Protter, *Stochastic Integration and Differential Equations*, Second Edition, Springer, New York, 2004.

[42] M.I. Reiman and R.J. Williams, *A boundary property of semimartingale reflecting Brownian motions*, Probab. Th. Rel. Fields, 77:87–97, 1988.

[43] K.I. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, 1999.

[44] C.E. Shannon, *A mathematical theory of communication*, Bell Sys. Tech. J., 27:379–423, 623–656, 1948.

[45] J.L. Solé, F. Utzet, and J. Vives, *Canonocal Lévy process and Malliavin Calculus*, Stochastic Processes and Their Applications, 117:165–187, 2007.

[46] H. Tanaka, *Stochastic differential equations with reflecting boundary condition in convex regions*, Hiroshima Math J., 9:163–177, 1979.

[47] L.M. Taylor and R.J. Williams, *Existence and uniqueness of semimartingale reflecting Brownian motions in an orthant*, Probability Theory and Related Fields, 96(3):283–317, 1993.

[48] R.J. Williams, *An invariance principle for semimartingale reflecting Brownian motions in an orthant*, Queueing Systems, 30:5–25, 1998.

[49] K. Yosida, *Functional Analysis*, Sixth Edition, Springer-Verlag, Berlin, 1980.