ESAIM: COCV 27 (2021) 30
https://doi.org/10.1051/cocv/2021031

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

CONTROLLABILITY GRAMIAN AND KALMAN RANK CONDITION FOR MEAN-FIELD CONTROL SYSTEMS*

Zhiyong Yu**

Abstract. This paper is concerned with the exact controllability of linear mean-field stochastic systems with deterministic coefficients. With the help of the theory of mean-field backward stochastic differential equations (MF-BSDEs, for short) and some delicate analysis, we obtain a mean-field version of the Gramian matrix criterion for the general time-variant case, and a mean-field version of the Kalman rank condition for the special time-invariant case.

Mathematics Subject Classification. 60H10, 93B05.

Received October 4, 2020. Accepted March 15, 2021.

1. Introduction

This paper is concerned with the following controlled linear stochastic differential equation with McKean-Vlasov type, a kind of mean-field type (MF-SDE, for short):

$$\begin{aligned} \mathrm{d}x(t) = &\Big\{A(t)x(t) + \bar{A}(t)\mathbb{E}[x(t)] + B(t)u(t) + \bar{B}(t)\mathbb{E}[u(t)]\Big\}\,\mathrm{d}t \\ &+ \Big\{C(t)x(t) + \bar{C}(t)\mathbb{E}[x(t)] + D(t)u(t) + \bar{D}(t)\mathbb{E}[u(t)]\Big\}\,\mathrm{d}W(t), \quad t\in[0,T], \end{aligned} \tag{1.1}$$

where $W(\cdot)$ is a 1-dimensional standard Brownian motion, $A(\cdot)$, $B(\cdot)$, etc. are suitable matrix-valued deterministic processes, $T>0$ is a fixed time horizon, $u(\cdot)$ is the control process, and $x(\cdot)$ is the state process. The system (1.1) is called *exact controllable* if for every initial state $x_0\in\mathbb{R}^n$, where $\mathbb{R}^n$ is the n-dimensional Euclidean space, and every terminal state $x_T\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$, whose rigorous definition will be given in the next section, there exists a control $u(\cdot)$ such that

$$x(0)=x_0, \qquad x(T)=x_T. \tag{1.2}$$

In this paper, we are interested in the exact controllability of (1.1).

* This work is supported by the National Natural Science Foundation of China (11871310) and the National Key R&D Program of China (2018YFA0703900).

Keywords and phrases: Backward stochastic differential equation, mean-field stochastic differential equation, exact controllability, controllability Gramian, Kalman rank condition.

School of Mathematics, Shandong University, Jinan 250100, PR China.

** Corresponding author: yuzhiyong@sdu.edu.cn

It is well known that controllability is one of the most important concepts in the control theory. For an extensive surveys of this issue on deterministic systems, one can refer to Lee and Markus [12] for ordinary differential equation (ODE, for short) systems and Russell [19], Lions [13], and Zuazua [27] for partial differential equation (PDE, for short) systems. In the research on ODE systems, the controllability Gramian and the Kalman rank condition are the most important results for time-variant systems and time-invariant systems, respectively.

Among the literature on stochastic differential equation (SDE, for short) systems, we would like to mention the following results which are more related to our present work. By virtue of the theory of backward stochastic differential equations (BSDEs, for short), Peng [17] proved a rank condition to characterize the exact controllability for stochastic systems when the coefficients are deterministic and time-invariant. However, the matrix related to Peng's rank condition consists of infinitely many columns, which brings some difficulty of verification. Recently, Lü and Zhang [15] simplified Peng's rank condition from infinite many columns to finite many columns and obtained a stochastic Kalman-type rank condition. When the coefficients are deterministic but allowed to be time-variant, Liu and Peng [14] obtained a stochastic version of controllability Gramian. Along this line, Bi *et al.* [1] further characterized the reachability of a particular point. For time-variant and random coefficients, Wang *et al.* [21] got some equivalent conditions. Recently, Wang and Yu [22] extended the notion of exact controllability and proposed a notion of *partial controllability*, then extended the results of [21]. Zhang and Guo [26] introduced a new kind of stochastic game-based control systems and investigated the related exact controllability. Different from the exact controllability, there exists another kind of notion, named *approximate controllability*, in the literature, which only requires that the control process steers the state to a neighborhood of the target. In this direction, one can refer to Buckdahn *et al.* [5] and Goreac [7]. We notice that in [7] the author proved that the approximate controllability and the exact controllability are not equivalent to each other.

The McKean-Vlasov type SDEs were first introduced when Kac [10] studied plasma dynamics equations and further developed by McKean [16]. Accompanying a recent surge of interest in mean-field games originally formulated by Huang *et al.* [9] and Lasry and Lions [11], the McKean-Vlasov SDEs and their numerous applications have received a great development in the past ten years. Especially for the corresponding optimal control and game theory, one can refer to Yong [25], Buckdahn *et al.* [2], Carmona and Delarue [6], Pham and Wei, Tian *et al.* [20], and so on. Meanwhile, there exists a little research on the controllability of McKean-Valsov SDEs. Goreac [8] investigated the approximate controllability of (1.1) with deterministic time-invariant coefficients. Recently, Ye and Yu [24] were interested in the exact controllability of (1.1) (but with random time-variant coefficients) and obtained the equivalence with exact observability of a dual system. As a continuation, the present paper is devoted to getting the controllability Gramian and the Kalman rank condition for (1.1).

We notice that, for the ODE version of controllability, the terminal state x_T is required to be every point of $\mathbb{R}^n$ (compared with (1.2)). This is a finite-dimensional space. Now, for our system (1.1), the space $\mathbb{R}^n$ is replaced by an infinite-dimensional one $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$. From finite to infinite dimensions makes the difficulty of the problem increase abruptly. Fortunately, this difficulty has already arisen in the case of SDE systems, and Peng [17] and Wang *et al.* [21] proposed two different approaches to overcome the difficulty for time-invariant and time-variant systems, respectively. In the present paper, the approach of Wang *et al.* [21] is adopted since the sytem (1.1) is time-variant. On the other hand, because Peng's approach for the special time-invariant case of (1.1) has some reference value from another viewpoint, then it is put in the appendix.

In this paper, in order for the approach in [21] (and the other approach in [17]) to be successfully applied to the MF-SDE case, we need to combine the following two characteristics of the system (1.1): (i) The setting of deterministic coefficients allows them to freely enter or exit the mathematical expectation operator, such as $\mathbb{E}[A(\cdot)x(\cdot)] = A(\cdot)\mathbb{E}[x(\cdot)]$; (ii) In many mean-field problems, the decomposition on a random variable $\xi = (\xi - \mathbb{E}[\xi]) + \mathbb{E}[\xi]$ often plays a key role, because the first term $(\xi - \mathbb{E}[\xi])$ has zero mean and the second term $\mathbb{E}[\xi]$ is deterministic.

In fact, under suitable conditions (see Assumption (H) in Sect. 3), the system (1.1) will be reduced to the following one:

$$\begin{cases} \mathrm{d}x(t) = \Big\{ \mathbf{A}(t)x(t) + \bar{\mathbf{A}}(t)\mathbb{E}[x(t)] + \mathbf{A}_1(t)q(t) + \bar{\mathbf{A}}_1(t)\mathbb{E}[q(t)] \\ \qquad\qquad + \mathbf{B}(t)v(t) + \bar{\mathbf{B}}(t)\mathbb{E}[v(t)] \Big\} \,\mathrm{d}t + q(t)\,\mathrm{d}W(t), \quad t \in [0,T], \\ x(T) = 0, \end{cases} \tag{1.3}$$

where $\mathbf{A}(\cdot)$, $\mathbf{A}_1(\cdot)$, $\mathbf{B}(\cdot)$, etc. will be defined in Section 3. Comparing it with the original system (1.1), a zero terminal condition is added into the system (1.3). Limiting the terminal state from every point in $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ to zero shifts the problem from an infinite-dimensional framework to a finite-dimensional one. In other words, now we only need, for every initial state $x_0 \in \mathbb{R}^n$, find a control $v(\cdot)$ such that (compared with (1.2))

$$x(0) = x_0. \tag{1.4}$$

Different from the system (1.1), the new system (1.3) should be understood as an MF-BSDE. In fact, the system (1.3) corresponds to an *exact null-controllability* in the classical sense. In this paper, we write the zero terminal value directly into the system (1.3) to highlight the viewpoint of MF-BSDEs.

For the ODE systems, it is well known that the proof of Kalman rank condition essentially depends on the following expression of a related fundamental matrix:

$$e^{tA} = \sum_{i=0}^{\infty} \frac{t^i A^i}{i!}, \qquad t \in [0,T]. \tag{1.5}$$

This expression is too restrictive such that the classical proof cannot be generalized to the MF-BSDE system (1.3) we care about. We notice that a similar difficulty has already arisen in the case of BSDE systems. But, unfortunately, the corresponding proof is missing in Peng [17]. In the present paper, although the fundamental matrix $\Phi(\cdot)$ (see the definition (5.22)) to a related MF-SDE does not admit an expression similar to (1.5), we instead introduce an approximation sequence $\{\Phi^{(k)}(\cdot)\}_{k=1}^{\infty}$ for $\Phi(\cdot)$. By proving the relevant properties of $\Phi^{(k)}(\cdot)$ ($k = 1, 2, \dots$) and sending $k \to \infty$, we overcome this difficulty and obtain the desired Kalman rank condition. As a corollary, our approach reduced to the special case of BSDE systems fills the gap in [17].

In order to state and prove the Kalman rank condition, especially to analyze the mean-field dependence, in this paper, we introduce a couple of new multiplication operations: "$\otimes$" and "$\circ$" (see Def. 5.1 and Eq. (5.31)). The first one is called a *block-tensor product* which can be regarded as a combination of the classical product and the tensor product of matrices. The second one can be regarded as a combination of the classical matrix product with a scalar and the classical inner product of vectors. Moreover, a family of auxiliary scalar-valued stochastic processes $\eta_j^{(k)}$ ($k = 1, 2, \dots;\ j = 0, 1, 2, \dots, 2^{k-1}$) are also introduced (see (5.29)). These operations and processes are subtlely designed and work together to complete the proofs.

The rest of this paper is organized as follows. In Section 2, we introduce some notations and do some preliminary works. In Section 3, we reduce the system (1.1) to the system (1.3). Then, for the exact controllability of (1.3), a mean-field version of controllability Gramian (see Thm. 4.1) is obtained in Section 4, and a mean-field version of Kalman rank condition (see Thm. 5.9) is also obtained in Section 5.

2. Preliminaries

Recall that $\mathbb{R}^n$ is the n-dimensional Euclidean space equipped with the Euclidean inner product $\langle \cdot,\ \cdot \rangle$ and the induced norm $|\cdot|$. Let $\mathbb{R}^{n\times m}$ be the set of all $(n \times m)$ matrices. With the inner product

$$\langle A,\ B \rangle = \mathrm{tr}\,\left[A^\top B\right] \quad \text{for any } A,\ B \in \mathbb{R}^{n\times m},$$

$\mathbb{R}^{n\times m}$ is also a Euclidean space. Here and hereafter, the superscript "$\top$" denotes the transpose of a vector or a matrix.

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space on which a 1-dimensional Brownian motion $W(\cdot)$ is defined. In this paper, we set the dimension of Brownian motion $d = 1$ for simplicity. The case $d > 1$ can be discussed similarly. Let $\mathbb{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$ be the natural filtration of $W(\cdot)$ augmented by all $\mathbb{P}$-null sets, and $\mathcal{F} = \mathcal{F}_T$.

For any Euclidean space $\mathbb{R}^n$, we introduce some Banach (sometimes more accurately, Hilbert) spaces as follows.

- $L^\infty(0,T;\mathbb{R}^n)$ is the set of all Lebesgue measurable deterministic processes valued in $\mathbb{R}^n$ and essentially bounded;
- $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ is the set of all $\mathcal{F}_T$-measurable random variables ξ valued in $\mathbb{R}^n$ such that
$$\mathbb{E}\big[|\xi|^2\big] < \infty;$$
- $L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$ is the set of all $\mathbb{F}$-progressively measurable processes $f(\cdot)$ valued in $\mathbb{R}^n$ such that
$$\mathbb{E}\int_0^T |f(t)|^2\,\mathrm{d}t < \infty;$$
- $L^2_{\mathbb{F}}(\Omega;C(0,T;\mathbb{R}^n))$ is the set of all $\mathbb{F}$-progressively measurable processes $f(\cdot)$ valued in $\mathbb{R}^n$ such that for almost all $\omega\in\Omega$, $t\mapsto f(t,\omega)$ is continuous and
$$\mathbb{E}\Bigg[\sup_{t\in[0,T]}|f(t)|^2\Bigg] < \infty.$$

We now return to our mean-field system (1.1) and give the rigorous assumptions on the coefficients: $A(\cdot)$, $\bar{A}(\cdot)$, $C(\cdot)$, $\bar{C}(\cdot) \in L^\infty(0,T;\mathbb{R}^{n\times n})$ and $B(\cdot)$, $\bar{B}(\cdot)$, $D(\cdot)$, $\bar{D}(\cdot)\in L^\infty(0,T;\mathbb{R}^{n\times m})$. By Proposition 2.6 in [25] (see also Prop. 2.2 in [23] for a wider situation), for any $x_0\in\mathbb{R}^n$ and any $u(\cdot)\in L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$, the MF-SDE (1.1) admits a unique solution $x(\cdot)\equiv x(\cdot;x_0,u(\cdot))\in L^2_{\mathbb{F}}(\Omega;C(0,T;\mathbb{R}^n))$. Next, we give the rigorous definition of exact controllability.

Definition 2.1. The system (1.1) is called exactly controllable on $[0,T]$, if for any $x_0\in\mathbb{R}^n$ and any $x_T\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$, there exists a $u(\cdot)\in L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$ such that the unique solution $x(\cdot)\equiv x(\cdot;x_0,u(\cdot))$ to (1.1) satisfies $x(T)\equiv x(T;x_0,u(\cdot)) = x_T$.

In this paper, we use the following notations for convenience:
$$\widehat{A}(\cdot)=A(\cdot)+\bar{A}(\cdot),\quad \widehat{B}(\cdot)=B(\cdot)+\bar{B}(\cdot),\quad \widehat{C}(\cdot)=C(\cdot)+\bar{C}(\cdot),\quad \widehat{D}(\cdot)=D(\cdot)+\bar{D}(\cdot). \tag{2.1}$$

Then, the controlled system (1.1) can be rewritten as (the argument t is suppressed for simplicity)
$$\begin{aligned}\mathrm{d}x &= \Big\{A\big(x-\mathbb{E}[x]\big)+\widehat{A}\mathbb{E}[x]+B\big(u-\mathbb{E}[u]\big)+\widehat{B}\mathbb{E}[u]\Big\}\,\mathrm{d}t\\ &\quad+\Big\{C\big(x-\mathbb{E}[x]\big)+\widehat{C}\mathbb{E}[x]+D\big(u-\mathbb{E}[u]\big)+\widehat{D}\mathbb{E}[u]\Big\}\,\mathrm{d}W,\quad t\in[0,T].\end{aligned} \tag{2.2}$$

First of all, we can easily get a couple of necessary conditions for the exact controllability of (1.1). Indeed, by taking the expectation to MF-SDE (1.1), we get an ODE:
$$\mathrm{d}\big(\mathbb{E}[x(t)]\big)=\Big\{\widehat{A}(t)\mathbb{E}[x(t)]+\widehat{B}(t)\mathbb{E}[u(t)]\Big\}\,\mathrm{d}t,\quad t\in[0,T]. \tag{2.3}$$

Let us focus temporarily on ODE (2.3). We introduce a deterministic matrix-valued process $\widehat{\Phi}(\cdot)$ to denote the solution of the following ODE:

$$\begin{cases} \mathrm{d}\widehat{\Phi}(t) = -\widehat{\Phi}(t)\widehat{A}(t)\,\mathrm{d}t, \quad t \in [0,T], \\ \widehat{\Phi}(0) = I. \end{cases} \tag{2.4}$$

It is clear that $\widehat{\Phi}(\cdot)$ is indeed the inverse of the classical fundamental matrix to ODE (2.3). Based on $\widehat{\Phi}(\cdot)$, we define a symmetric matrix as follows:

$$\widehat{\Psi} = \int_0^T \widehat{\Phi}(t)\widehat{B}(t)\widehat{B}(t)^\top \widehat{\Phi}(t)^\top \,\mathrm{d}t. \tag{2.5}$$

As corollaries of the controllability Gramian and the Kalman rank condition for the ODE system (2.3), we have

Proposition 2.2. *(i). If the system* (1.1) *is exactly controllable on* $[0,T]$*, then the matrix* $\widehat{\Psi}$ *is non-singular.*

(ii). If we have an additional condition that both $\widehat{A}(\cdot)$ *and* $\widehat{B}(\cdot)$ *are time-invariant, then the exact controllability of* (1.1) *implies*

$$\operatorname{Rank}\left(\widehat{B},\ \widehat{A}\widehat{B},\ \widehat{A}^2\widehat{B},\ \dots,\ \widehat{A}^{n-1}\widehat{B}\right) = n. \tag{2.6}$$

In the rest of this paper, we aim to find some non-trivial sufficient conditions or equivalent conditions for the exact controllability of (1.1).

3. Exact null-controllability and MF-BSDE

In this section, we shall reduce the original system (1.1) to the simpler one (1.3) along the approach in [21]. This reduction is a preparation for the controllability Gramian in Section 4 and the Kalman rank condition in Section 5.

For the reduction, we need the following assumptions on the coefficients.

Assumption (H). There exists a constant $\delta > 0$ such that

$$D(t)D(t)^\top,\ \widehat{D}(t)\widehat{D}(t)^\top \geq \delta I \quad \text{for almost all } t \in [0,T]. \tag{3.1}$$

In this case, $[D(t)D(t)^\top]^{-1}$ and $[\widehat{D}(t)\widehat{D}(t)^\top]^{-1}$ exist and essentially bounded. Moreover, it is clear that, in the special case that $D(\cdot)$ and $\widehat{D}(\cdot)$ are time-invariant, Assumption (H) is equivalent to

Assumption (H′). Rank $D = n$ and Rank $\widehat{D} = n$.

Obviously, Assumption (H′) (or Assumption (H)) implies that the dimension of the control is bigger or equal to that of the state. We notice that, when all of the coefficients in the system (1.1) are time-invariant, Goreac [8, Proposition 2] proved that Rank $\widehat{D} = n$ is a necessary condition for the exact controllability. Then he continued to obtain that Assumption (H′) is a sufficient condition for the *exact terminal controllability* (see [8], Prop. 5). Here, the exact terminal controllability is a weaker notion than exact controllability. In the present paper, we shall work under Assumption (H) to investigate the exact controllability of (1.1) with time-variant and time-invariant coefficients.

Let

$$\begin{cases} \mathbf{A}(\cdot) = A(\cdot) - B(\cdot)D(\cdot)^\top\big[D(\cdot)D(\cdot)^\top\big]^{-1}C(\cdot), \\ \mathbf{A}_1(\cdot) = B(\cdot)D(\cdot)^\top\big[D(\cdot)D(\cdot)^\top\big]^{-1}, \\ \mathbf{B}(\cdot) = B(\cdot)\Big(I - D(\cdot)^\top\big[D(\cdot)D(\cdot)^\top\big]^{-1}D(\cdot)\Big), \end{cases} \tag{3.2}$$

$$\begin{cases} \widehat{\mathbf{A}}(\cdot) = \widehat{A}(\cdot) - \widehat{B}(\cdot)\widehat{D}(\cdot)^\top\big[\widehat{D}(\cdot)\widehat{D}(\cdot)^\top\big]^{-1}\widehat{C}(\cdot), \\ \widehat{\mathbf{A}}_1(\cdot) = \widehat{B}(\cdot)\widehat{D}(\cdot)^\top\big[\widehat{D}(\cdot)\widehat{D}(\cdot)^\top\big]^{-1}, \\ \widehat{\mathbf{B}}(\cdot) = \widehat{B}(\cdot)\Big(I - \widehat{D}(\cdot)^\top\big[\widehat{D}(\cdot)\widehat{D}(\cdot)^\top\big]^{-1}\widehat{D}(\cdot)\Big), \end{cases} \tag{3.3}$$

and

$$\bar{\mathbf{A}}(\cdot) = \widehat{\mathbf{A}}(\cdot) - \mathbf{A}(\cdot), \quad \bar{\mathbf{A}}_1(\cdot) = \widehat{\mathbf{A}}_1(\cdot) - \mathbf{A}_1(\cdot), \quad \bar{\mathbf{B}}(\cdot) = \widehat{\mathbf{B}}(\cdot) - \mathbf{B}(\cdot). \tag{3.4}$$

We introduce the third controlled system:

$$\begin{aligned} \mathrm{d}x(t) = \Big\{ &\mathbf{A}(t)x(t) + \bar{\mathbf{A}}(t)\mathbb{E}[x(t)] + \mathbf{A}_1(t)q(t) + \bar{\mathbf{A}}_1(t)\mathbb{E}[q(t)] \\ &+ \mathbf{B}(t)v(t) + \bar{\mathbf{B}}(t)\mathbb{E}[v(t)]\Big\}\,\mathrm{d}t + q(t)\,\mathrm{d}W(t), \quad t \in [0,T], \end{aligned} \tag{3.5}$$

where $x(\cdot)$ is the state process and $(q(\cdot), v(\cdot))$ is a pair of control processes. Equivalently (the argument t is suppressed),

$$\begin{aligned} \mathrm{d}x = \Big\{ &\mathbf{A}\big(x - \mathbb{E}[x]\big) + \widehat{\mathbf{A}}\mathbb{E}[x] + \mathbf{A}_1\big(q - \mathbb{E}[q]\big) + \widehat{\mathbf{A}}_1\mathbb{E}[q] \\ &+ \mathbf{B}\big(v - \mathbb{E}[v]\big) + \widehat{\mathbf{B}}(t)\mathbb{E}[v]\Big\}\,\mathrm{d}t + q\,\mathrm{d}W, \quad t \in [0,T]. \end{aligned} \tag{3.6}$$

Proposition 3.1. *Under Assumption (H), the exact controllability of* (1.1) *is equivalent to that of* (3.5).

Proof. (Necessity). Let $(x_0, x_T) \in \mathbb{R}^n \times L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ be given. Since the system (1.1) is exactly controllable, then there exists a pair $(x(\cdot), u(\cdot)) \in L^2_{\mathbb{F}}(\Omega; C(0,T;\mathbb{R}^n)) \times L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$ satisfying

$$\begin{cases} \mathrm{d}x = \Big\{A\big(x - \mathbb{E}[x]\big) + \widehat{A}\mathbb{E}[x] + B\big(u - \mathbb{E}[u]\big) + \widehat{B}\mathbb{E}[u]\Big\}\,\mathrm{d}t \\ \qquad + \Big\{C\big(x - \mathbb{E}[x]\big) + \widehat{C}\mathbb{E}[x] + D\big(u - \mathbb{E}[u]\big) + \widehat{D}\mathbb{E}[u]\Big\}\,\mathrm{d}W, \quad t \in [0,T], \\ x(0) = x_0, \qquad x(T) = x_T. \end{cases} \tag{3.7}$$

Define

$$\begin{cases} q(\cdot) = C(\cdot)\big(x(\cdot) - \mathbb{E}[x(\cdot)]\big) + \widehat{C}(\cdot)\mathbb{E}[x(\cdot)] + D(\cdot)\big(u(\cdot) - E[u(\cdot)]\big) + \widehat{D}(\cdot)\mathbb{E}[u(\cdot)], \\ v(\cdot) = u(\cdot) \end{cases} \tag{3.8}$$

which is equivalent to

$$\begin{cases} \mathbb{E}[q] = \widehat{C}\mathbb{E}[x] + \widehat{D}\mathbb{E}[u], \qquad q - \mathbb{E}[q] = C\big(x - \mathbb{E}[x]\big) + D\big(u - E[u]\big), \\ \mathbb{E}[v] = \mathbb{E}[u], \qquad v - \mathbb{E}[v] = u - \mathbb{E}[u]. \end{cases}$$

It is easy to verify that $(q(\cdot), v(\cdot)) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n) \times L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$. We calculate

$$\begin{aligned} \mathbb{E}[u] &= \widehat{D}^\top \big[\widehat{D}\widehat{D}^\top\big]^{-1}\widehat{D}\mathbb{E}[u] + \Big(I - \widehat{D}^\top\big[\widehat{D}\widehat{D}^\top\big]^{-1}\widehat{D}\Big)\mathbb{E}[u] \\ &= \widehat{D}^\top \big[\widehat{D}\widehat{D}^\top\big]^{-1}\big(\mathbb{E}[q] - \widehat{C}\mathbb{E}[x]\big) + \Big(I - \widehat{D}^\top\big[\widehat{D}\widehat{D}^\top\big]^{-1}\widehat{D}\Big)\mathbb{E}[v]. \end{aligned} \tag{3.9}$$

By the definition (3.3),

$$\begin{aligned} \widehat{B}\mathbb{E}[u] &= \widehat{B}\widehat{D}^\top\big[\widehat{D}\widehat{D}^\top\big]^{-1}\mathbb{E}[q] - \widehat{B}\widehat{D}^\top\big[\widehat{D}\widehat{D}^\top\big]^{-1}\widehat{C}\mathbb{E}[x] + \widehat{B}\Big(I - \widehat{D}^\top\big[\widehat{D}\widehat{D}^\top\big]^{-1}\widehat{D}\Big)\mathbb{E}[v] \\ &= \widehat{\mathbf{A}}_1\mathbb{E}[q] + \big(\widehat{\mathbf{A}} - \widehat{A}\big)\mathbb{E}[x] + \widehat{\mathbf{B}}\mathbb{E}[v], \end{aligned}$$

i.e.,

$$\widehat{A}\mathbb{E}[x] + \widehat{B}\mathbb{E}[u] = \widehat{\mathbf{A}}\mathbb{E}[x] + \widehat{\mathbf{A}}_1\mathbb{E}[q] + \widehat{\mathbf{B}}\mathbb{E}[v]. \tag{3.10}$$

Similarly,

$$u - \mathbb{E}[u] = D^\top\big[DD^\top\big]^{-1}\Big\{\big(q - \mathbb{E}[q]\big) - C\big(x - \mathbb{E}[x]\big)\Big\} + \Big(I - D^\top\big[DD^\top\big]^{-1}D\Big)\big(v - \mathbb{E}[v]\big), \tag{3.11}$$

and by the definition (3.2),

$$A\big(x - \mathbb{E}[x]\big) + B\big(u - \mathbb{E}[u]\big) = \mathbf{A}\big(x - \mathbb{E}[x]\big) + \mathbf{A}_1\big(q - \mathbb{E}[q]\big) + \mathbf{B}\big(v - \mathbb{E}[v]\big). \tag{3.12}$$

Due to (3.8), (3.10) and (3.12), we rewrite (3.7) as

$$\begin{cases} \mathrm{d}x = \Big\{\mathbf{A}\big(x - \mathbb{E}[x]\big) + \widehat{\mathbf{A}}\mathbb{E}[x] + \mathbf{A}_1\big(q - \mathbb{E}[q]\big) + \widehat{\mathbf{A}}_1\mathbb{E}[q] \\ \qquad\qquad + \mathbf{B}\big(v - \mathbb{E}[v]\big) + \widehat{\mathbf{B}}(t)\mathbb{E}[v]\Big\}\,\mathrm{d}t + q\,\mathrm{d}W, \quad t \in [0,T], \\ x(0) = x_0, \qquad x(T) = x_T. \end{cases} \tag{3.13}$$

This means that the pair of the controls $(q(\cdot), v(\cdot))$ defined by (3.8) steers the state $x(\cdot)$ from $x(0) = x_0$ to $x(T) = x_T$. By the arbitrariness of (x_0, x_T), we prove the exact controllability of (3.5).

(Sufficiency). Let $(x_0, x_T) \in \mathbb{R}^n \times L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ be given. Since the system (3.5) is exactly controllable, then there exists a triple $(x(\cdot), q(\cdot), v(\cdot)) \in L^2_{\mathbb{F}}(\Omega; C(0,T;\mathbb{R}^n)) \times L^2_{\mathbb{F}}(0,T;\mathbb{R}^n) \times L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$ satisfying (3.13). Based on this triple, we define

$$\begin{aligned} u = {} & D^\top\big[DD^\top\big]^{-1}\Big\{\big(q - \mathbb{E}[q]\big) - C\big(x - \mathbb{E}[x]\big)\Big\} + \Big(I - D^\top\big[DD^\top\big]^{-1}D\Big)\big(v - \mathbb{E}[v]\big) \\ & + \widehat{D}^\top\big[\widehat{D}\widehat{D}^\top\big]^{-1}\big(\mathbb{E}[q] - \widehat{C}\mathbb{E}[x]\big) + \Big(I - \widehat{D}^\top\big[\widehat{D}\widehat{D}^\top\big]^{-1}\widehat{D}\Big)\mathbb{E}[v] \end{aligned} \tag{3.14}$$

which is equivalent to

$$\begin{cases} \mathbb{E}[u] = \widehat{D}^\top \big[\widehat{D}\widehat{D}^\top\big]^{-1}\big(\mathbb{E}[q] - \widehat{C}\mathbb{E}[x]\big) + \Big(I - \widehat{D}^\top \big[\widehat{D}\widehat{D}^\top\big]^{-1}\widehat{D}\Big)\mathbb{E}[v], \\ u - \mathbb{E}[u] = D^\top\big[DD^\top\big]^{-1}\Big\{\big(q - \mathbb{E}[q]\big) - C\big(x - \mathbb{E}[x]\big)\Big\} + \Big(I - D^\top\big[DD^\top\big]^{-1}D\Big)\big(v - \mathbb{E}[v]\big). \end{cases}$$

Obviously the above defined $u(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$. Comparing to (3.9) and (3.11), it is clear that the derivations of (3.10) and (3.12) are still valid. Moreover, a direct calculation leads to

$$\widehat{D}\mathbb{E}[u] = \mathbb{E}[q] - \widehat{C}\mathbb{E}[x] \tag{3.15}$$

and

$$D\big(u - \mathbb{E}[u]\big) = \big(q - \mathbb{E}[q]\big) - C\big(x - \mathbb{E}[x]\big). \tag{3.16}$$

Then, (3.10), (3.12), (3.15) and (3.16) work together to ensure that (3.13) can be rewritten as (3.7) with the setting (3.14). This proves that the exact controllability of (3.5) implies that of (1.1). The whole proof is finished. □

In the direction of the above proposition, we turn to consider the exact controllability of (3.5). Besides the notion of exact controllability, the following notion of exact null-controllability is also needed.

Definition 3.2. The system (3.5) is said to be exactly null-controllable on $[0,T]$, if for any $x_0 \in \mathbb{R}^n$, there exists a pair of control processes $(q(\cdot), v(\cdot)) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n) \times L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$ steering the state $x(\cdot)$ from $x(0) = x_0$ to $x(T) = 0$.

Proposition 3.3. *For the system* (3.5)*, the exact controllability is equivalent to the exact null-controllability.*

Proof. The necessity is obvious, and we focus on the sufficiency. Let $(x_0, x_T) \in \mathbb{R}^n \times L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ be given. on the one hand, by the theory of MF-BSDEs (see [3], Thm. 3.1, [4], Thm. 3.1, or [23], Prop. 2.6 for example), there exists a unique solution

$$(x^0(\cdot), q^0(\cdot)) \equiv (x^0(\cdot;x_T), q^0(\cdot;x_T)) \in L^2_{\mathbb{F}}(\Omega;C(0,T;\mathbb{R}^n)) \times L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$$

satisfying the following MF-BSDE:

$$\begin{cases} \mathrm{d}x^0(t) = \Big\{\mathbf{A}(t)x^0(t) + \bar{\mathbf{A}}(t)\mathbb{E}[x^0(t)] + \mathbf{A}_1(t)q^0(t) + \bar{\mathbf{A}}_1(t)\mathbb{E}[q^0(t)]\Big\}\,\mathrm{d}t \\ \qquad\qquad + q^0(t)\,\mathrm{d}W(t), \quad t \in [0,T], \\ x^0(T) = x_T. \end{cases} \tag{3.17}$$

On the other hand, since the system (3.5) is exactly null-controllable, then there exists a triple $(\widetilde{x}(\cdot), \widetilde{q}(\cdot), v(\cdot))$ satisfying

$$\begin{cases} \mathrm{d}\widetilde{x}(t) = \Big\{\mathbf{A}(t)\widetilde{x}(t) + \bar{\mathbf{A}}(t)\mathbb{E}[\widetilde{x}(t)] + \mathbf{A}_1(t)\widetilde{q}(t) + \bar{\mathbf{A}}_1(t)\mathbb{E}[\widetilde{q}(t)] \\ \qquad\qquad + \mathbf{B}(t)v(t) + \bar{\mathbf{B}}(t)\mathbb{E}[v(t)]\Big\}\,\mathrm{d}t + \widetilde{q}(t)\,\mathrm{d}W(t), \quad t \in [0,T], \\ \widetilde{x}(0) = x_0 - x^0(0), \qquad \widetilde{x}(T) = 0. \end{cases} \tag{3.18}$$

Let

$$x(\cdot) = x^0(\cdot) + \widetilde{x}(\cdot), \qquad q(\cdot) = q^0(\cdot) + \tilde{q}(\cdot).$$

Then the linearity of equations implies that $(x(\cdot), q(\cdot), v(\cdot))$ satisfies

$$\begin{cases} \mathrm{d}x(t) = \Big\{\mathbf{A}(t)x(t) + \bar{\mathbf{A}}(t)\mathbb{E}[x(t)] + \mathbf{A}_1(t)q(t) + \bar{\mathbf{A}}_1(t)\mathbb{E}[q(t)] \\ \qquad\qquad + \mathbf{B}(t)v(t) + \bar{\mathbf{B}}(t)\mathbb{E}[v(t)]\Big\}\,\mathrm{d}t + q(t)\,\mathrm{d}W(t), \quad t \in [0,T], \\ x(0) = x_0, \qquad x(T) = x_T. \end{cases}$$

In other words, the pair $(q(\cdot), v(\cdot))$ steers $x(\cdot)$ from $x(0) = x_0$ to $x(T) = x_T$. The exact controllability of (3.5) follows from the arbitrariness of (x_0, x_T). □

Remark 3.4. We notice that, in the above proof of Proposition 3.3, the result of MF-BSDEs can be used because the system (3.5) satisfies the form of MF-BSDEs, *i.e.*, its diffusion term is $q(\cdot)$. In other words, the system (3.5) has an advantage: It can be regarded as either an MF-SDE or an MF-BSDE. However, this proof of Proposition 3.3 cannot be applied to the original system (1.1).

At the end of this section, we describe the results we have obtained in a more intuitive form. By adding the zero terminal condition into the system (3.5), we get the system (1.3) which has been proposed in Section 1. The addition of the zero terminal condition means that we consider the system (1.3) as an MF-BSDE. From the system (1.1) to the system (3.5), then to the system (1.3), we transform an MF-SDE to an MF-BSDE.

Definition 3.5. (i). The system (1.3) is said to be exactly controllable on $[0,T]$ if and only if the system (3.5) is exactly null-controllable on $[0,T]$.

(ii). Due to the linearity, the set

$$J := \Big\{x(0; v(\cdot)) \;\Big|\; v(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)\Big\}, \tag{3.19}$$

where $(x(\cdot; v(\cdot)), q(\cdot; v(\cdot))) \in L^2_{\mathbb{F}}(\Omega; C(0,T;\mathbb{R}^n)) \times L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$ denotes the unique solution to MF-BSDE (1.3) under the control process $v(\cdot)$ (see [23], Prop. 2.6), forms a linear subspace of $\mathbb{R}^n$. We call it the controllable subspace of the system (1.3).

Consequently, the system (1.3) is exactly controllable if and only if the corresponding controllable subspace $J = \mathbb{R}^n$. With the help of the theory of MF-BSDEs (see the proof of Prop. 3.3), we have obtained the result:

Theorem 3.6. *Under Assumption (H), the exact controllability of* (1.1) *is equivalent to that of* (1.3).

4. Controllability Gramian in time-variant case

Now, we introduce a matrix-valued linear MF-SDE:

$$\begin{cases} \mathrm{d}Y(t) = -\Big\{Y(t)\mathbf{A}(t) + \mathbb{E}[Y(t)]\bar{\mathbf{A}}(t)\Big\}\,\mathrm{d}t \\ \qquad\qquad -\Big\{Y(t)\mathbf{A}_1(t) + \mathbb{E}[Y(t)]\bar{\mathbf{A}}_1(t)\Big\}\,\mathrm{d}W(t), \quad t \in [0,T], \\ Y(0) = I. \end{cases} \tag{4.1}$$

Based on it, we define a symmetrical matrix:

$$\mathbf{G} = \mathbb{E}\int_0^T \Big\{Y(t)\mathbf{B}(t) + \mathbb{E}[Y(t)]\bar{\mathbf{B}}(t)\Big\}\Big\{\mathbf{B}(t)^\top Y(t)^\top + \bar{\mathbf{B}}(t)^\top \mathbb{E}[Y(t)^\top]\Big\}\,\mathrm{d}t. \tag{4.2}$$

This matrix can be regarded as a mean-field version of the traditional Gramian matrix. Similar to the classical theory, we have the following result which can be regarded as a mean-field version of Gramian criterion for exact controllability.

Theorem 4.1. *Let the space J be defined by (3.19) and the matrix $\mathbf{G}$ be defined by (4.2). Then,*

$$J = \text{Span } \mathbf{G}. \tag{4.3}$$

Consequently, the system (1.3) is exactly controllable if and only if $\mathbf{G}$ is non-singular.

We notice that, in (4.3), Span $\mathbf{G}$ denotes the vector space spanned by the column vectors of the matrix $\mathbf{G}$.

Proof. Clearly, it is equivalent to proving the following result: For any given vector $\beta \in \mathbb{R}^n$,

$$\beta \perp J \quad \text{if and only if} \quad \beta \perp \mathbf{G}. \tag{4.4}$$

For any $v(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$, we denote the solution to MF-BSDE (1.3) by $(x(\cdot), q(\cdot)) \equiv (x(\cdot; v(\cdot)), q(\cdot; v(\cdot)))$. Applying Itô's formula to $Y(\cdot)x(\cdot)$ yields

$$\begin{aligned} -x(0) = \mathbb{E}[Y(T)x(T)] - Y(0)x(0) &= \mathbb{E}\int_0^T \Big\{ Y(t)\mathbf{B}(t)v(t) + Y(t)\bar{\mathbf{B}}(t)\mathbb{E}[v(t)] \Big\}\, \mathrm{d}t \\ &= \mathbb{E}\int_0^T \Big\{ Y(t)\mathbf{B}(t) + \mathbb{E}[Y(t)]\bar{\mathbf{B}}(t) \Big\} v(t)\, \mathrm{d}t. \end{aligned}$$

Making the inner product with β leads to

$$-\beta^\top x(0) = \mathbb{E}\int_0^T \beta^\top \Big\{ Y(t)\mathbf{B}(t) + \mathbb{E}[Y(t)]\bar{\mathbf{B}}(t) \Big\} v(t)\, \mathrm{d}t. \tag{4.5}$$

(Necessity of (4.4)). Since $\beta \perp J$, then (4.5) implies

$$\mathbb{E}\int_0^T \beta^\top \Big\{ Y(t)\mathbf{B}(t) + \mathbb{E}[Y(t)]\bar{\mathbf{B}}(t) \Big\} v(t)\, \mathrm{d}t = 0 \quad \text{for any} \quad v(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^m).$$

In particular, taking $v(\cdot)$ to be each column of $\{\mathbf{B}(\cdot)^\top Y(\cdot)^\top + \bar{\mathbf{B}}(\cdot)^\top \mathbb{E}[Y(\cdot)^\top]\}$ yields $\beta^\top \mathbf{G} = 0$ by the definition of $\mathbf{G}$.

(Sufficiency of (4.4)). Since $\beta \perp \mathbf{G}$, then

$$\begin{aligned} 0 = \beta^\top \mathbf{G}\beta &= \mathbb{E}\int_0^T \beta^\top \Big\{ Y(t)\mathbf{B}(t) + \mathbb{E}[Y(t)]\bar{\mathbf{B}}(t) \Big\}\Big\{ \mathbf{B}(t)^\top Y(t)^\top + \bar{\mathbf{B}}(t)^\top \mathbb{E}[Y(t)^\top] \Big\}\beta\, \mathrm{d}t \\ &= \mathbb{E}\int_0^T \Big| \Big\{ \mathbf{B}(t)^\top Y(t)^\top + \bar{\mathbf{B}}(t)^\top \mathbb{E}[Y(t)^\top] \Big\}\beta \Big|^2 \mathrm{d}t. \end{aligned}$$

Therefore,

$$\Big\{ \mathbf{B}(t)^\top Y(t)^\top + \bar{\mathbf{B}}(t)^\top \mathbb{E}[Y(t)^\top] \Big\}\beta = 0 \quad \text{for almost all } t \in [0,T] \text{ and almost all } \omega \in \Omega.$$

Substituting the above equation into (4.5) leads to $\beta^\top x(0) = 0$. By the arbitrariness of $v(\cdot)$, we have $\beta \perp J$. The proof is completed. $\square$

The non-singularity of the matrix $\mathbf{G}$ defined by (4.2) provides a nice criterion for the exact controllability of (1.3). Like the classical case, we call G the *controllability Gramian*. However, the definition of $\mathbf{G}$ is based on the matrix-valued MF-SDE (4.1). Obviously, reducing the complexity of (4.1) will make the use of this criterion easier. The next corollary will try to reduce the MF-SDE (4.1) to an ODE, then provide a sufficient condition and a necessary condition, respectively.

By taking the expectation on (4.1), we get an ODE:

$$\begin{cases} \mathrm{d}\big(\mathbb{E}[Y(t)]\big) = -\mathbb{E}[Y(t)]\widehat{\mathbf{A}}(t)\,\mathrm{d}t, \quad t\in[0,T], \\ \mathbb{E}[Y(0)] = I. \end{cases} \tag{4.6}$$

Based on the solution $\mathbb{E}[Y(\cdot)]$ to ODE (4.6), we define a couple of symmetrical matrices:

$$\widetilde{\mathbf{G}} = \int_0^T \mathbb{E}[Y(t)]\widehat{\mathbf{A}}_1(t)\widehat{\mathbf{A}}_1^\top(t)\mathbb{E}[Y(t)^\top]\,\mathrm{d}t \tag{4.7}$$

and

$$\widehat{\mathbf{G}} = \int_0^T \mathbb{E}[Y(t)]\widehat{\mathbf{B}}(t)\widehat{\mathbf{B}}(t)^\top\mathbb{E}[Y(t)^\top]\,\mathrm{d}t. \tag{4.8}$$

We notice that (4.6) and $\widehat{\mathbf{G}}$ are different from (2.4) and $\widehat{\Psi}$ since $\widehat{\mathbf{A}}(\cdot)$ and $\widehat{\mathbf{B}}(\cdot)$ are different from $\widehat{A}(\cdot)$ and $\widehat{B}(\cdot)$, respectively.

Corollary 4.2. *(i). If $\widehat{\mathbf{G}}$ is non-singular, then the system (1.3) is exactly controllable on $[0,T]$.*
(ii). If the system (1.3) is exactly controllable on $[0,T]$, then $(\widetilde{\mathbf{G}}+\widehat{\mathbf{G}})$ is non-singular.

Proof. (i). A direct calculation leads to

$$\begin{aligned} &\mathbb{E}\Big[\Big\{Y\mathbf{B}+\mathbb{E}[Y]\bar{\mathbf{B}}\Big\}\Big\{\mathbf{B}^\top Y^\top+\bar{\mathbf{B}}^\top\mathbb{E}[Y^\top]\Big\}\Big] \\ &=\mathbb{E}\Big[\Big\{\big(Y-\mathbb{E}[Y]\big)\mathbf{B}+\mathbb{E}[Y]\widehat{\mathbf{B}}\Big\}\Big\{\mathbf{B}^\top\big(Y-\mathbb{E}[Y]\big)^\top+\widehat{\mathbf{B}}^\top\mathbb{E}[Y^\top]\Big\}\Big] \\ &=\mathbb{E}\Big[\big(Y-\mathbb{E}[Y]\big)\mathbf{B}\mathbf{B}^\top\big(Y-\mathbb{E}[Y]\big)^\top\Big]+\mathbb{E}[Y]\widehat{\mathbf{B}}\widehat{\mathbf{B}}^\top\mathbb{E}[Y^\top] \\ &\geq \mathbb{E}[Y]\widehat{\mathbf{B}}\widehat{\mathbf{B}}^\top\mathbb{E}[Y^\top]. \end{aligned}$$

Here, for any symmetrical matrices $\mathbb{A}$ and $\mathbb{B}$, the inequality $\mathbb{A}\geq\mathbb{B}$ means that $(\mathbb{A}-\mathbb{B})$ is positive semi-definite. Then, by the definitions, $\mathbf{G}\geq\widehat{\mathbf{G}}$. Consequently, the non-singularity of $\widehat{\mathbf{G}}$ implies that of $\mathbf{G}$. Theorem 4.1 works to obtain the exact controllability of the system (1.3).

(ii). Taking the expectation on the system (3.5) yields

$$\mathrm{d}\big(\mathbb{E}[x(t)]\big) = \Big\{\widehat{\mathbf{A}}(t)\mathbb{E}[x(t)]+\widehat{\mathbf{A}}_1(t)\mathbb{E}[q(t)]+\widehat{\mathbf{B}}(t)\mathbb{E}[v(t)]\Big\}\,\mathrm{d}t, \quad t\in[0,T].$$

We notice that the pair $(\mathbb{E}[q(\cdot)],\mathbb{E}[v(\cdot)])$ is regarded as the control for the above deterministic system. Proposition 2.2-(i) shows that, when the system (3.5) is exactly controllable, then the matrix

$$\int_0^T \mathbb{E}[Y(t)]\begin{pmatrix}\widehat{\mathbf{A}}_1(t) & \widehat{\mathbf{B}}(t)\end{pmatrix}\begin{pmatrix}\widehat{\mathbf{A}}_1(t)^\top \\ \widehat{\mathbf{B}}(t)^\top\end{pmatrix}\mathbb{E}[Y(t)^\top]\,\mathrm{d}t = \widetilde{\mathbf{G}}+\widehat{\mathbf{G}}$$

(refer to the definitions (4.7) and (4.8)) is non-singular. Proposition 3.3 together with Definition 3.5-(i) shows that the exact controllability of (3.5) is equivalent to that of (1.3), which completes the proof. □

5. Kalman rank condition in time-invariant case

In this section, all the coefficients in the system (1.3) are restricted to be time-invariant, *i.e.*, $\mathbf{A}(\cdot)\equiv\mathbf{A}$, $\bar{\mathbf{A}}(\cdot)\equiv\bar{\mathbf{A}}$, $\mathbf{A}_1(\cdot)\equiv\mathbf{A}_1$, $\bar{\mathbf{A}}_1(\cdot)\equiv\bar{\mathbf{A}}_1$, $\mathbf{B}(\cdot)\equiv\mathbf{B}$, $\bar{\mathbf{B}}(\cdot)\equiv\bar{\mathbf{B}}$, and then $\widehat{\mathbf{A}}(\cdot)\equiv\widehat{\mathbf{A}}$, $\widehat{\mathbf{A}}_1(\cdot)\equiv\widehat{\mathbf{A}}_1$, $\widehat{\mathbf{B}}(\cdot)\equiv\widehat{\mathbf{B}}$. Obviously, this restriction can be implied by the setting: Let all the coefficients in the original system (1.1) be time-invariant. For convenience, we rewrite the time-invariant case of (1.3) as follows:

$$\begin{cases} \mathrm{d}x(t)=\Big\{\mathbf{A}x(t)+\bar{\mathbf{A}}\mathbb{E}[x(t)]+\mathbf{A}_1q(t)+\bar{\mathbf{A}}_1\mathbb{E}[q(t)]+\mathbf{B}v(t)+\bar{\mathbf{B}}\mathbb{E}[v(t)]\Big\}\,\mathrm{d}t\\ \qquad\qquad +q(t)\,\mathrm{d}W(t),\quad t\in[0,T],\\ x(T)=0. \end{cases}\tag{5.1}$$

Compared with the controllability Gramian studied in the previous section, this section is devoted to looking for a simpler criterion, named a Kalman rank condition, for the exact controllability of the time-invariant system (5.1).

As a start, we combine the classical product and the tensor product of matrices to define a new multiplication operation.

Definition 5.1. For any given block matrices

$$\mathbb{A}=\begin{pmatrix}\mathbb{A}_{11}&\mathbb{A}_{12}&\cdots&\mathbb{A}_{1n}\\ \mathbb{A}_{21}&\mathbb{A}_{22}&\cdots&\mathbb{A}_{2n}\\ \vdots&\vdots&\ddots&\vdots\\ \mathbb{A}_{m1}&\mathbb{A}_{m2}&\cdots&\mathbb{A}_{mn}\end{pmatrix}\quad\text{and}\quad \mathbb{B}=\begin{pmatrix}\mathbb{B}_{11}&\mathbb{B}_{12}&\cdots&\mathbb{B}_{1q}\\ \mathbb{B}_{21}&\mathbb{B}_{22}&\cdots&\mathbb{B}_{2q}\\ \vdots&\vdots&\ddots&\vdots\\ \mathbb{B}_{p1}&\mathbb{B}_{p2}&\cdots&\mathbb{B}_{pq}\end{pmatrix},\tag{5.2}$$

their block-tensor product is defined as follows:

$$\mathbb{A}\otimes\mathbb{B}=\begin{pmatrix}
\mathbb{A}_{11}\mathbb{B}_{11}&\mathbb{A}_{11}\mathbb{B}_{12}&\cdots&\mathbb{A}_{11}\mathbb{B}_{1q}&\cdots&\cdots&\mathbb{A}_{1n}\mathbb{B}_{11}&\mathbb{A}_{1n}\mathbb{B}_{12}&\cdots&\mathbb{A}_{1n}\mathbb{B}_{1q}\\
\mathbb{A}_{11}\mathbb{B}_{21}&\mathbb{A}_{11}\mathbb{B}_{22}&\cdots&\mathbb{A}_{11}\mathbb{B}_{2q}&\cdots&\cdots&\mathbb{A}_{1n}\mathbb{B}_{21}&\mathbb{A}_{1n}\mathbb{B}_{22}&\cdots&\mathbb{A}_{1n}\mathbb{B}_{2q}\\
\vdots&\vdots&\ddots&\vdots&&&\vdots&\vdots&\ddots&\vdots\\
\mathbb{A}_{11}\mathbb{B}_{p1}&\mathbb{A}_{11}\mathbb{B}_{p2}&\cdots&\mathbb{A}_{11}\mathbb{B}_{pq}&\cdots&\cdots&\mathbb{A}_{1n}\mathbb{B}_{p1}&\mathbb{A}_{1n}\mathbb{B}_{p2}&\cdots&\mathbb{A}_{1n}\mathbb{B}_{pq}\\
\vdots&\vdots&&\vdots&\ddots&&\vdots&\vdots&&\vdots\\
\vdots&\vdots&&\vdots&&\ddots&\vdots&\vdots&&\vdots\\
\mathbb{A}_{m1}\mathbb{B}_{11}&\mathbb{A}_{m1}\mathbb{B}_{12}&\cdots&\mathbb{A}_{m1}\mathbb{B}_{1q}&\cdots&\cdots&\mathbb{A}_{mn}\mathbb{B}_{11}&\mathbb{A}_{mn}\mathbb{B}_{12}&\cdots&\mathbb{A}_{mn}\mathbb{B}_{1q}\\
\mathbb{A}_{m1}\mathbb{B}_{21}&\mathbb{A}_{m1}\mathbb{B}_{22}&\cdots&\mathbb{A}_{m1}\mathbb{B}_{2q}&\cdots&\cdots&\mathbb{A}_{mn}\mathbb{B}_{21}&\mathbb{A}_{mn}\mathbb{B}_{22}&\cdots&\mathbb{A}_{mn}\mathbb{B}_{2q}\\
\vdots&\vdots&\ddots&\vdots&&&\vdots&\vdots&\ddots&\vdots\\
\mathbb{A}_{m1}\mathbb{B}_{p1}&\mathbb{A}_{m1}\mathbb{B}_{p2}&\cdots&\mathbb{A}_{m1}\mathbb{B}_{pq}&\cdots&\cdots&\mathbb{A}_{mn}\mathbb{B}_{p1}&\mathbb{A}_{mn}\mathbb{B}_{p2}&\cdots&\mathbb{A}_{mn}\mathbb{B}_{pq}
\end{pmatrix}\tag{5.3}$$

provided all the involved classical products $\mathbb{A}_{ij}\mathbb{B}_{hl}$ are well-posed ($i=1,2,\dots,m$; $j=1,2,\dots,n$; $h=1,2,\dots,p$; $l=1,2,\dots,q$).

We notice that, the block-tensor product (5.3) depends on the block form (5.2) of matrices $\mathbb{A}$ and $\mathbb{B}$. Different block forms will bring different calculation results.

Remark 5.2. In our analysis below, the following two simple properties will be used.

(i). The above defined block-tensor product obeys the following combination rule:

$$(\mathbb{A}\otimes\mathbb{B})\otimes\mathbb{C}=\mathbb{A}\otimes(\mathbb{B}\otimes\mathbb{C})=:\mathbb{A}\otimes\mathbb{B}\otimes\mathbb{C}.$$

For simplicity, the following notation will also be introduced:

$$\overbrace{\mathbb{A}\otimes\mathbb{A}\otimes\cdots\otimes\mathbb{A}}^{\text{the number is } k}=:\mathbb{A}^{\otimes k}.$$

(ii). We have the relationships:

$$\big(\mathbb{A}_{11}\otimes\mathbb{B}\big)^\top=\mathbb{B}^\top\otimes\mathbb{A}_{11}^\top \quad\text{and}\quad \big(\mathbb{A}\otimes\mathbb{B}_{11}\big)^\top=\mathbb{B}_{11}^\top\otimes\mathbb{A}^\top.$$

With the help of the above defined block-tensor operation, based on the coefficients $\mathbf{A}$, $\widehat{\mathbf{A}}$, $\mathbf{A}_1$, $\widehat{\mathbf{A}}_1$, $\mathbf{B}$, and $\widehat{\mathbf{B}}$ of the system (5.1), we recursively define

$$\begin{cases}\mathcal{A}_1:=\widehat{\mathbf{B}},\\ \mathcal{A}_2:=\Big(\ \widehat{\mathbf{A}}\otimes\mathcal{A}_1,\quad \widehat{\mathbf{A}}_1\otimes\mathbf{B}\ \Big),\\ \mathcal{A}_k:=\Big(\ \widehat{\mathbf{A}}\otimes\mathcal{A}_{k-1},\quad \widehat{\mathbf{A}}_1\otimes\big(\ \mathbf{A},\quad \mathbf{A}_1\ \big)^{\otimes(k-2)}\otimes\mathbf{B}\ \Big),\quad k=3,4,\dots\end{cases}\tag{5.4}$$

Here and hereafter, for the block-tensor operation, the matrices $\mathbf{A}$, $\widehat{\mathbf{A}}$, $\mathbf{A}_1$, $\widehat{\mathbf{A}}_1$, $\mathbf{B}$, $\widehat{\mathbf{B}}$, and their products are the basic block units. We will not repeat this point since it does not cause confusion. Clearly, we have

$$\begin{aligned}\mathcal{A}_k=\Bigg(\widehat{\mathbf{A}}^{k-1}\widehat{\mathbf{B}},\quad \widehat{\mathbf{A}}^{k-2}\widehat{\mathbf{A}}_1\mathbf{B},\quad &\overbrace{\dots,\quad \widehat{\mathbf{A}}^{k-2-i}\widehat{\mathbf{A}}_1\otimes\big(\mathbf{A},\quad \mathbf{A}_1\big)^{\otimes i}\otimes\mathbf{B},\quad \cdots}^{i=1,2,\dots,k-1}\\ &\widehat{\mathbf{A}}_1\otimes\big(\mathbf{A},\quad \mathbf{A}_1\big)^{\otimes(k-2)}\otimes\mathbf{B}\Big),\qquad k=2,3,\dots\end{aligned}\tag{5.5}$$

For more intuitiveness, besides $\mathcal{A}_1=\widehat{\mathbf{B}}$ given in the above (5.4), we would like to list also the calculation results of $\mathcal{A}_2$, $\mathcal{A}_3$, and $\mathcal{A}_4$ as follows:

$$\begin{aligned}\mathcal{A}_2&=\Big(\ \widehat{\mathbf{A}}\widehat{\mathbf{B}},\quad \widehat{\mathbf{A}}_1\mathbf{B}\ \Big),\\ \mathcal{A}_3&=\Big(\ \widehat{\mathbf{A}}\widehat{\mathbf{A}}\widehat{\mathbf{B}},\quad \widehat{\mathbf{A}}\widehat{\mathbf{A}}_1\mathbf{B},\quad \widehat{\mathbf{A}}_1\mathbf{A}\mathbf{B},\quad \widehat{\mathbf{A}}_1\mathbf{A}_1\mathbf{B}\ \Big),\\ \mathcal{A}_4&=\Big(\ \widehat{\mathbf{A}}\widehat{\mathbf{A}}\widehat{\mathbf{A}}\widehat{\mathbf{B}},\quad \widehat{\mathbf{A}}\widehat{\mathbf{A}}\widehat{\mathbf{A}}_1\mathbf{B},\quad \widehat{\mathbf{A}}\widehat{\mathbf{A}}_1\mathbf{A}\mathbf{B},\quad \widehat{\mathbf{A}}\widehat{\mathbf{A}}_1\mathbf{A}_1\mathbf{B},\\ &\qquad \widehat{\mathbf{A}}_1\mathbf{A}\mathbf{A}\mathbf{B},\quad \widehat{\mathbf{A}}_1\mathbf{A}\mathbf{A}_1\mathbf{B},\quad \widehat{\mathbf{A}}_1\mathbf{A}_1\mathbf{A}\mathbf{B},\quad \widehat{\mathbf{A}}_1\mathbf{A}_1\mathbf{A}_1\mathbf{B}\ \Big).\end{aligned}\tag{5.6}$$

The matrices $\mathcal{A}_k$ $(k=1,2,\dots)$ will be used to construct the desired Kalman rank condition below. Now, let us do some preparations on them.

Lemma 5.3. *(i). Let*

$$\mathcal{B}_1:=\mathbf{B},\qquad \begin{aligned}\mathcal{B}_k&:=\big(\ \mathbf{A},\quad \mathbf{A}_1\ \big)^{\otimes(k-1)}\otimes\mathbf{B}\\ &=\big(\ \mathbf{A}\otimes\mathcal{B}_{k-1},\quad \mathbf{A}_1\otimes\mathcal{B}_{k-1}\ \big)\end{aligned},\qquad k=2,3,\dots\tag{5.7}$$

Then, the following recursive relation hold:

$$\mathcal{A}_k = \left(\widehat{\mathbf{A}} \otimes \mathcal{A}_{k-1}, \quad \widehat{\mathbf{A}}_1 \otimes \mathcal{B}_{k-1} \right), \qquad k = 2, 3, \ldots \tag{5.8}$$

(ii). Let

$$\mathcal{C}_2 := \widehat{\mathbf{A}}_1, \qquad \begin{aligned} \mathcal{C}_k &:= \left(\widehat{\mathbf{A}} \otimes \mathcal{C}_{k-1}, \quad \widehat{\mathbf{A}}_1 \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right)^{\otimes(k-2)} \right) \\ &= \left(\widehat{\mathbf{A}}^{k-2}\widehat{\mathbf{A}}_1, \quad \mathcal{C}_{k-1} \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right) \right) \end{aligned}, \qquad k = 3, 4, \ldots \tag{5.9}$$

Then, $\mathcal{A}_k$ can be represented by $\mathcal{C}_k$ as

$$\mathcal{A}_k = \left(\widehat{\mathbf{A}}^{k-1}\widehat{\mathbf{B}}, \quad \mathcal{C}_k \otimes \mathbf{B} \right), \quad k = 2, 3, \ldots \tag{5.10}$$

Proof. (i). The equal sign in (5.7) can be easily derived by virtue of Definition 5.1 and Remark 5.2-(i). The relation (5.8) is obvious.

(ii). By an induction argument, the relation (5.10) can be easily obtained from the definition in (5.9). The remaining thing is to prove the equal sign in (5.9). In fact, the induction method will be employed for this aim.

Firstly, when $k = 3$, by the definitions of $\mathcal{C}_3$ and $\mathcal{C}_2$, we verify

$$\mathcal{C}_3 = \left(\widehat{\mathbf{A}} \otimes \mathcal{C}_2, \quad \widehat{\mathbf{A}}_1 \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right) \right) = \left(\widehat{\mathbf{A}}\widehat{\mathbf{A}}_1, \quad \mathcal{C}_2 \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right) \right),$$

i.e., the conclusion holds for $k = 3$.

Secondly, we assume that the conclusion holds for $k - 1$, *i.e.*,

$$\mathcal{C}_{k-1} = \left(\widehat{\mathbf{A}}^{k-3}\widehat{\mathbf{A}}_1, \quad \mathcal{C}_{k-2} \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right) \right). \tag{5.11}$$

Next we try to prove that the conclusion also holds for k. Indeed, by the definition of $\mathcal{C}_{k-1}$, we have ("RHS" is short for "the right hand side of the equal sign in (5.9)")

$$\begin{aligned} \text{RHS} &= \left(\widehat{\mathbf{A}}^{k-2}\widehat{\mathbf{A}}_1, \quad \mathcal{C}_{k-1} \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right) \right) \\ &= \left(\widehat{\mathbf{A}}^{k-2}\widehat{\mathbf{A}}_1, \quad \left(\widehat{\mathbf{A}} \otimes \mathcal{C}_{k-2}, \quad \widehat{\mathbf{A}}_1 \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right)^{\otimes(k-3)} \right) \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right) \right) \\ &= \left(\widehat{\mathbf{A}}^{k-2}\widehat{\mathbf{A}}_1, \quad \widehat{\mathbf{A}} \otimes \mathcal{C}_{k-2} \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right), \quad \widehat{\mathbf{A}}_1 \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right)^{\otimes(k-3)} \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right) \right) \\ &= \left(\widehat{\mathbf{A}} \otimes \left(\widehat{\mathbf{A}}^{k-3}\widehat{\mathbf{A}}_1, \quad \mathcal{C}_{k-2} \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right) \right), \quad \widehat{\mathbf{A}}_1 \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right)^{\otimes(k-2)} \right). \end{aligned}$$

By (5.11) and the definition of $\mathcal{C}_k$, we obtain

$$\text{RHS} = \left(\widehat{\mathbf{A}} \otimes \mathcal{C}_{k-1}, \quad \widehat{\mathbf{A}}_1 \otimes \left(\mathbf{A}, \quad \mathbf{A}_1 \right)^{\otimes(k-2)} \right) = \text{LHS},$$

where "LHS" is short for "the left hand side of the equal sign in (5.9)". The proof is completed. □

As the same style as Theorem 4.1, we propose

Theorem 5.4. *Let J defined by (3.19) be the controllable subspace of the system (5.1). Let $\mathcal{A}_k$ $(k=1,2,\dots)$ be defined by (5.4). Then*

$$J = \text{Span}\,\big(\mathcal{A}_1,\ \mathcal{A}_2,\ \mathcal{A}_3,\ \dots\big). \tag{5.12}$$

Consequently, the system (5.1) is exactly controllable if and only if

$$\text{Rank}\,\big(\mathcal{A}_1,\ \mathcal{A}_2,\ \mathcal{A}_3,\ \dots\big) = n. \tag{5.13}$$

Proof. It is equivalent to proving the following result: For any given $\beta \in \mathbb{R}^n$,

$$\beta \perp J \quad \text{if and only if} \quad \beta \perp \big(\mathcal{A}_1,\ \mathcal{A}_2,\ \mathcal{A}_3,\ \dots\big). \tag{5.14}$$

We split the whole proof into three steps.

Step 1: $\boldsymbol{\beta \perp J}$ is equivalent to (5.18) below.

Inspired by Peng [17], for any $\beta \in \mathbb{R}^n$, we introduce a linear MF-SDE which is called a dual equation to the system (5.1):

$$\begin{cases} \mathrm{d}p(t) = -\Big\{\mathbf{A}^\top p(t) + \bar{\mathbf{A}}^\top \mathbb{E}[p(t)]\Big\}\,\mathrm{d}t - \Big\{\mathbf{A}_1^\top p(t) + \bar{\mathbf{A}}_1^\top \mathbb{E}[p(t)]\Big\}\,\mathrm{d}W(t), \quad t \in [0,T], \\ p(0) = \beta. \end{cases} \tag{5.15}$$

Clearly, it is equivalent to the following system (the argument s is suppressed for simplicity):

$$\begin{cases} \mathbb{E}[p(t)] = \beta - \displaystyle\int_0^t \widehat{\mathbf{A}}^\top \mathbb{E}[p]\,\mathrm{d}s, \quad t \in [0,T] \\ p(t) - \mathbb{E}[p(t)] = -\displaystyle\int_0^t \mathbf{A}^\top\big(p - \mathbb{E}[p]\big)\,\mathrm{d}s - \int_0^t \Big\{\widehat{\mathbf{A}}_1^\top \mathbb{E}[p] + \mathbf{A}_1^\top\big(p - \mathbb{E}[p]\big)\Big\}\,\mathrm{d}W, \quad t \in [0,T]. \end{cases} \tag{5.16}$$

By applying Itô's formula to $\langle p(\cdot),\ x(\cdot)\rangle$ where $(x(\cdot), q(\cdot)) \equiv (x(\cdot; v(\cdot)), q(\cdot; v(\cdot)))$ denotes the unique solution to MF-BSDE (5.1), we have

$$\mathbb{E}\int_0^T \Big\langle v(t),\ \mathbf{B}^\top p(t) + \bar{\mathbf{B}}^\top \mathbb{E}[p(t)]\Big\rangle\,\mathrm{d}t = -\big\langle \beta,\ x(0; v(\cdot))\big\rangle.$$

Therefore, $\beta \perp J$ if and only if

$$\mathbb{E}\int_0^T \Big\langle v(t),\ \mathbf{B}^\top p(t) + \bar{\mathbf{B}}^\top \mathbb{E}[p(t)]\Big\rangle\,\mathrm{d}t = 0, \quad \forall\ v(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^m),$$

which is also equivalent to

$$\mathbf{B}^\top p(\cdot) + \bar{\mathbf{B}}^\top \mathbb{E}[p(\cdot)] = 0. \tag{5.17}$$

In other words,

$$\widehat{\mathbf{B}}^\top \mathbb{E}[p(\cdot)] = 0 \quad \text{and} \quad \mathbf{B}^\top\big(p(\cdot) - \mathbb{E}[p(\cdot)]\big) = 0. \tag{5.18}$$

Step 2: (5.18) implies $\boldsymbol{\beta \perp (\mathcal{A}_1,\ \mathcal{A}_2,\ \mathcal{A}_3,\ \dots)}$.

By Definitions (5.4) and (5.7), we rewrite (5.18) as

$$\mathcal{A}_1^\top \otimes \mathbb{E}[p(\cdot)] = 0, \qquad \mathcal{B}_1^\top \otimes \big(p(\cdot) - \mathbb{E}[p(\cdot)]\big) = 0. \tag{5.19}$$

Then, (5.19) and (5.16) lead to

$$\begin{cases} \mathcal{A}_1^\top \otimes \beta - \displaystyle\int_0^t \mathcal{A}_1^\top \otimes \widehat{\mathbf{A}}^\top \mathbb{E}[p]\,\mathrm{d}s = 0, \\ -\displaystyle\int_0^t \mathcal{B}_1^\top \otimes \mathbf{A}^\top \big(p - \mathbb{E}[p]\big)\,\mathrm{d}s - \int_0^t \Big\{\mathcal{B}_1^\top \otimes \widehat{\mathbf{A}}_1^\top \mathbb{E}[p] + \mathcal{B}_1^\top \otimes \mathbf{A}_1^\top \big(p - \mathbb{E}[p]\big)\Big\}\,\mathrm{d}W = 0. \end{cases}$$

By noticing Remark 5.2-(ii), (5.8) and (5.7), we have

$$\begin{aligned} \begin{pmatrix} \mathcal{A}_1^\top \otimes \widehat{\mathbf{A}}^\top \\ \mathcal{B}_1^\top \otimes \widehat{\mathbf{A}}_1^\top \end{pmatrix} &= \begin{pmatrix} \big(\mathcal{A}_1^\top \otimes \widehat{\mathbf{A}}^\top\big)^\top, & \big(\mathcal{B}_1^\top \otimes \widehat{\mathbf{A}}_1^\top\big)^\top \end{pmatrix}^\top \\ &= \begin{pmatrix} \widehat{\mathbf{A}} \otimes \mathcal{A}_1, & \widehat{\mathbf{A}}_1 \otimes \mathcal{B}_1 \end{pmatrix}^\top = \mathcal{A}_2^\top \end{aligned}$$

and

$$\begin{pmatrix} \mathcal{B}_1^\top \otimes \mathbf{A}^\top \\ \mathcal{B}_1^\top \otimes \mathbf{A}_1^\top \end{pmatrix} = \begin{pmatrix} \mathbf{A} \otimes \mathcal{B}_1, & \mathbf{A}_1 \otimes \mathcal{B}_1 \end{pmatrix}^\top = \mathcal{B}_2^\top.$$

Therefore,

$$\mathcal{A}_1^\top \otimes \beta = 0 \tag{5.20}$$

and

$$\mathcal{A}_2^\top \otimes \mathbb{E}[p(\cdot)] = 0, \qquad \mathcal{B}_2^\top \otimes \big(p(\cdot) - \mathbb{E}[p(\cdot)]\big) = 0. \tag{5.21}$$

The above procedure can be repeated infinitely. With the help of the induction method, we obtain

$$\mathcal{A}_k^\top \otimes \beta = 0, \quad k = 1, 2, \ldots,$$

which is equivalent to $\beta \perp (\mathcal{A}_1,\ \mathcal{A}_2,\ \mathcal{A}_3,\ \ldots)$.

Step 3: $\boldsymbol{\beta \perp (\mathcal{A}_1,\ \mathcal{A}_2,\ \mathcal{A}_3,\ \ldots)}$ implies (5.18).

We introduce a matrix-valued MF-SDE:

$$\begin{cases} \mathrm{d}\Phi(t) = -\Big\{\mathbf{A}^\top \Phi(t) + \bar{\mathbf{A}}^\top \mathbb{E}[\Phi(t)]\Big\}\,\mathrm{d}t - \Big\{\mathbf{A}_1^\top \Phi(t) + \bar{\mathbf{A}}_1^\top \mathbb{E}[\Phi(t)]\Big\}\,\mathrm{d}W(t), \quad t \in [0,T], \\ \Phi(0) = I, \end{cases} \tag{5.22}$$

which is equivalent to (the argument s is suppressed)

$$\begin{cases} \mathbb{E}[\Phi(t)] = I - \displaystyle\int_0^t \widehat{\mathbf{A}}^\top \mathbb{E}[\Phi]\,\mathrm{d}s, \quad t \in [0,T] \\ \Phi(t) - \mathbb{E}[\Phi(t)] = -\displaystyle\int_0^t \mathbf{A}^\top \big(\Phi - \mathbb{E}[\Phi]\big)\,\mathrm{d}s - \int_0^t \Big\{\widehat{\mathbf{A}}_1^\top \mathbb{E}[\Phi] + \mathbf{A}_1^\top \big(\Phi - \mathbb{E}[\Phi]\big)\Big\}\,\mathrm{d}W, \quad t \in [0,T]. \end{cases} \tag{5.23}$$

It is clear that $\Phi(\cdot)$ is indeed the fundamental matrix to SDE (5.15). Then we have $p(\cdot)=\Phi(\cdot)\beta$. Therefore, (5.18) becomes

$$\widehat{\mathbf{B}}^\top\mathbb{E}[\Phi(\cdot)]\beta=0\quad\text{and}\quad\mathbf{B}^\top\big(\Phi(\cdot)-\mathbb{E}[\Phi(\cdot)]\big)\beta=0. \tag{5.24}$$

We continue to recursively define

$$\Phi^{(1)}(\cdot)\equiv I \tag{5.25}$$

and

$$\begin{cases}\mathrm{d}\Phi^{(k)}(t)=-\Big\{\mathbf{A}^\top\Phi^{(k-1)}(t)+\bar{\mathbf{A}}^\top\mathbb{E}[\Phi^{(k-1)}(t)]\Big\}\,\mathrm{d}t\\ \qquad\qquad -\Big\{\mathbf{A}_1^\top\Phi^{(k-1)}(t)+\bar{\mathbf{A}}_1^\top\mathbb{E}[\Phi^{(k-1)}(t)]\Big\}\,\mathrm{d}W(t),\quad t\in[0,T],\\ \Phi^{(k)}(0)=I,\end{cases}\qquad k=2,3,\dots \tag{5.26}$$

Clearly, the above (5.26) is equivalent to

$$\begin{cases}\mathbb{E}[\Phi^{(k)}(t)]=I-\displaystyle\int_0^t\widehat{\mathbf{A}}^\top\mathbb{E}[\Phi^{(k-1)}]\,\mathrm{d}s,\quad t\in[0,T],\\ \Phi^{(k)}(t)-\mathbb{E}[\Phi^{(k)}(t)]=-\displaystyle\int_0^t\widehat{\mathbf{A}}_1^\top\mathbb{E}[\Phi^{(k-1)}]\,\mathrm{d}W\\ \qquad -\displaystyle\int_0^t\mathbf{A}^\top\big(\Phi^{(k-1)}-\mathbb{E}[\Phi^{(k-1)}]\big)\,\mathrm{d}s\\ \qquad -\displaystyle\int_0^t\mathbf{A}_1^\top\big(\Phi^{(k-1)}-\mathbb{E}[\Phi^{(k-1)}]\big)\,\mathrm{d}W,\quad t\in[0,T],\end{cases}\qquad k=2,3,\dots \tag{5.27}$$

Due to the theory of SDEs, it is clear that $\Phi^{(k)}(\cdot)\to\Phi(\cdot)$ in $L^2_{\mathbb{F}}(\Omega;C(0,T;\mathbb{R}^{n\times n}))$ as $k\to\infty$. Therefore, (5.24) (equivalently (5.18)) can be implied by

$$\widehat{\mathbf{B}}^\top\mathbb{E}[\Phi^{(k)}(\cdot)]\beta=0,\quad\mathbf{B}^\top\big(\Phi^{(k)}(\cdot)-\mathbb{E}[\Phi^{(k)}(\cdot)]\big)\beta=0,\qquad\text{for any }k=1,2,\dots \tag{5.28}$$

Corollary 5.6 below shows that $\beta\perp(\mathcal{A}_1,\ \mathcal{A}_2,\ \mathcal{A}_3,\ \dots)$ implies the above (5.28). Then, the proof is completed. However, the proofs of Corollary 5.6 and its main argument: Lemma 5.5 are a bit complicated. We would like to split them out of this proof. □

Let us introduce some notations and do some preparatory works for Lemma 5.5 and Corollary 5.6 below. We define a pair of scalar-valued processes:

$$\eta^{(1)}_{(0)}(t):=\int_0^t(-1)\,\mathrm{d}s,\quad\eta^{(1)}_{(1)}(t):=\int_0^t(-1)\,\mathrm{d}W(s),\qquad t\in[0,T].$$

It is clear that $\eta^{(1)}_{(0)}(\cdot),\ \eta^{(1)}_{(1)}(\cdot)\in L^2_{\mathbb{F}}(\Omega;C(0,T;\mathbb{R}))$. We continue to define four scalar-valued processes as follows:

$$\begin{cases}\eta^{(2)}_{(00)}(t):=\displaystyle\int_0^t\Big(-\eta^{(1)}_{(0)}(s)\Big)\,\mathrm{d}s,\quad\eta^{(2)}_{(01)}(t):=\int_0^t\Big(-\eta^{(1)}_{(0)}(s)\Big)\,\mathrm{d}W(s),\\ \eta^{(2)}_{(10)}(t):=\displaystyle\int_0^t\Big(-\eta^{(1)}_{(1)}(s)\Big)\,\mathrm{d}s,\quad\eta^{(2)}_{(11)}(t):=\int_0^t\Big(-\eta^{(1)}_{(1)}(s)\Big)\,\mathrm{d}W(s),\end{cases}\qquad t\in[0,T].$$

It is clear that the above defined four processes also belong to the space $L^2_{\mathbb{F}}(\Omega; C(0,T;\mathbb{R}))$. We notice the naming rules: (i) The number of integrals is written in the superscript of η; (ii) When we have a Lebesgue's integral of time s, we write a 0 in the subscript; Symmetrically, when we have a Itô's integral of Brownian motion $W(s)$, we write a 1 in the subscript; (iii) For the early integral, the corresponding 0 or 1 appears on the left.

The above recursive definitions can be repeated infinitely. However, a large superscript will cause a bad result: The related subscript will be too long to be convenient to use. Due to this, we regard the string of numbers consisting of 0 and 1 as a binary integer, and rewrite it as its decimal integer form. For example, $\eta^{(4)}_{(1001)}(\cdot)$ is rewritten as $\eta^{(4)}_9(\cdot)$. With these new notations, the recursive definitions are re-given as follows:

$$\begin{cases} \eta^{(1)}_0(t) := \int_0^t (-1)\,\mathrm{d}s, \quad \eta^{(1)}_1(t) := \int_0^t (-1)\,\mathrm{d}W(s), \qquad t\in[0,T], \\ \eta^{(k+1)}_{2j}(t) := \int_0^t \Big(-\eta^{(k)}_j(s)\Big)\,\mathrm{d}s, \quad \eta^{(k+1)}_{2j+1}(t) := \int_0^t \Big(-\eta^{(k)}_j(s)\Big)\,\mathrm{d}W(s), \qquad t\in[0,T], \\ \qquad\qquad k=1,2,\ldots, \quad j=0,1,2,\ldots,2^k-1. \end{cases} \tag{5.29}$$

Moreover, we denote

$$\begin{aligned} \eta^{(k)}(\cdot) &:= \Big(\eta^{(k)}_0(\cdot),\ \eta^{(k)}_1(\cdot),\ \ldots,\ \eta^{(k)}_{2^k-1}(\cdot)\Big) \\ &=: \Big(\eta^{(k)}_0(\cdot),\ \eta^{(k)}_{-0}\Big), \end{aligned} \qquad k=1,2,\ldots \tag{5.30}$$

We notice that, for any $k=2,3,\ldots$, the block matrix $\mathcal{C}_k$ defined by (5.9) consists of $2^{k-1}-1$ blocks. If we denote its j-th block by $\mathcal{C}_{k,j}$ $(j=1,2,\ldots,2^{k-1}-1)$, then

$$\mathcal{C}_k = \big(\mathcal{C}_{k,1},\ \mathcal{C}_{k,2},\ \ldots,\ \mathcal{C}_{k,2^{k-1}-1}\big).$$

For the convenience of subsequent analysis, we denote

$$\eta^{(k-1)}_{-0}(\cdot)\circ\mathcal{C}_k^\top = \sum_{j=1}^{2^{k-1}-1} \eta^{(k-1)}_j(\cdot)\mathcal{C}_{k,j}^\top, \qquad k=2,3,\ldots \tag{5.31}$$

The above defined operation "$\circ$" can be regarded as a certain combination of the classical matrix product with a scalar and the classical inner product of vectors.

With the above preparations, we give the following result:

Lemma 5.5. *Let $\Phi^{(k)}(\cdot)$ $(k=1,2,\ldots)$ be defined by (5.25) and (5.26). Then, we have*

$$\begin{cases} \mathbb{E}[\Phi^{(1)}(\cdot)] = I, \\ \Phi^{(1)}(\cdot) - \mathbb{E}[\Phi^{(1)}(\cdot)] = O \end{cases} \tag{5.32}$$

and

$$\begin{cases} \mathbb{E}[\Phi^{(k)}(\cdot)] = \mathbb{E}[\Phi^{(k-1)}(\cdot)] + \eta^{(k-1)}_0(\cdot)\Big(\widehat{\mathbf{A}}^{k-1}\Big)^\top, \\ \Phi^{(k)}(\cdot) - \mathbb{E}[\Phi^{(k)}(\cdot)] = \Phi^{(k-1)}(\cdot) - \mathbb{E}[\Phi^{(k-1)}(\cdot)] + \eta^{(k-1)}_{-0}(\cdot)\circ\mathcal{C}_k^\top, \end{cases} \qquad k=2,3,\ldots, \tag{5.33}$$

where the involved notations are given by (5.9), (5.29), (5.30), and (5.31).

Proof. (5.32) is obvious. To prove (5.33), we will once again employ the induction method.
Setp 1: Prove the first equation in (5.33).

When $k=2$, by the definition of $\mathbb{E}[\Phi^{(2)}(\cdot)]$, for any $t\in[0,T]$, we calculate

$$\mathbb{E}[\Phi^{(2)}(t)] = I - \int_0^t \widehat{\mathbf{A}}^\top \,\mathrm{d}s = \mathbb{E}[\Phi^{(1)}(t)] + \Big(\int_0^t (-1)\,\mathrm{d}s\Big)\widehat{\mathbf{A}}^\top = \mathbb{E}[\Phi^{(1)}(t)] + \eta_0^{(1)}(t)\widehat{\mathbf{A}}^\top,$$

i.e., the conclusion holds for $k=2$.

Let $k=3,4,\dots$ be given. We assume that the first equation in (5.33) holds for $k-1$, *i.e.*,

$$\mathbb{E}[\Phi^{(k-1)}(\cdot)] = \mathbb{E}[\Phi^{(k-2)}(\cdot)] + \eta_0^{(k-2)}(\cdot)\Big(\widehat{\mathbf{A}}^{k-2}\Big)^\top.$$

Based on this, we are going to prove the same conclusion also holds for k. Indeed, by the recursive definitions of $\mathbb{E}[\Phi^{(k)}(\cdot)]$ and $\mathbb{E}[\Phi^{(k-1)}]$ and the above equation, we calculate

$$\begin{aligned}
\mathbb{E}[\Phi^{(k)}(t)] &= I - \int_0^t \widehat{\mathbf{A}}^\top \mathbb{E}[\Phi^{(k-1)}(s)]\,\mathrm{d}s\\
&= I - \int_0^t \widehat{\mathbf{A}}^\top \Big\{\mathbb{E}[\Phi^{(k-2)}(s)] + \eta_0^{(k-2)}(s)\Big(\widehat{\mathbf{A}}^{k-2}\Big)^\top\Big\}\,\mathrm{d}s\\
&= \Big(I - \int_0^t \widehat{\mathbf{A}}^\top \mathbb{E}[\Phi^{(k-2)}(s)]\,\mathrm{d}s\Big) + \Big(\int_0^t \Big(-\eta_0^{(k-2)}(s)\Big)\,\mathrm{d}s\Big)\Big(\widehat{\mathbf{A}}^{k-1}\Big)^\top\\
&= \mathbb{E}[\Phi^{(k-1)}(t)] + \eta_0^{(k-1)}(t)\Big(\widehat{\mathbf{A}}^{k-1}\Big)^\top,
\end{aligned}$$

which is just the first equation in (5.33). The proof of Step 1 is finished.
Setp 2: Prove the second equation in (5.33).

Let $k=2$. By the definition of $\Phi^{(2)}(\cdot)-\mathbb{E}[\Phi^{(2)}(\cdot)]$, we calculate

$$\begin{aligned}
\Phi^{(2)}(t) - \mathbb{E}[\Phi^{(2)}(t)] &= -\int_0^t \widehat{\mathbf{A}}_1^\top \,\mathrm{d}W(s) = \Big(\Phi^{(1)}(t) - \mathbb{E}[\Phi^{(1)}(t)]\Big) + \Big(\int_0^t (-1)\,\mathrm{d}W(s)\Big)\widehat{\mathbf{A}}_1^\top\\
&= \Big(\Phi^{(1)}(t) - \mathbb{E}[\Phi^{(1)}(t)]\Big) + \eta_1^{(1)}(t)\widehat{\mathbf{A}}_1^\top = \Big(\Phi^{(1)}(t) - \mathbb{E}[\Phi^{(1)}(t)]\Big) + \eta_{-0}^{(1)}(t)\circ\mathcal{C}_2^\top.
\end{aligned}$$

We have verified that the second equation in (5.33) holds for $k=2$. Since Step 2 is a bit complicated and the case of $k=2$ is too simple to reveal the essence, we continue to calculate the case of $k=3$ for a better comprehension.

Let $k=3$. From the recursive definitions of $\Phi^{(3)}(\cdot)-\mathbb{E}[\Phi^{(3)}(\cdot)]$ and $\Phi^{(2)}(\cdot)-\mathbb{E}[\Phi^{(2)}(\cdot)]$, the first equation in (5.33), and the above equation,

$$
\begin{aligned}
&\Phi^{(3)}(t)-\mathbb{E}[\Phi^{(3)}(t)]\\
=&\ -\int_0^t \widehat{\mathbf{A}}_1^\top \mathbb{E}[\Phi^{(2)}]\,\mathrm{d}W-\int_0^t \mathbf{A}^\top\big(\Phi^{(2)}-\mathbb{E}[\Phi^{(2)}]\big)\,\mathrm{d}s-\int_0^t \mathbf{A}_1^\top\big(\Phi^{(2)}-\mathbb{E}[\Phi^{(2)}]\big)\,\mathrm{d}W\\
=&\ -\int_0^t \widehat{\mathbf{A}}_1^\top\Big\{\mathbb{E}[\Phi^{(1)}]+\eta_0^{(1)}\widehat{\mathbf{A}}^\top\Big\}\,\mathrm{d}W-\int_0^t \mathbf{A}^\top\Big\{\big(\Phi^{(1)}-\mathbb{E}[\Phi^{(1)}]\big)+\eta_1^{(1)}\mathcal{C}_2^\top\Big\}\,\mathrm{d}s\\
&\ -\int_0^t \mathbf{A}_1^\top\Big\{\big(\Phi^{(1)}-\mathbb{E}[\Phi^{(1)}]\big)+\eta_1^{(1)}\mathcal{C}_2^\top\Big\}\,\mathrm{d}W\\
=&\ \Big(\Phi^{(2)}(t)-\mathbb{E}[\Phi^{(2)}(t)]\Big)+\Big[\int_0^t\Big(-\eta_0^{(1)}\Big)\,\mathrm{d}W\Big]\widehat{\mathbf{A}}_1^\top\widehat{\mathbf{A}}^\top\\
&\ +\Big[\int_0^t\Big(-\eta_1^{(1)}\Big)\,\mathrm{d}s\Big]\circ\big(\mathbf{A}^\top\mathcal{C}_2^\top\big)+\Big[\int_0^t\Big(-\eta_1^{(1)}\Big)\,\mathrm{d}W\Big]\circ\big(\mathbf{A}_1^\top\mathcal{C}_2^\top\big).
\end{aligned}
$$

By the definitions of $\eta_j^{(k)}$ and $\eta_{-0}^{(k)}$ and the notation "$\circ$", we have

$$
\begin{aligned}
&\Big(\Phi^{(3)}(t)-\mathbb{E}[\Phi^{(3)}(t)]\Big)-\Big(\Phi^{(2)}(t)-\mathbb{E}[\Phi^{(2)}(t)]\Big)\\
&=\eta_1^{(2)}(t)\big(\widehat{\mathbf{A}}\widehat{\mathbf{A}}_1\big)^\top+\eta_2^{(2)}(t)\big(\mathcal{C}_2\mathbf{A}\big)^\top+\eta_3^{(2)}(t)\big(\mathcal{C}_2\mathbf{A}_1\big)^\top\\
&=\eta_{-0}^{(2)}(t)\circ\Big(\widehat{\mathbf{A}}\widehat{\mathbf{A}}_1,\ \mathcal{C}_2\mathbf{A},\ \mathcal{C}_2\mathbf{A}_1\Big)^\top\\
&=\eta_{-0}^{(2)}(t)\circ\Big(\widehat{\mathbf{A}}\widehat{\mathbf{A}}_1,\ \mathcal{C}_2\otimes\big(\mathbf{A},\ \mathbf{A}_1\big)\Big)^\top.
\end{aligned}
$$

We notice that, in the above derivation, we have used $(\mathcal{C}_2\otimes\mathbf{A},\ \mathcal{C}_2\otimes\mathbf{A}_1)=(\mathcal{C}_2\mathbf{A},\ \mathcal{C}_2\mathbf{A}_1)=\mathcal{C}_2\otimes\big(\mathbf{A},\ \mathbf{A}_1\big)$. This equation holds true since $\mathcal{C}_2=\widehat{\mathbf{A}}_1$ which consists of only one block. In the analysis for $k\geq 4$, the related $\mathcal{C}_k$ consists of many blocks. Then, for the corresponding equation holds true, we must exchange some columns of matrices. This will be explained in detail later. Now let us return to the case of $k=3$ and finish the verification of this case. In fact, (5.9) in Lemma 5.3 implies

$$
\Big(\Phi^{(3)}(t)-\mathbb{E}[\Phi^{(3)}(t)]\Big)-\Big(\Phi^{(2)}(t)-\mathbb{E}[\Phi^{(2)}(t)]\Big)=\eta_{-0}^{(2)}(t)\circ\mathcal{C}_3^\top.
$$

Let $k=4,5,\dots$ be given. We assume that the second equation in (5.33) holds for $k-1$, *i.e.*,

$$
\Phi^{(k-1)}(\cdot)-\mathbb{E}[\Phi^{(k-1)}(\cdot)]=\Phi^{(k-2)}(\cdot)-\mathbb{E}[\Phi^{(k-2)}(\cdot)]+\eta_{-0}^{(k-2)}(\cdot)\circ\mathcal{C}_{k-1}^\top. \tag{5.34}
$$

Similar to the case of $k=3$, by the recursive definitions of $\Phi^{(k)}(\cdot)-\mathbb{E}[\Phi^{(k)}(\cdot)]$ and $\Phi^{(k-1)}(\cdot)-\mathbb{E}[\Phi^{(k-1)}(\cdot)]$, the first equation in (5.33), and the above equation, we have

$$
\begin{aligned}
&\Phi^{(k)}(t)-\mathbb{E}[\Phi^{(k)}(t)]\\
=&\ \Big(\Phi^{(k-1)}(t)-\mathbb{E}[\Phi^{(k-1)}(t)]\Big)+\Big[\int_0^t\Big(-\eta_0^{(k-2)}\Big)\,\mathrm{d}W\Big]\Big(\widehat{\mathbf{A}}_1^\top\Big(\widehat{\mathbf{A}}^{k-2}\Big)^\top\Big)\\
&\ +\Big[\int_0^t\Big(-\eta_{-0}^{(k-2)}\Big)\,\mathrm{d}s\Big]\circ\big(\mathbf{A}^\top\otimes\mathcal{C}_{k-1}^\top\big)+\Big[\int_0^t\Big(-\eta_{-0}^{(k-2)}\Big)\,\mathrm{d}W\Big]\circ\big(\mathbf{A}_1^\top\otimes\mathcal{C}_{k-1}^\top\big).
\end{aligned}
$$

From the definitions of (5.29) and (5.30),

$$\begin{aligned}&\int_0^t \Big(-\eta_0^{(k-2)}\Big)\,\mathrm{d}W=\eta_1^{(k-1)}(t),\\&\int_0^t \Big(-\eta_{-0}^{(k-2)}\Big)\,\mathrm{d}s=\bigg(\int_0^t\Big(-\eta_1^{(k-2)}\Big)\,\mathrm{d}s,\ \int_0^t\Big(-\eta_2^{(k-2)}\Big)\,\mathrm{d}s,\ \ldots,\ \int_0^t\Big(-\eta_{2^{k-2}-1}^{(k-2)}\Big)\,\mathrm{d}s\bigg)\\&\qquad=\Big(\eta_2^{(k-1)}(t),\ \eta_4^{(k-1)}(t),\ \ldots,\ \eta_{2^{k-1}-2}^{(k-1)}(t)\Big),\end{aligned}$$

and

$$\begin{aligned}&\int_0^t \Big(-\eta_{-0}^{(k-2)}\Big)\,\mathrm{d}W=\bigg(\int_0^t\Big(-\eta_1^{(k-2)}\Big)\,\mathrm{d}W,\ \int_0^t\Big(-\eta_2^{(k-2)}\Big)\,\mathrm{d}W,\ \ldots,\ \int_0^t\Big(-\eta_{2^{k-2}-1}^{(k-2)}\Big)\,\mathrm{d}W\bigg)\\&\qquad=\Big(\eta_3^{(k-1)}(t),\ \eta_5^{(k-1)}(t),\ \ldots,\ \eta_{2^{k-1}-1}^{(k-1)}(t)\Big).\end{aligned}$$

Therefore,

$$\begin{aligned}&\Big(\Phi^{(k)}(t)-\mathbb{E}[\Phi^{(k)}(t)]\Big)-\Big(\Phi^{(k-1)}(t)-\mathbb{E}[\Phi^{(k-1)}(t)]\Big)\\&=\eta_1^{(k-1)}(t)\big(\widehat{\mathbf{A}}^{k-2}\widehat{\mathbf{A}}_1\big)^\top+\Big(\eta_2^{(k-1)},\ \eta_4^{(k-1)},\ \ldots,\ \eta_{2^{k-1}-2}^{(k-1)}\Big)(t)\circ\big(\mathcal{C}_{k-1}\otimes\mathbf{A}\big)^\top\\&\quad+\Big(\eta_3^{(k-1)},\ \eta_5^{(k-1)},\ \ldots,\ \eta_{2^{k-1}-1}^{(k-1)}\Big)(t)\circ\big(\mathcal{C}_{k-1}\otimes\mathbf{A}_1\big)^\top\\&=\Big(\eta_1^{(k-1)},\ \eta_2^{(k-1)},\ \eta_4^{(k-1)},\ \ldots,\ \eta_{2^{k-1}-2}^{(k-1)},\ \eta_3^{(k-1)},\ \eta_5^{(k-1)},\ \ldots,\ \eta_{2^{k-1}-1}^{(k-1)}\Big)(t)\\&\quad\circ\Big(\widehat{\mathbf{A}}^{k-2}\widehat{\mathbf{A}}_1,\ \mathcal{C}_{k-1}\otimes\mathbf{A},\ \mathcal{C}_{k-1}\otimes\mathbf{A}_1\Big)^\top.\end{aligned}$$

Rearranging the subscripts of $\eta^{(k-1)}(\cdot)$ in the natural order (*i.e.*, from $j=1$ to $j=2^{k-1}-1$) leads to

$$\begin{aligned}&\Big(\Phi^{(k)}(t)-\mathbb{E}[\Phi^{(k)}(t)]\Big)-\Big(\Phi^{(k-1)}(t)-\mathbb{E}[\Phi^{(k-1)}(t)]\Big)\\&=\Big(\eta_1^{(k-1)},\ \eta_2^{(k-1)},\ \eta_3^{(k-1)},\ \ldots,\ \eta_{2^{k-1}-2}^{(k-1)},\ \eta_{2^{k-1}-1}^{(k-1)}\Big)(t)\circ\Big(\widehat{\mathbf{A}}^{k-2}\widehat{\mathbf{A}}_1,\ \mathcal{C}_{k-1}\otimes\big(\mathbf{A},\ \mathbf{A}_1\big)\Big)^\top.\end{aligned}$$

Thanks to (5.9) in Lemma 5.3 once again, we can obtain the desired result and complete the proof. □

Corollary 5.6. *Let $\beta\in\mathbb{R}^n$ be given and $\beta\perp(\mathcal{A}_1,\ \mathcal{A}_2,\ \mathcal{A}_3,\ \ldots)$. Let $\Phi^{(k)}(\cdot)$ $(k=1,2,\ldots)$ be defined by* (5.25) *and* (5.26)*. Then,* (5.28) *holds true.*

Proof. The induction method will be used once again. First of all, we directly verify

$$\widehat{\mathbf{B}}^\top\mathbb{E}[\Phi^{(1)}(\cdot)]\beta=\widehat{\mathbf{B}}^\top I\beta=\mathcal{A}_1^\top\beta=0$$

and

$$\mathbf{B}^\top\Big(\Phi^{(1)}(\cdot)-\mathbb{E}[\Phi^{(1)}(\cdot)]\Big)\beta=\mathbf{B}^\top O\beta=0,$$

i.e., (5.28) holds for $k=1$.

Let $k \geq 2$. We assume that (5.28) holds for $k-1$, and try to prove that (5.28) also holds for k. In fact, with the help of Lemma 5.5, Equation (5.10), and $\beta \perp (\mathcal{A}_1,\ \mathcal{A}_2,\ \mathcal{A}_3,\ \dots)$, we derive

$$\widehat{\mathbf{B}}^\top \mathbb{E}[\Phi^{(k)}(\cdot)]\beta = \widehat{\mathbf{B}}^\top \Big\{\mathbb{E}[\Phi^{(k-1)}(\cdot)] + \eta_0^{(k-1)}(\cdot)\big(\widehat{\mathbf{A}}^{k-1}\big)^\top\Big\}\beta = \eta_0^{(k-1)}(\cdot)\Big(\widehat{\mathbf{A}}^{k-1}\widehat{\mathbf{B}}\Big)^\top \beta = 0,$$

and

$$\begin{aligned}\mathbf{B}^\top\Big(\Phi^{(k)}(\cdot) - \mathbb{E}[\Phi^{(k)}(\cdot)]\Big)\beta &= \mathbf{B}^\top\Big\{\Phi^{(k-1)}(\cdot) - \mathbb{E}[\Phi^{(k-1)}(\cdot)] + \eta_{-0}^{(k-1)}(\cdot)\circ\mathcal{C}_k^\top\Big\}\beta\\ &= \Big[\eta_{-0}^{(k-1)}(\cdot)\circ\big(\mathbf{B}^\top\otimes\mathcal{C}_k^\top\big)\Big]\beta = \eta_{-0}^{(k-1)}(\cdot)\circ\Big[\big(\mathcal{C}_k\otimes\mathbf{B}\big)^\top\otimes\beta\Big] = 0.\end{aligned}$$

The proof is finished. □

Theorem 5.4 provides a rank condition to characterize the exact controllability of (5.1). Now, we give a pair of corollaries of it.

Corollary 5.7. *(i). Let $\mathbf{B} = O$. The system (5.1) is exact controllable if and only if*

$$\operatorname{Rank}\Big(\widehat{\mathbf{B}},\ \widehat{\mathbf{A}}\widehat{\mathbf{B}},\ \widehat{\mathbf{A}}^2\widehat{\mathbf{B}},\ \dots,\ \widehat{\mathbf{A}}^{n-1}\widehat{\mathbf{B}}\Big) = n. \tag{5.35}$$

We notice that, in this case, there is actually no random items involved. The condition (5.35) is exactly the classical Kalman rank condition for ODE system.

(ii). Let $\bar{\mathbf{A}} = \bar{\mathbf{A}}_1 = O$ and $\bar{\mathbf{B}} = O$. The system (5.1) is exact controllable if and only if

$$\operatorname{Rank}\Big(\mathbf{B},\ (\ \mathbf{A},\ \ \mathbf{A}_1\ )\otimes\mathbf{B},\ (\ \mathbf{A},\ \ \mathbf{A}_1\ )^{\otimes 2}\otimes\mathbf{B},\ \dots,\ (\ \mathbf{A},\ \ \mathbf{A}_1\ )^{\otimes k}\otimes\mathbf{B},\ \dots\Big) = n. \tag{5.36}$$

We notice that, in this case, there is no mean-field items involved. The condition (5.36) is exactly the result of Theorem 3.2 in Peng [17].

Proof. (i). Let $\mathbf{B} = O$. Due to (5.4) and (5.10), the Kalman rank condition (5.13) in this case is rewritten as

$$\operatorname{Rank}\big(\widehat{\mathbf{B}},\ \widehat{\mathbf{A}}\widehat{\mathbf{B}},\ \widehat{\mathbf{A}}^2\widehat{\mathbf{B}},\ \dots,\ \widehat{\mathbf{A}}^{n-1}\widehat{\mathbf{B}},\ \cdots\big) = n.$$

By virtue of the classical Hamilton-Cayley Theorem, we know that

$$\operatorname{Rank}\big(\widehat{\mathbf{B}},\ \widehat{\mathbf{A}}\widehat{\mathbf{B}},\ \widehat{\mathbf{A}}^2\widehat{\mathbf{B}},\ \dots,\ \widehat{\mathbf{A}}^{n-1}\widehat{\mathbf{B}},\ \cdots\big) = \operatorname{Rank}\big(\widehat{\mathbf{B}},\ \widehat{\mathbf{A}}\widehat{\mathbf{B}},\ \widehat{\mathbf{A}}^2\widehat{\mathbf{B}},\ \dots,\ \widehat{\mathbf{A}}^{n-1}\widehat{\mathbf{B}}\big).$$

This proves (5.35).

(ii). Let $\bar{\mathbf{A}} = \bar{\mathbf{A}}_1 = O$ and $\bar{\mathbf{B}} = O$. By employing the induction method, we can easily derive (5.36) from (5.4) and (5.7). □

We notice that, the matrix $(\mathcal{A}_1,\ \mathcal{A}_2,\ \mathcal{A}_3,\ \cdots)$ appearing in Theorem 5.4 consists of infinitely many columns, which brings some difficulty of verification for actual use. Recently, for the stochastic system without mean-field items, Lü and Zhang [15] simplified Peng's rank condition (see (5.36)) from infinite many columns to finite many columns. Now let us recall this result as follows:

Lemma 5.8 (Lü and Zhang [15], Thm. 6.10)**.**

$$\begin{aligned}&\operatorname{Span}\Big(\mathbf{B},\ \ (\mathbf{A},\ \ \mathbf{A}_1)\otimes\mathbf{B},\ \ (\mathbf{A},\ \ \mathbf{A}_1)^{\otimes 2}\otimes\mathbf{B},\ \ \dots,\ \ (\mathbf{A},\ \ \mathbf{A}_1)^{\otimes k}\otimes\mathbf{B},\ \ \cdots\Big)\\ =\ &\operatorname{Span}\Big(\mathbf{B},\ \ (\mathbf{A},\ \ \mathbf{A}_1)\otimes\mathbf{B},\ \ (\mathbf{A},\ \ \mathbf{A}_1)^{\otimes 2}\otimes\mathbf{B},\ \ \dots,\ \ (\mathbf{A},\ \ \mathbf{A}_1)^{\otimes(n-1)}\otimes\mathbf{B}\Big).\end{aligned} \tag{5.37}$$

With the help of the above Lemma 5.8, we can also simplify the rank condition in Theorem 5.4 to obtain a mean-field version of the Kalman rank condition to characterize the exact controllability of (5.1).

Based on the matrix $\mathcal{A}_k$ ($k = 1, 2, \ldots$) (see (5.4) and (5.5)), let us define

$$\mathcal{D}_k := \begin{cases} \mathcal{A}_k, \qquad k = 1, 2, \ldots, n+1, \\ \Big(\widehat{\mathbf{A}}^{k-1}\widehat{\mathbf{B}}, \quad \widehat{\mathbf{A}}^{k-2}\widehat{\mathbf{A}}_1\mathbf{B}, \quad \widehat{\mathbf{A}}^{k-3}\widehat{\mathbf{A}}_1 \otimes (\mathbf{A}, \quad \mathbf{A}_1) \otimes \mathbf{B}, \\ \qquad \ldots, \quad \widehat{\mathbf{A}}^{k-(n+1)}\widehat{\mathbf{A}}_1 \otimes (\mathbf{A}, \quad \mathbf{A}_1)^{\otimes(n-1)} \otimes \mathbf{B}\Big), \qquad k = n+2, n+3, \ldots \end{cases} \tag{5.38}$$

Then we have

Theorem 5.9. *Let J defined by (3.19) be the controllable subspace of the system (5.1). Let $\mathcal{D}_k$ ($k = 1, 2, \ldots$) be defined by (5.38). Then*

$$J = \text{Span}\,\big(\mathcal{D}_1, \ \mathcal{D}_2, \ \mathcal{D}_3, \ \ldots, \ \mathcal{D}_{2n}\big). \tag{5.39}$$

Consequently, the system (5.1) is exactly controllable if and only if

$$\text{Rank}\,\big(\mathcal{D}_1, \ \mathcal{D}_2, \ \mathcal{D}_3, \ \ldots, \ \mathcal{D}_{2n}\big) = n. \tag{5.40}$$

Proof. Due to Theorem 5.4, it is clear that the remaining this is to prove

$$\text{Span}\,\big(\mathcal{A}_1, \quad \mathcal{A}_2, \quad \mathcal{A}_3, \quad \cdots\big) = \text{Span}\,\big(\mathcal{D}_1, \quad \mathcal{D}_2, \quad \mathcal{D}_3, \quad \ldots, \quad \mathcal{D}_{2n}\big). \tag{5.41}$$

Step 1: Prove that Span $\big(\mathcal{A}_1, \quad \mathcal{A}_2, \quad \cdots\big) =$ Span $\big(\mathcal{D}_1, \quad \mathcal{D}_2, \quad \cdots\big)$.

Obviously, by the definitions, we only need to prove Span $\big(\mathcal{A}_1, \quad \mathcal{A}_2, \quad \cdots\big) \subset$ Span $\big(\mathcal{D}_1, \quad \mathcal{D}_2, \quad \cdots\big)$ when $k \geq n+2$. In fact, we shall prove that, for any $k \geq n+2$,

$$\text{Span}\,\big(\mathcal{A}_1, \quad \mathcal{A}_2, \quad \ldots, \quad \mathcal{A}_k\big) \subset \text{Span}\,\big(\mathcal{D}_1, \quad \mathcal{D}_2, \quad \ldots, \quad \mathcal{D}_k\big). \tag{5.42}$$

Firstly, we consider the case: $k = n+2$. It is clear that

$$\begin{aligned} &\big(\mathcal{A}_1, \quad \mathcal{A}_2, \quad \ldots, \quad \mathcal{A}_{n+1}, \quad \mathcal{A}_{n+2}\big) \\ =& \Big(\mathcal{A}_1, \quad \mathcal{A}_2, \quad \ldots, \quad \mathcal{A}_{n+1}, \quad \mathcal{D}_{n+2}, \quad \widehat{\mathbf{A}}_1 \otimes (\mathbf{A}, \quad \mathbf{A}_1)^{\otimes n} \otimes \mathbf{B}\Big). \end{aligned}$$

Since Lemma 5.8 implies that

$$\begin{aligned} &\text{Span}\,\big(\mathbf{A}, \quad \mathbf{A}_1\big)^{\otimes n} \otimes \mathbf{B} \\ \subset\, &\text{Span}\,\Big(\mathbf{B}, \quad (\mathbf{A}, \quad \mathbf{A}_1) \otimes \mathbf{B}, \quad (\mathbf{A}, \quad \mathbf{A}_1)^{\otimes 2} \otimes \mathbf{B}, \quad \ldots, \quad (\mathbf{A}, \quad \mathbf{A}_1)^{\otimes(n-1)} \otimes \mathbf{B}\Big), \end{aligned}$$

then

$$\begin{aligned} &\text{Span}\,\widehat{\mathbf{A}}_1 \otimes (\mathbf{A}, \quad \mathbf{A}_1)^{\otimes n} \otimes \mathbf{B} \\ \subset\, &\text{Span}\,\Big(\widehat{\mathbf{A}}_1\mathbf{B}, \quad \widehat{\mathbf{A}}_1 \otimes (\mathbf{A}, \quad \mathbf{A}_1) \otimes \mathbf{B}, \quad \ldots, \quad \widehat{\mathbf{A}}_1 \otimes (\mathbf{A}, \quad \mathbf{A}_1)^{\otimes(n-1)} \otimes \mathbf{B}\Big) \\ \subset\, &\text{Span}\,\big(\mathcal{A}_2, \quad \mathcal{A}_3, \quad \ldots, \quad \mathcal{A}_{n+1}\big). \end{aligned}$$

Therefore, (5.42) is proved when $k = n+2$.

With the help of the induction argument, we can prove in the same way that (5.42) holds for all $k \geq n+2$.

Step 2: Prove that Span $\big(\boldsymbol{\mathcal{D}_1},\ \ \boldsymbol{\mathcal{D}_2},\ \ \cdots\big)$ = Span $\big(\boldsymbol{\mathcal{D}_1},\ \ \boldsymbol{\mathcal{D}_2},\ \ \ldots,\ \ \boldsymbol{\mathcal{D}_{2n}}\big)$.

Firstly, for the first parts of all $\mathcal{D}_k$ $(k = 1, 2, 3, \ldots)$:

$$\widehat{\mathbf{B}}, \quad \widehat{\mathbf{A}}\widehat{\mathbf{B}}, \quad \widehat{\mathbf{A}}^2\widehat{\mathbf{B}}, \quad \ldots, \quad \widehat{\mathbf{A}}^{k-1}\widehat{\mathbf{B}}, \quad \ldots,$$

The classical Hamilton-Cayley Theorem leads to

$$\begin{aligned} &\text{Span}\ \Big(\widehat{\mathbf{B}}, \quad \widehat{\mathbf{A}}\widehat{\mathbf{B}}, \quad \widehat{\mathbf{A}}^2\widehat{\mathbf{B}}, \quad \ldots, \quad \widehat{\mathbf{A}}^{k-1}\widehat{\mathbf{B}}, \quad \cdots\Big) \\ =\ &\text{Span}\ \Big(\widehat{\mathbf{B}}, \quad \widehat{\mathbf{A}}\widehat{\mathbf{B}}, \quad \widehat{\mathbf{A}}^2\widehat{\mathbf{B}}, \quad \ldots, \quad \widehat{\mathbf{A}}^{n-1}\widehat{\mathbf{B}}\Big) \\ \subset\ &\text{Span}\ \big(\mathcal{D}_1, \quad \mathcal{D}_2, \quad \ldots, \quad \mathcal{D}_n\big). \end{aligned} \tag{5.43}$$

The same argument can be applied to the j-th parts of all $\mathcal{D}_k$ $(k = 1, 2, \ldots)$ to yield

$$\begin{aligned} &\text{Span}\ \Big(\widehat{\mathbf{A}}_1 \otimes \big(\mathbf{A}, \quad \mathbf{A}_1\big)^{\otimes(j-2)} \otimes \mathbf{B}, \quad \widehat{\mathbf{A}}\widehat{\mathbf{A}}_1 \otimes \big(\mathbf{A}, \quad \mathbf{A}_1\big)^{\otimes(j-2)} \otimes \mathbf{B}, \\ &\qquad \ldots, \quad \widehat{\mathbf{A}}^{k-j}\widehat{\mathbf{A}}_1 \otimes \big(\mathbf{A}, \quad \mathbf{A}_1\big)^{\otimes(j-2)} \otimes \mathbf{B}, \quad \cdots\Big) \\ \subset\ &\text{Span}\ \big(\mathcal{D}_j, \quad \mathcal{D}_{j+1}, \quad \ldots, \quad \mathcal{D}_{j+n-1}\big), \qquad j = 2, 3, \ldots, n+1, \end{aligned} \tag{5.44}$$

where we have used the notation $\big(\mathbf{A}, \quad \mathbf{A}_1\big)^{\otimes 0} := I$ for simplicity.

From (5.43) and (5.44), we conclude that

$$\text{Span}\ \big(\mathcal{D}_1, \quad \mathcal{D}_2, \quad \cdots\big) = \text{Span}\ \big(\mathcal{D}_1, \quad \mathcal{D}_2, \quad \ldots, \quad \mathcal{D}_{2n}\big). \tag{5.45}$$

We finish the whole proof. □

At the end of this paper, we would like to give an example to illustrate the use of Theorem 5.9.

Example 5.10. Let $n = 3$ and $m = 1$. Let

$$\mathbf{A} = \mathbf{A}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \bar{\mathbf{A}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \bar{\mathbf{A}}_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{B} = -\bar{\mathbf{B}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Then the system (5.1) turns to

$$\begin{cases} \mathrm{d}x_1(t) = \Big(\mathbb{E}[q_1(t)] + v(t) - \mathbb{E}[v(t)]\Big)\,\mathrm{d}t + q_1(t)\,\mathrm{d}W(t), \quad t \in [0, T], \\ \mathrm{d}x_2(t) = \Big(x_1(t) + q_1(t) - \mathbb{E}[q_1(t)] + \mathbb{E}[q_2(t)]\Big)\,\mathrm{d}t + q_2(t)\,\mathrm{d}W(t), \quad t \in [0, T], \\ \mathrm{d}x_3(t) = \Big(\mathbb{E}[x_2(t)] + \mathbb{E}[q_3(t)]\Big)\,\mathrm{d}t + q_3(t)\,\mathrm{d}W(t), \quad t \in [0, T], \\ x_1(T) = x_2(T) = x_3(T) = 0. \end{cases} \tag{5.46}$$

We directly calculate

$$\widehat{\mathbf{A}}_1\mathbf{B} = \mathbf{B} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \widehat{\mathbf{A}}_1\mathbf{A}_1\mathbf{B} = \mathbf{A}_1\mathbf{B} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

and

$$\widehat{\mathbf{A}}\widehat{\mathbf{A}}\widehat{\mathbf{A}}_1\mathbf{B} = \widehat{\mathbf{A}}\widehat{\mathbf{A}}\mathbf{B} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Therefore (see (5.6)),

$$\text{Rank}\,\big(\widehat{\mathbf{A}}_1\mathbf{B},\ \widehat{\mathbf{A}}_1\mathbf{A}_1\mathbf{B},\ \widehat{\mathbf{A}}\widehat{\mathbf{A}}\widehat{\mathbf{A}}_1\mathbf{B}\big) = 3.$$

By Theorem 5.9 (or Thm. 5.4), the system (5.46) is exactly controllable.

In this example, we only have a 1-dimensional control $v(\cdot) - \mathbb{E}[v(\cdot)]$ with zero-mean. (i) The control directly affects the first component $x_1(\cdot)$ of the state; (ii) Moreover, it affects the second component $x_2(\cdot)$ of the state through $(x_1(\cdot), q_1(\cdot) - \mathbb{E}[q_1(\cdot)])$; (iii) Furthermore, it affects the third component $x_3(\cdot)$ of the state through $\mathbb{E}[x_2(\cdot)]$. The overall effect is that the control process $v(\cdot) - \mathbb{E}[v(\cdot)]$ can exactly control all three components of the state.

Appendix A. Another reduction from infinite to finite dimensions in time-invariant case

In this appendix, all the coefficients in (1.1) are restricted to be time-invariant. For clarity, we represent the time-invariant version of (1.1) as follows:

$$\begin{aligned} \mathrm{d}x(t) = &\Big\{ Ax(t) + \bar{A}\mathbb{E}[x(t)] + Bu(t) + \bar{B}\mathbb{E}[u(t)] \Big\}\,\mathrm{d}t \\ &+ \Big\{ Cx(t) + \bar{C}\mathbb{E}[x(t)] + Du(t) + \bar{D}\mathbb{E}[u(t)] \Big\}\,\mathrm{d}W(t), \quad t \in [0, T]. \end{aligned} \tag{A.1}$$

In Section 3, inspired by Wang *et al.* [21], we have reduced the general time-variant system (1.1) to the system (1.3). However, for the time-invariant stochastic system without mean-field terms (*i.e.*, (A.1) without mean-field terms), Peng [17] introduced a different approach to reduce his problem from the infinite to the finite dimensions. Although Peng's approach is not suitable for time-variant stochastic systems (with or without mean-field terms), but it can be extended to deal with the time-invariant mean-field system (A.1). In this appendix, we shall show this different approach.

In fact, the system (A.1) will be reduced to the following one under Assumption (H$'$):

$$\begin{cases} \mathrm{d}x(t) = \Big\{ \mathbf{C}x(t) + \bar{\mathbf{C}}\mathbb{E}[x(t)] + \mathbf{C}_1 z(t) + \bar{\mathbf{C}}_1\mathbb{E}[z(t)] + \mathbf{D}\alpha(t) + \bar{\mathbf{D}}\mathbb{E}[\alpha(t)] \Big\}\,\mathrm{d}t \\ \qquad\qquad + z(t)\,\mathrm{d}W(t), \quad t \in [0, T], \\ x(T) = 0, \end{cases} \tag{A.2}$$

where $\mathbf{C}$, $\bar{\mathbf{C}}$, $\mathbf{C}_1$, $\bar{\mathbf{C}}_1 \in \mathbb{R}^{n\times n}$ and $\mathbf{D}$, $\bar{\mathbf{D}} \in \mathbb{R}^{n\times(m-n)}$ will be defined below (see (A.7), (A.8) and the related (A.3), (A.4)). Similar to (5.1) or its time-variant version (1.3), the system (A.2) is also an MF-BSDE, in which $\alpha(\cdot) \in L^2_{\mathbb{F}}(0, T; \mathbb{R}^{m-n})$ is the control process and $(x(\cdot), z(\cdot)) \equiv (x(\cdot;\alpha(\cdot)), z(\cdot;\alpha(\cdot))) \in L^2_{\mathbb{F}}(\Omega; C(0, T; \mathbb{R}^n)) \times L^2_{\mathbb{F}}(0, T; \mathbb{R}^n)$ is the solution to (A.2). We notice that, compared with the dimension m of the control $v(\cdot)$ in (5.1), the dimension of the control $\alpha(\cdot)$ in (A.2) is $(m - n)$. This is a difference between these two approaches. Besides, this new approach will provide another understanding from a different viewpoint.

Now, we begin our derivation to reduce (A.1) to (A.2) which is inspired by Peng [17] and Goreac [8]. Let Assumption (H′) hold. Then, there exist two non-singular matrices M, $\widehat{M} \in \mathbb{R}^{m\times m}$ such that

$$DM = \Big(I_n,\ O_{n\times(m-n)}\Big) \quad \text{and} \quad \widehat{D}\widehat{M} = \Big(I_n,\ O_{n\times(m-n)}\Big). \tag{A.3}$$

Let a pair of processes $(x(\cdot), u(\cdot)) \in L^2_{\mathbb{F}}(\Omega; C(0,T;\mathbb{R}^n)) \times L^2_{\mathbb{F}}(0,T;\mathbb{R}^m)$ satisfy (A.1). We calculate

$$\begin{aligned} \mathrm{d}x = &\Big\{A\big(x-\mathbb{E}[x]\big) + \widehat{A}\mathbb{E}[x] + BMM^{-1}\big(u-\mathbb{E}[u]\big) + \widehat{B}\widehat{M}\widehat{M}^{-1}\mathbb{E}[u]\Big\}\,\mathrm{d}t \\ &+ \Big\{C\big(x-\mathbb{E}[x]\big) + \widehat{C}\mathbb{E}[x] + DMM^{-1}\big(u-\mathbb{E}[u]\big) + \widehat{D}\widehat{M}\widehat{M}^{-1}\mathbb{E}[u]\Big\}\,\mathrm{d}W. \end{aligned}$$

We denote the following block representations for matrices:

$$\begin{aligned} &BM = \big(\ \{BM\}_1,\ \ \{BM\}_2\ \big), \qquad \widehat{B}\widehat{M} = \Big(\ \{\widehat{B}\widehat{M}\}_1,\ \ \{\widehat{B}\widehat{M}\}_2\ \Big), \\ &M^{-1}\big(u-\mathbb{E}[u]\big) = \begin{pmatrix} \{M^{-1}(u-\mathbb{E}[u])\}_1 \\ \{M^{-1}(u-\mathbb{E}[u])\}_2 \end{pmatrix}, \qquad \widehat{M}^{-1}\mathbb{E}[u] = \begin{pmatrix} \{\widehat{M}^{-1}\mathbb{E}[u]\}_1 \\ \{\widehat{M}^{-1}\mathbb{E}[u]\}_2 \end{pmatrix} \end{aligned} \tag{A.4}$$

with suitable dimensions. With the help of (A.3), we have

$$\begin{aligned} \mathrm{d}x = &\Big\{A\big(x-\mathbb{E}[x]\big) + \widehat{A}\mathbb{E}[x] + \{BM\}_1\{M^{-1}\big(u-\mathbb{E}[u]\big)\}_1 + \{BM\}_2\{M^{-1}\big(u-\mathbb{E}[u]\big)\}_2 \\ &+ \{\widehat{B}\widehat{M}\}_1\{\widehat{M}^{-1}\mathbb{E}[u]\}_1 + \{\widehat{B}\widehat{M}\}_2\{\widehat{M}^{-1}\mathbb{E}[u]\}_2\Big\}\,\mathrm{d}t \\ &+ \Big\{C\big(x-\mathbb{E}[x]\big) + \widehat{C}\mathbb{E}[x] + \{M^{-1}\big(u-\mathbb{E}[u]\big)\}_1 + \{\widehat{M}^{-1}\mathbb{E}[u]\}_1\Big\}\,\mathrm{d}W. \end{aligned}$$

Let us define

$$\begin{cases} z = C\big(x-\mathbb{E}[x]\big) + \widehat{C}\mathbb{E}[x] + \{M^{-1}\big(u-\mathbb{E}[u]\big)\}_1 + \{\widehat{M}^{-1}\mathbb{E}[u]\}_1, \\ \alpha = \{M^{-1}\big(u-\mathbb{E}[u]\big)\}_2 + \{\widehat{M}^{-1}\mathbb{E}[u]\}_2 \end{cases} \tag{A.5}$$

which is equivalent to

$$\begin{cases} \mathbb{E}[z] = \widehat{C}\mathbb{E}[x] + \{\widehat{M}^{-1}\mathbb{E}[u]\}_1, \\ z - \mathbb{E}[z] = C\big(x-\mathbb{E}[x]\big) + \{M^{-1}\big(u-\mathbb{E}[u]\big)\}_1, \\ \mathbb{E}[\alpha] = \{\widehat{M}^{-1}\mathbb{E}[u]\}_2, \\ \alpha - \mathbb{E}[\alpha] = \{M^{-1}\big(u-\mathbb{E}[u]\big)\}_2. \end{cases} \tag{A.6}$$

Clearly, $(z(\cdot), \alpha(\cdot)) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n) \times L^2_{\mathbb{F}}(0,T;\mathbb{R}^{m-n})$. Therefore,

$$\begin{aligned} \mathrm{d}x = &\bigg\{\Big(A - \{BM\}_1 C\Big)\big(x-\mathbb{E}[x]\big) + \Big(\widehat{A} - \{\widehat{B}\widehat{M}\}_1\widehat{C}\Big)\mathbb{E}[x] + \{BM\}_1\big(z-\mathbb{E}[z]\big) \\ &+ \{\widehat{B}\widehat{M}\}_1\mathbb{E}[z] + \{BM\}_2\big(\alpha-\mathbb{E}[\alpha]\big) + \{\widehat{B}\widehat{M}\}_2\mathbb{E}[\alpha]\bigg\}\,\mathrm{d}t + z\,\mathrm{d}W. \end{aligned}$$

Denote

$$\begin{aligned} &\mathbf{C} = A - \{BM\}_1 C, && \mathbf{C}_1 = \{BM\}_1, && \mathbf{D} = \{BM\}_2, \\ &\widehat{\mathbf{C}} = \widehat{A} - \{\widehat{B}\widehat{M}\}_1 \widehat{C}, && \widehat{\mathbf{C}}_1 = \{\widehat{B}\widehat{M}\}_1, && \widehat{\mathbf{D}} = \{\widehat{B}\widehat{M}\}_2, \end{aligned} \tag{A.7}$$

and

$$\begin{aligned} &\bar{\mathbf{C}} = \widehat{\mathbf{C}} - \mathbf{C} = \bar{A} - \Big(\{\widehat{B}\widehat{M}\}_1 \widehat{C} - \{BM\}_1 C\Big), \\ &\bar{\mathbf{C}}_1 = \widehat{\mathbf{C}}_1 - \mathbf{C}_1 = \{\widehat{B}\widehat{M}\}_1 - \{BM\}_1, \\ &\bar{\mathbf{D}} = \widehat{\mathbf{D}} - \mathbf{D} = \{\widehat{B}\widehat{M}\}_2 - \{BM\}_2. \end{aligned} \tag{A.8}$$

With these notations, we have

$$\mathrm{d}x = \Big\{\mathbf{C}\big(x - \mathbb{E}[x]\big) + \widehat{\mathbf{C}}\mathbb{E}[x] + \mathbf{C}_1\big(z - \mathbb{E}[z]\big) + \widehat{\mathbf{C}}_1\mathbb{E}[z] + \mathbf{D}\big(\alpha - \mathbb{E}[\alpha]\big) + \widehat{\mathbf{D}}\mathbb{E}[\alpha]\Big\}\,\mathrm{d}t + z\,\mathrm{d}W.$$

Equivalently,

$$\mathrm{d}x = \Big\{\mathbf{C}x + \bar{\mathbf{C}}\mathbb{E}[x] + \mathbf{C}_1 z + \bar{\mathbf{C}}_1\mathbb{E}[z] + \mathbf{D}\alpha + \bar{\mathbf{D}}\mathbb{E}[\alpha]\Big\}\,\mathrm{d}t + z\,\mathrm{d}W, \quad t \in [0,T]. \tag{A.9}$$

Proposition A.1. *Under Assumption (H′), the exact controllability of* (A.1) *is equivalent to that of* (A.9).

Proof. The necessity is obtained from the above derivation. For the sufficiency, we notice that the transformation (A.5) from $u(\cdot)$ to $(z(\cdot), \alpha(\cdot))$ is invertible. In fact,

$$u = \widehat{M}\begin{pmatrix} \mathbb{E}[z] - \widehat{C}\mathbb{E}[x] \\ \mathbb{E}[\alpha] \end{pmatrix} + M\begin{pmatrix} \big(z - \mathbb{E}[z]\big) - C\big(x - \mathbb{E}[x]\big) \\ \alpha - \mathbb{E}[\alpha] \end{pmatrix}. \tag{A.10}$$

The remaining calculation is in the opposite direction of the previous derivation, and we would like to omit them. □

Next, the same procedure as Proposition 3.3 leads to

Theorem A.2. *Under Assumption (H′), the exact controllability of* (A.1) *is equivalent to that of* (A.2).

We notice that, if we replace the system (5.1) by (A.2), then all the derivations and results in Section 5 will be still valid.

Acknowledgements. The author thanks Professor Qi Lü, Sichuan University, for telling the author the result of Lemma 5.8 during the revision of this paper. This significantly helped the author improve the quality of this work. The author also acknowledges the two anonymous referees for their constructive and insightful comments which also improved the quality of this work.

References

[1] X. Bi, J. Sun and J. Xiong, Optimal control for controllable stochastic linear systems. *ESAIM: COCV* **26** (2020) 98.

[2] R. Buckdahn, B. Djehiche and J. Li, A general stochastic maximum principle for SDEs of mean-field type. *Appl. Math. Optim.* **64** (2011) 197–216.

[3] R. Buckdahn, B. Djehiche, J. Li and S. Peng, Mean-field backward stochastic differential equations: a limit approach. *Ann. Probab.* **37** (2009) 1524–1565.

[4] R. Buckdahn, J. Li and S. Peng, Mean-field backward stochastic differential equations and related partial differential equations. *Stochastic Process. Appl.* **119** (2009) 3133–3154.

[5] R. Buckdahn, M. Quincampoix and G. Tessitore, A characterization of approximately controllable linear stochastic differential equations. Stochastic partial differential equations and applications–VII. *Lect. Notes Pure Appl. Math.,* 245 Chapman & Hall/CRC, Boca Raton, FL (2006) 53–60.

[6] R. Carmona and F. Delarue, Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics. *Ann. Probab.* **43** (2015) 2647–2700.

[7] D. Goreac, A Kalman-type condition for stochastic approximate controllability. *C. R. Math. Acad. Sci. Paris* **346** (2008) 183–188.

[8] D. Goreac, Controllability properties of linear mean-field stochastic systems. *Stoch. Anal. Appl.* **32** (2014) 280–297.

[9] M. Huang, R.P. Malhamé and P.E. Caines, Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.* **6** (2006) 221–251.

[10] M. Kac, Foundations of kinetic theory. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability,* 19541955, vol. III. University of California Press, Berkeley and Los Angeles (1956) 171–197.

[11] J.M. Lasry and P.L. Lions, Mean field games. *Jpn. J. Math.* **2** (2007) 229–260.

[12] E.B. Lee and L. Markus, Foundations of optimal control theory. John Wiley & Sons, Inc., New York-London-Sydney (1967).

[13] J.L. Lions, Exact controllability, stabilization and perturbations for distributed systems. *SIAM Rev.* **30** (1988) 1–68.

[14] F. Liu and S. Peng, On controllability for stochastic control systems when the coefficient is time-variant. *J. Syst. Sci. Complex* **23** (2010) 270–278.

[15] Q. Lü and X. Zhang, Mathematical Theory for Stochastic Distributed Parameter Control Systems. Preprint.

[16] H.P. McKean, Jr., A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. U.S.A.* **56** (1966) 1907–1911.

[17] S. Peng, Backward stochastic differential equation and exact controllability of stochastic control systems. *Progr. Natur. Sci.* **4** (1994) 274-284.

[18] H. Pham and X. Wei, Bellman equation and viscosity solutions for mean-field stochastic control problem. *ESAIM: COCV* **24** (2018) 437–461.

[19] D.L. Russell, Controllability and stabilizability theory for linear partial differential equations: recent progress and open questions. *SIAM Rev.* **20** (1978) 639–739.

[20] R. Tian, Z. Yu and R. Zhang, A closed-loop saddle point for zero-sum linear-quadratic stochastic differential games with mean-field type. *Systems Control Lett.* **136** (2020) 104624.

[21] Y. Wang, D. Yang, J. Yong and Z. Yu, Exact controllability of linear stochastic differential equations and related problems. *Math. Control Relat. Fields* **7** (2017) 305–345.

[22] Y. Wang and Z. Yu, On the partial controllability of SDEs and the exact controllability of FBSDEs. *ESAIM: COCV* **26** (2020) Paper No. 68, 27 pp.

[23] Q. Wei, J. Yong and Z.Yu, Linear quadratic stochastic optimal control problems with operator coefficients: open-loop solutions. *ESAIM: COCV* **25** (2019) 17.

[24] W. Ye and Z. Yu, Exact controllability of linear mean-field stochastic systems and observability inequality for mean-field backward stochastic differential equations. *Asian J. Control* (2020) 1–12.

[25] J. Yong, Linear-quadratic optimal control problems for mean-field stochastic differential equations. *SIAM J. Control Optim.* **51** (2013) 2809–2838.

[26] R. Zhang and L. Guo, Controllability of stochastic game-based control systems. *SIAM J. Control Optim.* **57** (2019) 3799–3826.

[27] E. Zuazua, Controllability and observability of partial differential equations: some results and open problems. Vol. III of Handbook of differential equations: evolutionary equations. *Handb. Differ. Equ.,* Elsevier/North-Holland, Amsterdam (2007) 527–621.