Linear Forward-Backward Stochastic Differential Equations *

Jiongmin Yong

Department of Mathematics, Fudan University, Shanghai 200433, China

Abstract. The problem of finding adapted solutions to systems of coupled linear forward-backward stochastic differential equations (FBSDEs, for short) is investigated. A necessary condition of solvability leads to a reduction of general linear FBSDEs to a special one. By some ideas from controllability in control theory, using some functional analysis, we obtain a necessary and sufficient condition for the solvability of the linear FBSDEs with the processes Z (serves as a correction, see §1) being absent in the drift. Then a Riccati type equation for matrix-valued (not necessarily square) functions is derived using the idea of the Four-Step-Scheme (introduced in [11] for general FBSDEs). The solvability of such a Riccati type equation is studied which leads to a representation of adapted solutions to linear FBSDEs.

Keywords. Linear forward-backward stochastic differential equations, adapted solution, Riccati type equation.

AMS Mathematics subject classification. 60H10.

§1. Introduction.

Let $(\Omega, P, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{P})$ be a complete probability space on which defined a one dimensional standard Brownian motion $W(t)$, such that $\{\mathcal{F}_t\}_{t\geq 0}$ is the natural filtration generated by $W(t)$, augmented by all the $\mathcal{P}$-null sets in $\mathcal{F}$. In this paper, we consider the following system of coupled linear *forward-backward stochastic differential equations*

* This work is supported in part by the NNSF of China, the Chinese State Education Commission Science Foundation and the Trans-Century Training Programme Foundation for the Talents by the State Education Commission of China. Part of the work has been done while the author was visiting Department of Mathematics and Statistics, University of Minnesota, Duluth. The author would like to thank Professor Z. Liu for his hospitality. Some stimulating discussions with Professor J. Ma of Purdue University deserves a special acknowledgment.

(FBSDEs for short) on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{P})$:

$$
(1.1) \qquad \begin{cases} dX(t) = \big\{AX(t) + BY(t) + CZ(t) + Db(t)\big\}dt \\ \qquad\qquad + \big\{A_1X(t) + B_1Y(t) + C_1Z(t) + D_1\sigma(t)\big\}dW(t), \\ dY(t) = \big\{\widehat{A}X(t) + \widehat{B}Y(t) + \widehat{C}Z(t) + \widehat{D}\widehat{b}(t)\big\}dt \\ \qquad\qquad + \big\{\widehat{A}_1X(t) + \widehat{B}_1Y(t) + \widehat{C}_1Z(t) + \widehat{D}_1\widehat{\sigma}(t)\big\}dW(s), \\ X(0) = x, \qquad Y(T) = GX(T) + Fg. \end{cases}
$$

In the above, A, B, C etc. are (deterministic) matrices of suitable sizes, b, σ, $\widehat{b}$ and $\widehat{\sigma}$ are stochastic processes and g is a random variable. We are looking for $\{\mathcal{F}_t\}$-adapted processes $X(\cdot)$, $Y(\cdot)$ and $Z(\cdot)$, valued in $\mathbb{R}^n$, $\mathbb{R}^m$ and $\mathbb{R}^\ell$, respectively, satisfying the above.

We see that (1.1) is a kind of two-point boundary value problem for a system of linear stochastic differential equations. The key issue is that we want the processes X and Y to be $\{\mathcal{F}_t\}$-adapted. This is by no means obviously possible since $Y(T)$ is given as an $\mathcal{F}_T$-measurable random variable. Thanks to the introduction of the $\{\mathcal{F}_t\}$-adapted process Z, one obtains an extra freedom, which makes it possible to find $\{\mathcal{F}_t\}$-adapted processes (X, Y) satisfying (1.1), under certain mild conditions. We see that Z serves as a *correction*.

If there only is the equation for $Y(\cdot)$ in (1.1) (with $\widehat{A} = \widehat{A}_1 = 0$ and $G = 0$), we have the so-called *backward stochastic differential equation* (BSDE, for short). The study of such an equation can be traced back to Bismut [3] and the general solvability result was obtained by Bensoussan [2] using the Martingale Representation Theorem. Nonlinear BSDEs were studied by Pardoux and Peng [15] using the contraction mapping theorem. For general nonlinear FBSDEs, one can find several works. Let us briefly list them and mention the methods used. See [8] for a survey of BSDEs. Antonelli used the contraction mapping theorem to prove the solvability of FBSDEs in *small* time duration ([1]). See [17] also. In [12], Ma and Yong proved the weak solvability of a class of FBSDEs over *any* finite time duration via the stochastic optimal control theory. Later, Ma, Protter and Yong, inspired by [12], introduced the so-called Four-Step-Scheme ([11]) to obtain the solvability of FBSDEs with deterministic coefficients and with nondegenerate diffusion in the forward equation. See also [7], [5] and [6] for related results. Further development along this direction is still undergoing (see [13], [14]). In [9], Hu and Peng introduced the monotonicity condition, under which the FBSDEs can be solved. See [18] and [4] also. In [20], Yong introduced the method of continuation and the concept of bridge to treat the solvability of FBSDEs in a very general way. Pardoux and Tang studied the solvability of FBSDEs under some structure conditions [16]. All the above-mentioned works gave solvability for different classes of FBSDEs. We point out that the general solvability problem, however, is far away from completely solved.

In [20], among other things, this author studied a special class of linear FBSDEs via which, together with the bridge technique, some new classes of solvable FBSDEs were obtained. Inspired by this, in the present paper, we would like to study the solvability of general linear FBSDEs. Due to the linearity of the equations, it is expected to obtain relatively satisfactory solvability results than the general nonlinear situation. It is our hope that via such a study, one might get some new classes of solvable FBSDEs by combining the bridge technique introduced in [20].

In the first part of this paper, we present some necessary conditions for (1.1) to be solvable. These will lead to some reductions of (1.1) to a (seemingly) special one. Then for the reduced problem, we introduce two methods to study the solvability. Using functional analysis together with some control theoretic idea, among other thing, we obtain a necessary and sufficient conditions for the solvability of linear FBSDEs with the process Z does not appear in the drift. This result extends the relevant one in [20]. Our result reveals a significant difference between the solvability of FBSDEs and two-point boundary value problems for ordinary differential equations from the viewpoint of solvable time durations (see §4 for details). Next, we use the idea of Four-Step-Scheme [11] to derive a Riccati type differential equation for $(m \times n)$-matrix valued functions and a backward SDE associated with the reduced linear FBSDEs. It is shown that the solvability of such a Riccati type equation gives the unique solvability of the linear FBSDEs and moreover, the adapted solution is represented explicitly in terms of the solutions of the Riccati type equation and the corresponding backward SDE. Thus, this method is more constructive. In the case that Z does not appear in the drift, we obtain a necessary and sufficient condition for the Riccati type equation to be solvable and explicitly construct the solution to this equation. Finally, we extend our results to the case with multi-dimensional Brownian motion.

§2. A Necessary Condition for Solvability.

Let us introduce some notations.

For any sub-σ-field $\mathcal{G}$ of $\mathcal{F}$, we denote $L^2_{\mathcal{G}}(\Omega;\mathbb{R}^m)$ to be the set of all $\mathcal{G}$-measurable $\mathbb{R}^m$-valued square-integrable random variables. Let $L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ be the set of all $\{\mathcal{F}_t\}$-progressively measurable processes $X(\cdot)$ valued in $\mathbb{R}^n$ such that

$$\int_0^T E|X(t)|^2 dt < \infty.$$

Also, let $L^2_{\mathcal{F}}(\Omega;C([0,T];\mathbb{R}^n))$ be the set of all $\{\mathcal{F}_t\}$-progressively measurable continuous processes $X(\cdot)$ valued in $\mathbb{R}^n$, such that

$$E \sup_{t\in[0,T]} |X(t)|^2 < \infty.$$

Further, we define

$$\mathcal{M}[0,T] \stackrel{\Delta}{=} L^2_{\mathcal{F}}(\Omega;C([0,T];\mathbb{R}^n)) \times L^2_{\mathcal{F}}(\Omega;C([0,T];\mathbb{R}^m)) \times L^2_{\mathcal{F}}(0,T;\mathbb{R}^\ell). \tag{2.1}$$

The norm of this space is defined by

$$\begin{aligned} \|(X(\cdot),Y(\cdot),Z(\cdot))\| = \Big\{E \sup_{t\in[0,T]} |X(t)|^2 + E \sup_{t\in[0,T]} |Y(t)|^2 + E\int_0^T |Z(t)|^2 dt\Big\}^{1/2}, \\ \forall (X(\cdot),Y(\cdot),Z(\cdot)) \in \mathcal{M}[0,T]. \end{aligned} \tag{2.2}$$

Clearly, $\mathcal{M}[0,T]$ is a Banach space under norm (2.2). Let us introduce the following definition.

Definition 2.1. A triple $(X,Y,Z) \in \mathcal{M}[0,T]$ is called an *adapted solution* of (1.1) if the following holds for all $t \in [0,T]$, almost surely:

$$\left\{\begin{aligned} X(t) &= x + \int_0^t \big\{AX(s) + BY(s) + CZ(s) + Db(s)\big\} ds \\ &\qquad + \int_0^t \big\{A_1X(s) + B_1Y(s) + C_1Z(s) + D_1\sigma(s)\big\} dW(s), \\ Y(t) &= GX(T) + Fg - \int_t^T \big\{\widehat{A}X(s) + \widehat{B}Y(s) + \widehat{C}Z(s) + \widehat{D}\widehat{b}(s)\big\} ds \\ &\qquad - \int_t^T \big\{\widehat{A}_1X(s) + \widehat{B}_1Y(s) + \widehat{C}_1Z(s) + \widehat{D}_1\widehat{\sigma}(s)\big\} dW(s). \end{aligned}\right. \tag{2.3}$$

When (1.1) admits an adapted solution, we say that (1.1) is solvable.

In what follows, we will let

$$
\left\{
\begin{array}{l}
A, A_1 \in \mathbb{R}^{n\times n};\ \ B, B_1 \in \mathbb{R}^{n\times m};\ \ \ C, C_1 \in \mathbb{R}^{n\times \ell};\\
\widehat{A}, \widehat{A}_1, G \in \mathbb{R}^{m\times n};\ \ \ \widehat{B}, \widehat{B}_1 \in \mathbb{R}^{m\times m};\ \ \ \widehat{C}, \widehat{C}_1 \in \mathbb{R}^{m\times \ell};\\
D \in \mathbb{R}^{n\times \bar n};\ \ \ D_1 \in \mathbb{R}^{n\times \bar n_1};\ \ \ \widehat{D} \in \mathbb{R}^{m\times \bar m};\ \ \ \widehat{D}_1 \in \mathbb{R}^{m\times \bar m_1};\ \ \ F \in \mathbb{R}^{m\times k};\\
b \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^{\bar n});\ \ \ \sigma \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^{\bar n_1});\\
\widehat{b} \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^{\bar m});\ \ \ \widehat{\sigma} \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^{\bar m_1});\\
g \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^{k});\ \ \ x \in \mathbb{R}^n.
\end{array}
\right.
\tag{2.4}
$$

Following result gives a necessary condition for (1.1) to be solvable.

Theorem 2.2. *Suppose there exists a* $T>0$*, such that for all* b*,* σ*,* $\widehat{b}$*,* $\widehat{\sigma}$*,* g *and* x *satisfying (2.4), (1.1) admits an adapted solution* $(X,Y,Z)\in\mathcal{M}[0,T]$*. Then,*

$$
\mathcal{R}(\widehat{C}_1 - GC_1) \supseteq \mathcal{R}(F) + \mathcal{R}(\widehat{D}_1) + \mathcal{R}(GD_1),
\tag{2.5}
$$

where $\mathcal{R}(S)$ *is the range of operator* S*. In particular, if*

$$
\mathcal{R}(F) + \mathcal{R}(\widehat{D}_1) + \mathcal{R}(GD_1) = \mathbb{R}^m,
\tag{2.6}
$$

then $\widehat{C}_1 - GC_1 \in \mathbb{R}^{m\times\ell}$ *is onto and thus* $\ell \ge m$*.*

To prove the above result, we need the following lemma, which is interesting by itself.

Lemma 2.3. *Suppose that for any* $\bar\sigma \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^{\bar k})$ *and any* $g\in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^k)$*, there exist* $h\in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ *and* $f\in L^2_{\mathcal{F}}(\Omega;C([0,T];\mathbb{R}^m))$*, such that the following BSDE admits an adapted solution* $(\overline{Y},Z)\in L^2_{\mathcal{F}}(\Omega;C([0,T];\mathbb{R}^m))\times L^2_{\mathcal{F}}(0,T;\mathbb{R}^\ell)$*:*

$$
\left\{
\begin{array}{ll}
d\overline{Y}(t) = h(t)dt + [f(t) + \overline{C}_1 Z(t) + \overline{D}\bar\sigma(t)]dW(t), & t\in[0,T],\\
\overline{Y}(T) = Fg,
\end{array}
\right.
\tag{2.7}
$$

where $\overline{C}_1 \in \mathbb{R}^{m\times\ell}$ *and* $\overline{D}\in\mathbb{R}^{m\times\bar k}$*. Then,*

$$
\mathcal{R}(\overline{C}_1) \supseteq \mathcal{R}(F) + \mathcal{R}(\overline{D}).
\tag{2.8}
$$

Proof. Suppose (2.8) does not hold. Then, we can find an $\eta\in\mathbb{R}^m$, such that

$$
\eta^T\overline{C}_1 = 0, \quad \text{but} \ \ \eta^T F \ne 0, \quad \text{or} \quad \eta^T\overline{D}\ne 0.
\tag{2.9}
$$

Let $\zeta(t) = \eta^T \overline{Y}(t)$. Then, $\zeta(\cdot)$ satisfies

$$
(2.10) \qquad \begin{cases} d\zeta(t) = \bar{h}(t)dt + [\bar{f}(t) + \eta^T \overline{D}\bar{\sigma}(t)]dW(t), \\ \zeta(T) = \eta^T F g, \end{cases}
$$

where $\bar{h}(t) = \eta^T h(t)$, $\bar{f}(t) = \eta^T f(t)$. We claim that for some g and $\bar{\sigma}(\cdot)$, (2.10) does not admit an adapted solution $\zeta(\cdot)$ for any $\bar{h} \in L^2_{\mathcal{F}}(0,T;\mathbb{R})$ and $\bar{f} \in L^2_{\mathcal{F}}(\Omega; C([0,T];\mathbb{R}))$. To show this, we construct a deterministic Lebesgue measurable function β satisfying the following:

$$
(2.11) \qquad \begin{cases} \beta(s) = \pm 1, \qquad \forall s \in [0,T], \\ |\{s \in [T_i, T] \mid \beta(s) = 1\}| = |\{s \in [T_i, T] \mid \beta(s) = -1\}| = \dfrac{T - T_i}{2}, \qquad i \geq 1, \end{cases}
$$

for a sequence $T_i \uparrow T$, where $|\{\cdots\}|$ stands for the Lebesgue measure of $\{\cdots\}$. Such a function exists by some elementary construction. Now, we separate two cases.

Case 1. $\eta^T F \neq 0$. We may assume that $|F^T \eta| = 1$.

Let us choose

$$
(2.12) \qquad g = \Big(\int_0^T \beta(s) dW(s) \Big) F^T \eta, \qquad \bar{\sigma}(t) \equiv 0.
$$

Then, by defining

$$
(2.13) \qquad \widehat{\zeta}(t) = \Big(\int_0^t \beta(s) dW(s) \Big) F^T \eta, \qquad t \in [0,T],
$$

we have

$$
(2.14) \qquad \begin{cases} d[\zeta(t) - \widehat{\zeta}(t)] = \bar{h}(t)dt + [\bar{f}(t) - \beta(t)]dW(t), \\ \zeta(T) - \widehat{\zeta}(T) = 0. \end{cases}
$$

Applying Itô's formula to $|\zeta(t) - \widehat{\zeta}(t)|^2$, we obtain

$$
(2.15) \qquad \begin{aligned} & E|\zeta(t) - \widehat{\zeta}(t)|^2 + E\int_t^T |\bar{f}(s) - \beta(s)|^2 ds \\ & = -2E \int_t^T \langle\, \zeta(s) - \widehat{\zeta}(s), \bar{h}(s) \,\rangle\, ds \\ & = 2E \int_t^T \langle \int_s^T \bar{h}(r)dr + \int_s^T [\bar{f}(r) - \beta(r)]dW(r), \bar{h}(s) \,\rangle\, ds \\ & = 2E \int_t^T \langle \int_s^T \bar{h}(r)dr, \bar{h}(s) \,\rangle\, ds \\ & = E\Big| \int_t^T \bar{h}(s)ds \Big|^2 \leq (T-t) \int_t^T E|\bar{h}(s)|^2 ds. \end{aligned}
$$

Consequently, (note $\bar h \in L^2_{\mathcal{G}}(0,T;\mathbb{R})$ and $\bar f \in L^2_{\mathcal{F}}(\Omega; C([0,T];\mathbb{R}))$)

$$
\begin{aligned}
E\int_t^T |\bar f(T)-\beta(s)|^2 ds &\le 2E\int_t^T |\bar f(s)-\beta(s)|^2 ds + 2E\int_t^T |\bar f(T)-\bar f(s)|^2 ds \\
&\le 2(T-t)\int_t^T E|\bar h(s)|^2 ds + 2E\int_t^T |\bar f(T)-\bar f(s)|^2 ds = o(T-t).
\end{aligned}
\tag{2.16}
$$

On the other hand, by the definition of $\beta(\cdot)$, we have

$$
E\int_{T_i}^T |\bar f(T)-\beta(s)|^2 ds = \frac{T-T_i}{2}\Big(E|\bar f(T)-1|^2 + E|\bar f(T)+1|^2\Big), \quad \forall i \ge 1. \tag{2.17}
$$

Clearly, (2.17) contradicts (2.16), which means $\eta^T F \neq 0$ is not possible.

Case 2. $\eta^T F = 0$ and $\eta^T \overline{D} \neq 0$. We may assume that $|\overline{D}^T \eta| = 1$.

In this case, we choose $\bar\sigma(t) = \beta(t)\overline{D}^T\eta$. Thus, (2.10) becomes

$$
\begin{cases}
d\zeta(t) = \bar h(t)dt + [\bar f(t) + \beta(t)]dW(t), & t \in [0,T], \\
\zeta(T) = 0.
\end{cases}
\tag{2.18}
$$

Then, the argument used in Case 1 applies. Thus, $\eta^T\overline{D} \neq 0$ is impossible either. Hence, (2.8) follows. $\square$

Proof of Theorem 2.2. Let $(X,Y,Z) \in \mathcal{M}[0,T]$ be an adapted solution of (1.1). Set $\overline{Y}(t) = Y(t) - GX(t)$. Then, $\overline{Y}(\cdot)$ satisfies the following backward SDE:

$$
\begin{cases}
d\overline{Y} = \big\{(\widehat{A}-GA)X + (\widehat{B}-GB)Y + (\widehat{C}-GC)Z + \widehat{D}\widehat{b} - GDb\big\}dt \\
\qquad\qquad + \big\{(\widehat{A}_1 - GA_1)X + (\widehat{B}_1 - GB_1)Y \\
\qquad\qquad\qquad + (\widehat{C}_1 - GC_1)Z + \widehat{D}_1\widehat{\sigma} - GD_1\sigma\big\}dW(t), \\
\overline{Y}(T) = Fg.
\end{cases}
\tag{2.19}
$$

Denote

$$
\begin{cases}
h = (\widehat{A}-GA)X + (\widehat{B}-GB)Y + (\widehat{C}-GC)Z + \widehat{D}\widehat{b} - GDb, \\
f = (\widehat{A}_1 - GA_1)X + (\widehat{B}_1 - GB_1)Y.
\end{cases}
\tag{2.20}
$$

We see that $h \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ and $f \in L^2_{\mathcal{F}}(\Omega; C([0,T];\mathbb{R}^m))$. One can rewrite (2.19) as follows:

$$
\begin{cases}
d\overline{Y} = hdt + \{f + (\widehat{C}_1 - GC_1)Z + \widehat{D}_1\widehat{\sigma} - GD_1\sigma\}dW(t), \\
\overline{Y}(T) = Fg.
\end{cases}
\tag{2.21}
$$

Then, by Lemma 2.3, we obtain (2.5). The final conclusion is obvious. $\square$

To conclude this section, let us present the following further result, for completeness of the above technique.

Proposition 2.4. *Suppose the assumption of Theorem 2.2 holds. For any b, σ, $\widehat{b}$, $\widehat{\sigma}$, g and x satisfying (2.4), let $(X,Y,Z) \in \mathcal{M}[0,T]$ be an adapted solution of (1.1). Then, it holds*

$$[\widehat{A}_1 - GA_1 + (\widehat{B}_1 - GB_1)G]X(T) + (\widehat{B}_1 - GB_1)Fg \in \mathcal{R}(\widehat{C}_1 - GC_1), \qquad \text{a.s.} \tag{2.22}$$

If, in addition, the following holds:

$$\begin{cases} \mathcal{R}(A+BG) + \mathcal{R}(BF) \subseteq \mathcal{R}(D), \quad \mathcal{R}(A_1 + B_1 G) + \mathcal{R}(B_1 F) \subseteq \mathcal{R}(D_1), \\ \mathcal{R}(\widehat{A} + \widehat{B}G) + \mathcal{R}(\widehat{B}F) \subseteq \mathcal{R}(\widehat{D}), \quad \mathcal{R}(\widehat{A}_1 + \widehat{B}_1 G) + \mathcal{R}(\widehat{B}_1 F) \subseteq \mathcal{R}(\widehat{D}_1), \end{cases} \tag{2.23}$$

then

$$\mathcal{R}\Big(\widehat{A}_1 - GA_1 + (\widehat{B}_1 - GB_1)G\Big) + \mathcal{R}\Big((\widehat{B}_1 - GB_1)F\Big) \subseteq \mathcal{R}(\widehat{C}_1 - GC_1). \tag{2.24}$$

Proof. Suppose $\eta \in \mathbb{R}^m$ such that

$$\eta^T(\widehat{C}_1 - GC_1) = 0. \tag{2.25}$$

Then, by (2.5), one has

$$\eta^T F = 0, \quad \eta^T \widehat{D}_1 = 0, \quad \eta^T G D_1 = 0. \tag{2.26}$$

Hence, from (2.21), we obtain

$$\begin{cases} d[\eta^T \overline{Y}(t)] = \eta^T h(t)dt + \eta^T f(t)dW(t), \quad t \in [0,T], \\ \eta^T \overline{Y}(T) = 0. \end{cases} \tag{2.27}$$

Applying Itô's formula to $|\eta^T \overline{Y}(t)|^2$, we have (similar to (2.15))

$$\begin{aligned} & E|\eta^T \overline{Y}(t)|^2 + E\int_t^T |\eta^T f(s)|^2 ds \\ & \quad = E\Big|\int_t^T \eta^T h(s)ds\Big|^2 \le (T-t)\int_t^T E|\eta^T h(s)|^2 ds. \end{aligned} \tag{2.28}$$

Dividing both sides by $T-t$ and then sending $t \to T$, we obtain

$$E|\eta^T f(T)|^2 = 0. \tag{2.29}$$

By (2.20), and the relation $Y(T) = GX(T) + Fg$, we obtain

$$\eta^T[\widehat{A}_1 - GA_1 + (\widehat{B}_1 - GB_1)G]X(T) + \eta^T(\widehat{B}_1 - GB_1)Fg = 0, \qquad \text{a.s.} \tag{2.30}$$

Thus, (2.22) follows. In the case (2.23) holds, for any $x \in \mathbb{R}^n$ and $g \in \mathbb{R}^m$ (deterministic), by some choice of b, σ, $\widehat{b}$ and $\widehat{\sigma}$, (1.1) admits an adapted solution $(X, Y, Z) \equiv (x, Gx+Fg, 0)$. Then, (2.22) implies (2.24). $\square$

§3. Some Reductions.

In this section, we are going to make some reductions under condition (2.6). We note that (2.6) is very general. It is true if, for example, $F = I \in \mathbb{R}^{m\times m}$, which is the case in many applications. Now, we assume (2.6). By Theorem 2.2, if we want (1.1) to be solvable for all given data, we must have $\widehat{C}_1 - GC_1$ to be onto (and thus $\ell \geq m$). Thus, it is reasonable to make the following assumption:

Assumption A. Let $\ell = m$ and $\widehat{C}_1 - GC_1 \in \mathbb{R}^{m\times m}$ be is invertible.

Let us make some reductions under Assumption A. Set $\overline{Y} = Y - GX$. Then, $\overline{Y}(T) = Fg$ and (see (2.19))

$$\begin{aligned}
d\overline{Y} &= (\widehat{A}X + \widehat{B}Y + \widehat{C}Z + \widehat{D}\widehat{b})dt + (\widehat{A}_1 X + \widehat{B}_1 Y + \widehat{C}_1 Z + \widehat{D}_1\widehat{\sigma})dW \\
&\quad - G(AX + BY + CZ + Db)dt - G(A_1X + B_1Y + C_1Z + D_1\sigma)dW \\
&= \Big\{\big[\widehat{A} - GA + (\widehat{B} - GB)G\big]X + (\widehat{B} - GB)\overline{Y} + (\widehat{C} - GC)Z + \widehat{D}\widehat{b} - GDb\Big\}dt \\
&\quad + \Big\{\big[\widehat{A}_1 - GA_1 + (\widehat{B}_1 - GB_1)G\big]X + (\widehat{B}_1 - GB_1)\overline{Y} \\
&\qquad + (\widehat{C}_1 - GC_1)Z + \widehat{D}_1\widehat{\sigma} - GD_1\sigma\Big\}dW.
\end{aligned} \tag{3.1}$$

Define

$$\begin{aligned}
\overline{Z} = &\big[\widehat{A}_1 - GA_1 + (\widehat{B}_1 - GB_1)G\big]X + (\widehat{B}_1 - GB_1)\overline{Y} \\
&+ (\widehat{C}_1 - GC_1)Z + \widehat{D}_1\widehat{\sigma} - GD_1\sigma.
\end{aligned} \tag{3.2}$$

Since $(\widehat{C}_1 - GC_1)$ is invertible, we have

$$
\begin{aligned}
Z = (\widehat{C}_1 - GC_1)^{-1}\big\{\overline{Z} - [\widehat{A}_1 - GA_1 + (\widehat{B}_1 - GB_1)G]X \\
- (\widehat{B}_1 - GB_1)\overline{Y} - (\widehat{D}_1\widehat{\sigma} - GD_1\sigma)\big\}.
\end{aligned} \tag{3.3}
$$

Then, it follows that

$$
\begin{cases}
dX = \Big(\overline{A}X + \overline{B}\,\overline{Y} + \overline{C}\,\overline{Z} + \overline{b}\Big)dt + \Big(\overline{A}_1 X + \overline{B}_1\overline{Y} + \overline{C}_1\overline{Z} + \overline{\sigma}\Big)dW, \\
d\overline{Y} = \Big(\overline{A}_0 X + \overline{B}_0\overline{Y} + \overline{C}_0\overline{Z} + \overline{h}\Big)dt + \overline{Z}dW, \\
X(0) = x, \qquad \overline{Y}(T) = Fg,
\end{cases} \tag{3.4}
$$

where

$$
\begin{cases}
\overline{A} = A + BG - C(\widehat{C}_1 - GC_1)^{-1}[\widehat{A}_1 - GA_1 + (\widehat{B}_1 - GB_1)G], \\
\overline{B} = B - C(\widehat{C}_1 - GC_1)^{-1}(\widehat{B}_1 - GB_1), \\
\overline{C} = C(\widehat{C}_1 - GC_1)^{-1}, \\
\overline{b} = Db - C(\widehat{C}_1 - GC_1)^{-1}(\widehat{D}_1\widehat{\sigma} - GD_1\sigma), \\
\overline{A}_1 = A_1 + B_1 G - C_1(\widehat{C}_1 - GC_1)^{-1}[\widehat{A}_1 - GA_1 + (\widehat{B}_1 - GB_1)G], \\
\overline{B}_1 = B_1 - C_1(\widehat{C}_1 - GC_1)^{-1}(\widehat{B}_1 - GB_1), \\
\overline{C}_1 = C_1(\widehat{C}_1 - GC_1)^{-1}, \\
\overline{\sigma} = D_1\sigma - C_1(\widehat{C}_1 - GC_1)^{-1}(\widehat{D}_1\widehat{\sigma} - GD_1\sigma), \\
\overline{A}_0 = \widehat{A} - GA + (\widehat{B} - GB)G \\
\qquad\qquad - (\widehat{C} - GC)(\widehat{C}_1 - GC_1)^{-1}[\widehat{A}_1 - GA_1 + (\widehat{B}_1 - GB_1)G], \\
\overline{B}_0 = \widehat{B} - GB - (\widehat{C} - GC)(\widehat{C}_1 - GC_1)^{-1}(\widehat{B}_1 - GB_1), \\
\overline{C}_0 = (\widehat{C} - GC)(\widehat{C}_1 - GC_1)^{-1}, \\
\overline{h} = \widehat{D}\widehat{b} - GDb - (\widehat{C} - GC)(\widehat{C}_1 - GC_1)^{-1}(\widehat{D}_1\widehat{\sigma} - GD_1\sigma).
\end{cases} \tag{3.5}
$$

The above tells us that under Assumption A, (1.1) and (3.4) are equivalent. Next, we denote

$$
\begin{cases}
\overline{\mathcal{A}} = \begin{pmatrix} \overline{A} & \overline{B} \\ \overline{A}_0 & \overline{B}_0 \end{pmatrix}, \quad \overline{\mathcal{C}} = \begin{pmatrix} \overline{C} \\ \overline{C}_0 \end{pmatrix}, \\
\overline{\mathcal{A}}_1 = \begin{pmatrix} \overline{A}_1 & \overline{B}_1 \\ 0 & 0 \end{pmatrix}, \quad \overline{\mathcal{C}}_1 = \begin{pmatrix} \overline{C}_1 \\ I \end{pmatrix}.
\end{cases} \tag{3.6}
$$

Let $\Psi(\cdot)$ be the solution of the following:

$$
\tag{3.7}
\begin{cases}
d\Psi(t) = \overline{\mathcal{A}}\Psi(t)dt + \overline{\mathcal{A}}_1\Psi(t)dW(t), & t \ge 0, \\
\Psi(0) = I.
\end{cases}
$$

Then, (3.4) is equivalent to the following: For some $y \in \mathbb{R}^m$,

$$
\tag{3.8}
\begin{aligned}
\begin{pmatrix} X(t) \\ \overline{Y}(t) \end{pmatrix} &= \Psi(t)\begin{pmatrix} x \\ y \end{pmatrix} + \Psi(t)\int_0^t \Psi(s)^{-1}\Big[\overline{\mathcal{C}}\,\overline{Z}(s) + \begin{pmatrix} \overline{b}(s) \\ \overline{h}(s) \end{pmatrix}\Big]ds \\
&\quad + \Psi(t)\int_0^t \Psi(s)^{-1}\Big[\overline{\mathcal{C}}_1\overline{Z}(s) + \begin{pmatrix} \overline{\sigma}(s) \\ 0 \end{pmatrix}\Big]dW(s), \quad t \in [0,T],
\end{aligned}
$$

with the property that

$$
\tag{3.9}
\begin{aligned}
Fg &= (0,I)\Psi(T)\begin{pmatrix} x \\ y \end{pmatrix} + (0,I)\Psi(T)\int_0^T \Psi(s)^{-1}\Big[\overline{\mathcal{C}}\,\overline{Z}(s) + \begin{pmatrix} \overline{b}(s) \\ \overline{h}(s) \end{pmatrix}\Big]ds \\
&\quad + (0,I)\Psi(T)\int_0^T \Psi(s)^{-1}\Big[\overline{\mathcal{C}}_1\overline{Z}(s) + \begin{pmatrix} \overline{\sigma}(s) \\ 0 \end{pmatrix}\Big]dW(s).
\end{aligned}
$$

Clearly, (3.9) is equivalent to the following: For some $y \in \mathbb{R}^m$ and $\overline{Z}(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$, it holds

$$
\tag{3.10}
\begin{aligned}
\eta \equiv Fg &- (0,I)\Psi(T)\begin{pmatrix} x \\ 0 \end{pmatrix} - (0,I)\Psi(T)\int_0^T \Psi(s)^{-1}\begin{pmatrix} \overline{b}(s) \\ \overline{h}(s) \end{pmatrix}ds \\
&- (0,I)\Psi(T)\int_0^T \Psi(s)^{-1}\begin{pmatrix} \overline{\sigma}(s) \\ 0 \end{pmatrix}dW(s) \\
= (0,I)&\Psi(T)\begin{pmatrix} 0 \\ y \end{pmatrix} + (0,I)\Psi(T)\int_0^T \Psi(s)^{-1}\overline{\mathcal{C}}\,\overline{Z}(s)ds \\
&+ (0,I)\Psi(T)\int_0^T \Psi(s)^{-1}\overline{\mathcal{C}}_1\,\overline{Z}(s)dW(s).
\end{aligned}
$$

Thus, if we can solve the following:

$$
\tag{3.11}
\begin{cases}
d\begin{pmatrix} \widetilde{X} \\ \widetilde{Y} \end{pmatrix} = \Big\{\overline{\mathcal{A}}\begin{pmatrix} \widetilde{X} \\ \widetilde{Y} \end{pmatrix} + \overline{\mathcal{C}}\,\widetilde{Z}\Big\}dt + \Big\{\overline{\mathcal{A}}_1\begin{pmatrix} \widetilde{X} \\ \widetilde{Y} \end{pmatrix} + \overline{\mathcal{C}}_1\widetilde{Z}\Big\}dW, \\
\widetilde{X}(0) = 0, \qquad \widetilde{Y}(T) = \eta,
\end{cases}
$$

with η being given by (3.10), then for such a pair $y \equiv \widetilde{Y}(0)$ and $\overline{Z}(\cdot) \equiv \widetilde{Z}(\cdot)$, by setting $(X, \overline{Y})$ as (3.8), we obtain an adapted solution $(X, \overline{Y}, \overline{Z}) \in \mathcal{M}[0,T]$ of (3.4). The above procedure is reversible. Thus, by the equivalence between (3.4) and (1.1), we actually have the equivalence between the solvability of (1.1) and (3.11). Let us state this result as follows.

Theorem 3.1. *Let $F = I \in \mathbb{R}^{m\times m}$ and $\ell = m$. Then, (1.1) is solvable for all b, σ, $\widehat{b}$, $\widehat{\sigma}$, x and g satisfying (2.4) if and only if (3.11) is solvable for all $\eta \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^m)$.*

We note that by Theorem 2.2, $F = I$ and $\ell = m$ imply Assumption A. Based on the above reduction, in what follows, we concentrate on the following FBSDEs:

$$
\begin{cases}
dX = \Big(AX + BY + CZ\Big)dt + \Big(A_1X + B_1Y + C_1Z\Big)dW, \\
dY = \Big(\widehat{A}X + \widehat{B}Y + \widehat{C}Z\Big)dt + ZdW, \\
X(0) = 0, \qquad Y(T) = g.
\end{cases} \qquad t \in [0,T], \tag{3.12}
$$

Also, we will denote

$$
\begin{cases}
\mathcal{A} = \begin{pmatrix} A & B \\ \widehat{A} & \widehat{B} \end{pmatrix}, \quad \mathcal{C} = \begin{pmatrix} C \\ \widehat{C} \end{pmatrix}, \\
\mathcal{A}_1 = \begin{pmatrix} A_1 & B_1 \\ 0 & 0 \end{pmatrix}, \quad \mathcal{C}_1 = \begin{pmatrix} C_1 \\ I \end{pmatrix}.
\end{cases} \tag{3.13}
$$

and will let Φ be the solution of the following:

$$
\begin{cases}
d\Phi(t) = \mathcal{A}\Phi(t)dt + \mathcal{A}_1\Phi(t)dW(t), \qquad t \in [0,T], \\
\Phi(0) = I.
\end{cases} \tag{3.14}
$$

If we regard (X,Y) as the *state* and Z as the *control*, (3.12) is called a (linear) *stochastic control system.* Then, the solvability of (3.12) becomes the following *controllability* problem: For give $g \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^m)$, find a control $Z \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$, such that some initial state $(X(0),Y(0)) \in \{0\}\times\mathbb{R}^m$ can be steered to the final state $(X(T),Y(T)) \in \mathbb{R}^n \times \{g\}$. This can be referred to as the controllability of the system (3.12) from $\{0\}\times\mathbb{R}^m$ to $\mathbb{R}^n\times\{g\}$. We note that g is an $\mathcal{F}_T$-measurable square integrable random Φ vector, and we need exactly control $Y(T)$ to g. To the best knowledge of this author, such a controllability problem has not been discussed in the literature.

§4. Solvability of Linear FBSDEs.

In this section, we are going to present some solvability results for linear FBSDEs (3.12). The basic idea is adopted from the study of controllability in control theory. For convenience, we denote hereafter that $H = L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^m)$ and $\mathcal{H} = L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ (which are Hilbert spaces to which the final datum g and the process $Z(\cdot)$ belong, respectively).

First of all, we recall that if Φ is the solution of (3.14), then, Φ^{-1} exists and it satisfies the following linear SDE:

$$
\begin{cases} d\Phi^{-1} = \big[- \Phi^{-1}\mathcal{A} + \Phi^{-1}\mathcal{A}_1\mathcal{A}_1\big]dt - \Phi^{-1}\mathcal{A}_1 dW(t), & t \ge 0, \\ \Phi^{-1}(0) = I. \end{cases} \tag{4.1}
$$

Moreover, $(X,Y,Z) \in \mathcal{M}[0,T]$ is an adapted solution of (3.12) if and only if the following variation of constant formula holds:

$$
\begin{aligned} \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = \Phi(t)\begin{pmatrix} 0 \\ y \end{pmatrix} &+ \Phi(t)\int_0^t \Phi(s)^{-1}\mathcal{C}Z(s)ds \\ &+ \Phi(t)\int_0^t \Phi(s)^{-1}\mathcal{C}_1 Z(s)dW(s), \quad t \in [0,T], \end{aligned} \tag{4.2}
$$

for some $y \in \mathbb{R}^m$ and with the property:

$$
\begin{aligned} g = (0,I)\Big\{\Phi(T)\begin{pmatrix} 0 \\ y \end{pmatrix} &+ \Phi(T)\int_0^T \Phi(s)^{-1}\mathcal{C}Z(s)ds \\ &+ \Phi(T)\int_0^T \Phi(s)^{-1}\mathcal{C}_1 Z(s)dW(s)\Big\}. \end{aligned} \tag{4.3}
$$

Let us introduce an operator $\mathcal{K} : \mathcal{H} \to H$ as follows:

$$
\mathcal{K}Z = (0,I)\Big\{\Phi(T)\int_0^T \Phi(s)^{-1}\mathcal{C}Z(s)ds + \Phi(T)\int_0^T \Phi(s)^{-1}\mathcal{C}_1 Z(s)dW(s)\Big\}. \tag{4.4}
$$

Then, for given $g \in H$, finding adapted solutions to (3.12) amounts to the following: Find $y \in \mathbb{R}^m$ and $Z \in \mathcal{H}$, such that

$$
g = (0,I)\Phi(T)\begin{pmatrix} 0 \\ I \end{pmatrix} y + \mathcal{K}Z, \tag{4.5}
$$

and define (X,Y) as in (4.2), then $(X,Y,Z) \in \mathcal{M}[0,T]$ is an adapted solution of (3.12). Hence, the study of operators $\Phi(T)$ and $\mathcal{K}$ is crucial to the solvability of linear FBSDEs (3.12). We now make some investigations on $\Phi(\cdot)$ and $\mathcal{K}$. Let us first give the following lemma.

Lemma 4.1. *For any $f \in L^1_{\mathcal{F}}(0,T;\mathbb{R}^{n+m})$ and $h \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^{n+m})$, it holds*

$$
(4.6)\qquad \begin{cases} E\Phi(t) = e^{\mathcal{A}t}, \\ E\Big\{\Phi(t)\displaystyle\int_0^t \Phi(s)^{-1} f(s)ds\Big\} = \displaystyle\int_0^t e^{\mathcal{A}(t-s)} Ef(s)ds, \\ E\Big\{\Phi(t)\displaystyle\int_0^t \Phi(s)^{-1} h(s)dW(s)\Big\} = 0, \end{cases} \qquad t \in [0,T].
$$

Also, it holds that

$$
(4.7)\qquad E\sup_{0\le t\le T}|\Phi(t)|^{2k}, \quad E\sup_{0\le t\le T}|\Phi(t)^{-1}|^{2k} < \infty, \qquad \forall k \ge 1.
$$

Proof. Let us first prove the second equality in (4.6). The other two in (4.6) can be proved similarly. Set

$$
(4.8)\qquad \xi(t) = \Phi(t)\int_0^t \Phi(s)^{-1} f(s)ds, \qquad t \in [0,T].
$$

Then, $\xi(\cdot)$ satisfies the following SDE:

$$
(4.9)\qquad \begin{cases} d\xi(t) = [\mathcal{A}\xi(t) + f(t)]dt + \mathcal{A}_1\xi(t)dW(t), \qquad t \in [0,T], \\ \xi(0) = 0. \end{cases}
$$

Taking expectation in (4.9), we obtain

$$
(4.10)\qquad \begin{cases} d[E\xi(t)] = [\mathcal{A}E\xi(t) + Ef(t)]dt, \qquad t \in [0,T], \\ E\xi(0) = 0. \end{cases}
$$

Thus,

$$
(4.11)\qquad E\xi(t) = \int_0^t e^{\mathcal{A}(t-s)} Ef(s)ds, \qquad t \in [0,T],
$$

proving our claim.

Now, we prove (4.7). For any $\xi_0 \in \mathbb{R}^{n+m}$, process $\xi(t) \stackrel{\Delta}{=} \Phi(t)\xi_0$ satisfies the following SDE:

$$
(4.12)\qquad \begin{cases} d\xi(t) = \mathcal{A}\xi(t)dt + \mathcal{A}_1\xi(t)dW(t), \qquad t \in [0,T], \\ \xi(0) = \xi_0. \end{cases}
$$

Then, by Itô's formula, Burkholder-Davis-Gundy's inequality ([10]) and Gronwall's inequality, we can show that

$$E \sup_{0\le t\le T} |\xi(t)|^{2k} \le C|\xi_0|^{2k}, \qquad k \ge 1. \tag{4.13}$$

Thus, the first inequality in (4.7) follows. The second one can be proved in the same way. $\square$

From (4.7), we see that $\mathcal{K} : \mathcal{H} \to H$ is a bounded linear operator. Now, applying (4.6) to (4.3), we obtain that (3.12) admits an adapted solution, then

$$Eg = (0, I)\Big\{ e^{\mathcal{A}T} \begin{pmatrix} 0 \\ I \end{pmatrix} y + \int_0^T e^{\mathcal{A}(T-s)} \mathcal{C} EZ(s) ds \Big\}, \tag{4.14}$$

for some $y \in \mathbb{R}^m$ and $EZ(\cdot) \in L^2(0, T; \mathbb{R}^m)$. This leads to the following necessary condition for the solvability of (3.12).

Theorem 4.2. *Suppose (3.12) is solvable for all $g \in H$. Then,*

$$\operatorname{rank}\Big\{ (0, I)\Big(e^{\mathcal{A}T} \begin{pmatrix} 0 \\ I \end{pmatrix}, \mathcal{C}, \mathcal{A}\mathcal{C}, \cdots, \mathcal{A}^{n+m-1}\mathcal{C} \Big) \Big\} = m. \tag{4.15}$$

In particular, if $\mathcal{C} = 0$, then, (4.15) can be replaced by

$$\det \Big\{ (0, I) e^{\mathcal{A}T} \begin{pmatrix} 0 \\ I \end{pmatrix} \Big\} \ne 0. \tag{4.16}$$

Proof. It suffices to note that (see [19], for example) the range of the operator

$$u(\cdot) \mapsto \int_0^T e^{\mathcal{A}(T-s)} \mathcal{C} u(s) ds, \qquad \forall u(\cdot) \in L^2(0, T; \mathbb{R}^m),$$

is given by

$$\mathcal{R}(\mathcal{C}) + \mathcal{R}(\mathcal{A}\mathcal{C}) + \cdots + \mathcal{R}(\mathcal{A}^{n+m-1}\mathcal{C}).$$

Then, we have (4.15). $\square$

We note that in the case $\mathcal{C} = 0$, (4.16) amounts to say that the FBSDEs (3.12) is solvable for all $g \in H$ implies that the corresponding two-point boundary value problem for ODEs:

$$\begin{cases} \begin{pmatrix} \dot X(t) \\ \dot Y(t) \end{pmatrix} = \mathcal{A} \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix}, & t \in [0, T], \\ X(0) = 0, \qquad Y(T) = \bar g, \end{cases} \tag{4.17}$$

admits a solution for all $\bar g \in \mathbb{R}^m$. In [20], it was proved that a little stronger condition than (4.16) is also sufficient for the solvability of (3.12) if A_1, B_1, C_1, C and $\widehat{C}$ are all zero (note since $g \in H$, (3.12) is still a FBSDEs). We will extend that result below.

On the other hand, we note that condition (4.15) implies that the (deterministic) control system $[\mathcal{A}, \mathcal{C}]$:

$$
\begin{pmatrix} \dot X(t) \\ \dot Y(t) \end{pmatrix} = \mathcal{A} \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} + \mathcal{C} Z(t), \tag{4.18}
$$

is controllable from $\{0\} \times \mathbb{R}^m$ to $\mathbb{R}^n \times \{\bar g\}$ for any $\bar g \in \mathbb{R}^m$.

Let us now present another necessary condition for the solvability of (3.12).

Theorem 4.3. *Let $\mathcal{C} = 0$. Suppose (3.12) is solvable for all $g \in H$. Then,*

$$
\det\left\{(0, I) e^{\mathcal{A}t} \mathcal{C}_1\right\} > 0, \qquad \forall t \in [0, T]. \tag{4.19}
$$

Consequently, if

$$
\widehat{T} = \inf\{T > 0 \mid \det\left[(0, I) e^{\mathcal{A}T} \mathcal{C}_1\right] = 0\} < \infty, \tag{4.20}
$$

then, for any $T \geq \widehat{T}$, there exists a $g \in H$, such that (3.12) is not solvable.

Remark 4.4. The above result reveals a significant difference between the solvability of FBSDEs and that of two-point boundary value problems for ordinary differential equations. We note that for (4.17) to be solvable for all $\bar g \in \mathbb{R}^m$, if and only if (4.16) holds. Since the function $t \mapsto \det\left\{(0, I) e^{\mathcal{A}t} \begin{pmatrix} 0 \\ I \end{pmatrix}\right\}$ is analytic (and it is equal to 1 at $t = 0$), except at most a discrete set of T's, (4.16) holds. That implies that for any $T_0 \in (0, \infty)$, if it happens that (4.17) is not solvable for $T = T_0$ with some $\bar g \in \mathbb{R}^m$, then, at some later time $T > T_0$, (4.17) will be solvable again for all $\bar g \in \mathbb{R}^m$. But, in the above FBSDEs case, if $\widehat{T} < \infty$, then for any $T \geq \widehat{T}$, we can always find a $g \in H$, such that (3.12) (with $\mathcal{C} = 0$) is not solvable. Thus, besides other differences, FBSDEs and the two-point boundary value problem for ODEs is also significantly different as far as the solvable duration is concerned.

Proof of Theorem 4.3. Suppose there exists an $s_0 \in [0, T)$, such that

$$
\det\left\{(0, I) e^{\mathcal{A}(T - s_0)} \mathcal{C}_1\right\} = 0. \tag{4.21}
$$

Note that $s_0 < T$. Then, there exists an $\eta \in \mathbb{R}^m$, $|\eta| = 1$, such that

$$\eta^T(0,I)e^{\mathcal{A}(T-s_0)}\mathcal{C}_1 = 0. \tag{4.22}$$

We are going to prove that for any $\varepsilon > 0$ with $s_0 + \varepsilon \le T$, there exists a $g \in L^2_{\mathcal{F}_{s_0+\varepsilon}}(\Omega;\mathbb{R}^m) \subseteq H$, such that (3.12) has no adapted solutions. to this end, we let $\beta : [0,T] \to \mathbb{R}$ be a Lebesgue measurable function such that

$$\begin{cases} \beta(s) = \pm 1, \quad \forall s \in [0,T]; \\ |\{s \in [s_0, s_k] \bigm| \beta(s) = 1\}| = |\{s \in [s_0, s_k] \bigm| \beta(s) = -1\}| = \dfrac{s_k - s_0}{2}, \qquad k \ge 1, \end{cases} \tag{4.23}$$

for some sequence $s_k \downarrow s_0$ and $s_k \le T - \varepsilon$. Next, we define

$$\zeta(t) = \int_0^t \beta(s)dW(s), \qquad t \in [0,T], \tag{4.24}$$

and take $g = \zeta(T)\eta \in L^2_{\mathcal{F}_{s_0+\varepsilon}}(\Omega;\mathbb{R}^m) \subseteq H$. Suppose (3.12) admits an adapted solution $(X,Y,Z) \in \mathcal{M}[0,T]$ for this g. Then, for some $y \in \mathbb{R}^m$, we have (remember $\mathcal{C} = 0$)

$$\zeta(T)\eta = (0,I)\Big\{ e^{\mathcal{A}T}\begin{pmatrix} 0 \\ y \end{pmatrix} + \int_0^T e^{\mathcal{A}(T-s)}\Big[\mathcal{A}_1\begin{pmatrix} X(s) \\ Y(s) \end{pmatrix} + \mathcal{C}_1 Z(s)\Big]dW(s)\Big\}. \tag{4.25}$$

Applying η^T from left to (4.25) gives the following:

$$\zeta(T) = \alpha + \int_0^T \big\{\gamma(s) + \langle\, \psi(s), Z(s)\,\rangle\,\big\}dW(s), \tag{4.26}$$

where

$$\begin{cases} \alpha = \eta^T(0,I)e^{\mathcal{A}T}\begin{pmatrix} 0 \\ y \end{pmatrix} \in \mathbb{R}, \\ \gamma(\cdot) = \eta^T(0,I)e^{\mathcal{A}(T-\cdot)}\mathcal{A}_1\begin{pmatrix} X(\cdot) \\ Y(\cdot) \end{pmatrix} \in L^2_{\mathcal{F}}(\Omega; C([0,T];\mathbb{R})), \\ \psi(\cdot) = \Big[\eta^T(0,I)e^{\mathcal{A}(T-\cdot)}\mathcal{C}_1\Big]^T \text{ is analytic, } \psi(s_0) = 0. \end{cases} \tag{4.27}$$

Let us denote

$$\theta(t) = \alpha + \int_0^t [\gamma(s) + \langle\, \psi(s), Z(s)\,\rangle]dW(s), \qquad t \in [0,T]. \tag{4.28}$$

Then, it follows that

$$\begin{cases} d[\theta(t) - \zeta(t)] = [\gamma(t) + \langle\, \psi(t), Z(t)\,\rangle - \beta(t)]dW(t), \quad t \in [0,T], \\ [\theta(T) - \zeta(T)] = 0. \end{cases} \tag{4.29}$$

By Itô's formula, we have

$$0 = E|\theta(t) - \zeta(t)|^2 + E\int_t^T |\gamma(s) + \langle\, \psi(s), Z(s)\,\rangle - \beta(s)|^2 ds, \quad t \in [0,T]. \tag{4.30}$$

Thus,

$$\beta(s) - \gamma(s) = \langle\, \psi(s), Z(s)\,\rangle, \qquad \text{a.e. } s \in [0,T], \text{ a.s.} \tag{4.31}$$

which yields

$$\int_{s_0}^{s_k} E|\beta(s) - \gamma(s)|^2 ds = \int_{s_0}^{s_k} E|\,\langle\, \psi(s), Z(s)\,\rangle\,|^2 ds, \quad \forall k \geq 1. \tag{4.32}$$

Now, we observe that (note $\gamma \in L^2_{\mathcal{F}}(\Omega; C([0,T]; \mathbb{R}))$ and (4.23))

$$\begin{aligned} \int_{s_0}^{s_k} E|\beta(s) - \gamma(s)|^2 ds &\geq \frac{1}{2}\int_{s_0}^{s_k} E|\beta(s) - \gamma(s_0)|^2 ds - \int_{s_k}^{s_0} E|\gamma(s) - \gamma(s_0)|^2 ds \\ &\geq \frac{s_k - s_0}{4} E\Big[|1 - \gamma(s_0)|^2 + |1 + \gamma(s_0)|^2\Big] - o(s_k - s_0), \quad k \geq 1. \end{aligned} \tag{4.33}$$

On the other hand, since $\psi(\cdot)$ is analytic with $\psi(s_0) = 0$, we must must have

$$\psi(s) = (s - s_0)\widetilde{\psi}(s), \qquad s \in [0,T], \tag{4.34}$$

for some $\widetilde{\psi}(\cdot)$ which is analytic and hence bounded on $[0,T]$. Consequently,

$$\int_{s_0}^{s_k} E|\,\langle\, \psi(s), Z(s)\,\rangle\,|^2 ds \leq C(s_k - s_0)^2 \int_{s_0}^{s_k} E|Z(s)|^2 ds. \tag{4.35}$$

Hence, (4.32)–(4.33) and (4.35) imply

$$\begin{aligned} &\frac{s_k - s_0}{4} E\Big[|1 - \gamma(s_0)|^2 + |1 + \gamma(s_0)|^2\Big] - o(s_k - s_0) \\ &\qquad \leq C(s_k - s_0)^2 \int_{s_0}^{s_k} E|Z(s)|^2 ds, \quad \forall k \geq 1. \end{aligned} \tag{4.36}$$

This is impossible. Finally, noting the fact that $\det\{(0, I)e^{\mathcal{A}t}\mathcal{C}_1\}\big|_{t=0} = 1$, we obtain (4.19). The final assertion is clear. □

It is not clear if the above result holds for the case $\mathcal{C} \neq 0$ since the assumption $\mathcal{C} = 0$ is crucial in the proof.

Let us now present some results on the operator $\mathcal{K}$.

Lemma 4.5. *The range $\mathcal{R}(\mathcal{K})$ of $\mathcal{K}$ is closed in H.*

Proof. Let us denote $H_0 = L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ and $\widehat{H} = H_0 \times H \equiv L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^{n+m})$. Define

$$\widehat{\mathcal{K}}Z = \Phi(T)\int_0^T \Phi(s)^{-1}\mathcal{C}Z(s)ds + \Phi(T)\int_0^T \Phi(s)^{-1}\mathcal{C}_1 Z(s)dW(s), \quad Z \in \mathcal{H}. \tag{4.37}$$

Then, by (4.7), $\widehat{\mathcal{K}}$ is a bounded linear operator and $\mathcal{K} = (0, I)\widehat{\mathcal{K}}$. We claim that the range $\mathcal{R}(\widehat{\mathcal{K}})$ of $\widehat{\mathcal{K}}$ is closed in $\widehat{H}$. To show this, let us take any convergence sequence

$$\begin{pmatrix} X_k(T) \\ Y_k(T) \end{pmatrix} \equiv \widehat{\mathcal{K}}Z_k \to \zeta, \qquad \text{in } \widehat{H}, \tag{4.38}$$

where (X_k, Y_k) is the solution of the following:

$$\begin{cases} d\begin{pmatrix} X_k \\ Y_k \end{pmatrix} = \Big\{\mathcal{A}\begin{pmatrix} X_k \\ Y_k \end{pmatrix} + \mathcal{C}Z_k\Big\}dt + \Big\{\begin{pmatrix} X_k \\ Y_k \end{pmatrix} + \mathcal{C}_1 Z_k\Big\}dW(t), \\ \begin{pmatrix} X_k(0) \\ Y_k(0) \end{pmatrix} = 0. \end{cases} \tag{4.39}$$

Then, by Itô's formula, we have

$$\begin{aligned} &E\Big\{|X_k(t)|^2 + |Y_k(t)|^2 + \int_t^T \Big|\mathcal{A}_1\begin{pmatrix} X_k(s) \\ Y_k(s) \end{pmatrix} + \mathcal{C}_1 Z_k(s)\Big|^2 ds\Big\} \\ &= E\Big\{|X_k(T)|^2 + |Y_k(T)|^2 - 2\int_t^T \langle \begin{pmatrix} X_k(s) \\ Y_k(s) \end{pmatrix}, \mathcal{A}\begin{pmatrix} X_k(s) \\ Y_k(s) \end{pmatrix} + \mathcal{C}Z_k(s) \rangle\, ds\Big\}. \end{aligned} \tag{4.40}$$

We note that (recall $\mathcal{C}_1 = \begin{pmatrix} C_1 \\ I \end{pmatrix}$)

$$\begin{aligned} &\Big|\mathcal{A}_1\begin{pmatrix} X_k \\ Y_k \end{pmatrix} + \mathcal{C}_1 Z_k\Big|^2 \\ &= \langle (I + C_1^T C_1)Z_k, Z_k \rangle + \Big|\mathcal{A}_1\begin{pmatrix} X_k \\ Y_k \end{pmatrix}\Big|^2 + 2\langle \mathcal{C}_1^T\mathcal{A}_1\begin{pmatrix} X_k \\ Y_k \end{pmatrix}, Z_k \rangle \\ &\geq \frac{1}{2}|Z_k|^2 - C(|X_k|^2 + |Y_k|^2), \end{aligned} \tag{4.41}$$

for some constant $C > 0$. Thus, (4.40) implies

$$\begin{aligned} &E\{|X_k(t)|^2 + |Y_k(t)|^2 + \int_t^T |Z_k(s)|^2 ds\} \\ &\leq CE\{|X_k(T)|^2 + |Y_k(T)|^2 + \int_t^T \big(|X_k(s)|^2 + |Y_k(s)|^2\big)ds\}, \quad t \in [0, T]. \end{aligned} \tag{4.42}$$

Using Gronwall's inequality, we obtain

$$(4.43)\quad E\Big\{|X_k(t)|^2+|Y_k(t)|^2+\int_t^T|Z_k(s)|^2ds\Big\}\le CE\Big\{|X_k(T)|^2+|Y_k(T)|^2\Big\},\quad t\in[0,T].$$

From the convergence (4.38), we see that Z_k is bounded in $\mathcal{H}$. Thus, we may assume that $Z_k\to\widetilde{Z}$ weakly in $\mathcal{H}$. Then, it is easy to see that $\widehat{\mathcal{K}}\widetilde{Z}=\zeta$, proving the closeness of $\mathcal{R}(\widehat{\mathcal{K}})$.

Now, $\mathcal{R}(\widehat{\mathcal{K}})$ is a Hilbert space with the induced inner product of $\widehat{H}$. In this space, we define an orthogonal projection $P_H:\widehat{H}\to H$ by the following:

$$(4.44)\qquad P_H\begin{pmatrix}\xi\\ \eta\end{pmatrix}=\begin{pmatrix}0\\ \eta\end{pmatrix},\qquad\forall\begin{pmatrix}\xi\\ \eta\end{pmatrix}\in\widehat{H}\equiv H_0\times H.$$

Then, the space

$$(4.45)\qquad P_H(\mathcal{R}(\widehat{\mathcal{K}}))=\{0\}\times\mathcal{R}(\mathcal{K})$$

is closed in $\mathcal{R}(\widehat{\mathcal{K}})$ and so is in $\widehat{H}$. Hence, $\mathcal{R}(\mathcal{K})$ is closed in H. $\square$

The following result gives some more information for the operator $\mathcal{K}$ when $\mathcal{C}=0$.

Lemma 4.6. *Let $\mathcal{C}=0$ and let (4.19) hold. Then,*

$$(4.46)\qquad \mathcal{R}(\mathcal{K})=\{\eta\in H\bigm| E\eta=0\}\stackrel{\Delta}{=}\mathcal{N}(E),$$

$$(4.47)\qquad \mathcal{N}(\mathcal{K})\stackrel{\Delta}{=}\{Z\in\mathcal{H}\bigm|\mathcal{K}Z=0\}=\{0\}.$$

Proof. First of all, by Lemma 4.4 (with $\mathcal{C}=0$), we see that $\mathcal{R}(\mathcal{K})$ is closed. Also, by Lemma 4.1, $\mathcal{R}(\mathcal{K})\subseteq\mathcal{N}(E)$. Thus, to show (4.46), it suffices to show that

$$(4.48)\qquad \mathcal{N}(E)\bigcap\mathcal{R}(\mathcal{K})^\perp=\{0\}.$$

We now prove (4.48). Take $\eta\in\mathcal{N}(E)$. Suppose

$$(4.49)\qquad 0=E\,\langle\,\eta,\mathcal{K}Z\,\rangle=E\,\langle\,\eta,(0,I)\Phi(T)\int_0^T\Phi(s)^{-1}\mathcal{C}_1Z(s)dW(s)\,\rangle,\quad\forall Z\in\mathcal{H}.$$

Denote

$$(4.50)\qquad \begin{pmatrix}\overline{X}(t)\\ \overline{Y}(t)\end{pmatrix}=\Phi(t)\int_0^t\Phi(s)^{-1}\mathcal{C}_1Z(s)dW(s),\qquad t\in[0,T].$$

Then,

$$
\left\{
\begin{aligned}
& d\begin{pmatrix} \overline{X} \\ \overline{Y} \end{pmatrix} = \mathcal{A}\begin{pmatrix} \overline{X} \\ \overline{Y} \end{pmatrix} dt + \Big\{\mathcal{A}_1 \begin{pmatrix} \overline{X} \\ \overline{Y} \end{pmatrix} + \mathcal{C}_1 Z\Big\} dW(t), \\
& \begin{pmatrix} \overline{X}(0) \\ \overline{Y}(0) \end{pmatrix} = 0.
\end{aligned}
\right.
\tag{4.51}
$$

By Itô's formula and Gronwall's inequality, we obtain

$$
E\{|\overline{X}(t)|^2 + |\overline{Y}(t)|^2\} \le C \int_0^t E|Z(s)|^2 ds, \qquad t \in [0,T]. \tag{4.52}
$$

Also, we have

$$
\begin{pmatrix} \overline{X}(t) \\ \overline{Y}(t) \end{pmatrix} = \int_0^t e^{\mathcal{A}(t-s)} \Big\{\mathcal{A}_1 \begin{pmatrix} \overline{X}(s) \\ \overline{Y}(s) \end{pmatrix} + \mathcal{C}_1 Z(s)\Big\} dW(s), \quad t \in [0,T]. \tag{4.53}
$$

Since $E\eta = 0$ and $\eta \in H$, by Martingale Representation Theorem, there exists a $\zeta \in \mathcal{H}$, such that

$$
\eta = \int_0^T \zeta(s) dW(s). \tag{4.54}
$$

Then, from (4.49) and (4.53), we have

$$
\begin{aligned}
0 &= E\,\langle\, \eta, (0,I) \begin{pmatrix} \overline{X}(T) \\ \overline{Y}(T) \end{pmatrix} \rangle \\
&= \int_0^T E\,\langle\, \zeta(s), (0,I) e^{\mathcal{A}(T-s)} \Big\{\mathcal{A}_1 \begin{pmatrix} \overline{X}(s) \\ \overline{Y}(s) \end{pmatrix} + \mathcal{C}_1 Z(s)\Big\} \,\rangle\, ds.
\end{aligned}
\tag{4.55}
$$

This yields

$$
\begin{aligned}
& \int_0^T E\,\langle\, \mathcal{C}_1^T e^{\mathcal{A}^T(T-s)} \begin{pmatrix} 0 \\ I \end{pmatrix} \zeta(s), Z(s) \,\rangle\, ds \\
&\qquad = -\int_0^T E\,\langle\, \mathcal{A}_1^T e^{\mathcal{A}^T(T-s)} \begin{pmatrix} 0 \\ I \end{pmatrix} \zeta(s), \begin{pmatrix} \overline{X}(s) \\ \overline{Y}(s) \end{pmatrix} \rangle\, ds.
\end{aligned}
\tag{4.56}
$$

Now, let $0 < \delta < T$ and take

$$
Z(s) = \mathcal{C}_1^T e^{\mathcal{A}^T(T-s)} \begin{pmatrix} 0 \\ I \end{pmatrix} \zeta(s) \chi_{[T-\delta,T]}(s), \qquad s \in [0,T]. \tag{4.57}
$$

Then, $\overline{X}(s) = 0$, $\overline{Y}(s) = 0$ for all $s \in [0, T-\delta]$. Consequently, (4.56) and (4.52) result in

$$
\begin{aligned}
& \int_{T-\delta}^T E\Big|\mathcal{C}_1^T e^{\mathcal{A}^T(T-s)} \begin{pmatrix} 0 \\ I \end{pmatrix} \zeta(s)\Big|^2 ds \\
&\qquad \le C \int_{T-\delta}^T \Big(E|\zeta(s)|^2\Big)^{1/2} \Big(\int_{T-\delta}^s E|Z(r)|^2 dr\Big)^{1/2} ds \\
&\qquad \le C \int_{T-\delta}^T \Big(E|\zeta(s)|^2\Big)^{1/2} \Big(\int_{T-\delta}^s E|\zeta(r)|^2 dr\Big)^{1/2} ds.
\end{aligned}
\tag{4.58}
$$

By (4.19), we obtain

$$
\begin{aligned}
\int_{T-\delta}^{T} E|\zeta(s)|^2 ds &\le C\int_{T-\delta}^{T}\Big(E|\zeta(s)|^2\Big)^{1/2}\Big(\int_{T-\delta}^{s} E|\zeta(r)|^2 dr\Big)^{1/2} ds \\
&\le \frac{1}{2}\int_{T-\delta}^{T} E|\zeta(s)|^2 ds + C\int_{T-\delta}^{T}\int_{T-\delta}^{s} E|\zeta(r)|^2 dr ds.
\end{aligned}
\tag{4.59}
$$

Thus, it follows that

$$
\int_{T-\delta}^{T} E|\zeta(s)|^2 ds \le C\delta\int_{T-\delta}^{T} E|\zeta(s)|^2 ds,
\tag{4.60}
$$

with $C > 0$ being an absolute constant (independent of δ). Therefore, for $\delta > 0$ small, we must have

$$
\zeta(s) = 0, \qquad \text{a.e. } s \in [T-\delta, T], \text{ a.s.}
\tag{4.61}
$$

This together with (4.56) implies that

$$
\begin{aligned}
&\int_0^{T-\delta} E\,\langle\, \mathcal{C}_1^T e^{\mathcal{A}^T(T-s)}\begin{pmatrix}0\\ I\end{pmatrix}\zeta(s), Z(s)\,\rangle\, ds \\
&\quad = -\int_0^{T-\delta} E\,\langle\, \mathcal{A}_1^T e^{\mathcal{A}^T(T-s)}\begin{pmatrix}0\\ I\end{pmatrix}\zeta(s), \begin{pmatrix}\overline{X}(s)\\ \overline{Y}(s)\end{pmatrix}\rangle\, ds.
\end{aligned}
\tag{4.62}
$$

Then, thanks to (4.19), we can continue the above procedure to conclude that (4.61) holds over $[0,T]$ and hence $\eta = 0$. This proves (4.48).

We now prove (4.47). Suppose $\mathcal{K}Z = 0$. Again, we let $(\overline{X}(\cdot), \overline{Y}(\cdot))$ be defined by (4.50). Then, for any $\zeta \in \mathcal{H}$, by (4.53), we have

$$
\begin{aligned}
0 &= E\,\langle \int_0^T \zeta(s)dW(s), \mathcal{K}Z\,\rangle \\
&= E\int_0^T \langle\, \zeta(s), (0,I)e^{\mathcal{A}(T-s)}\Big\{\mathcal{A}_1\begin{pmatrix}\overline{X}(s)\\ \overline{Y}(s)\end{pmatrix} + \mathcal{C}_1 Z(s)\Big\}\,\rangle\, ds.
\end{aligned}
\tag{4.63}
$$

This implies that

$$
(0,I)e^{\mathcal{A}(T-s)}\Big\{\mathcal{A}_1\begin{pmatrix}\overline{X}(s)\\ \overline{Y}(s)\end{pmatrix} + \mathcal{C}_1 Z(s)\Big\} = 0, \quad \text{a.e. } s\in[0,T], \text{ a.s.}
\tag{4.64}
$$

By (4.19), we easily see that

$$
\mathcal{B}(s) \stackrel{\Delta}{=} \big\{(0,I)e^{\mathcal{A}(T-s)}\mathcal{C}_1\big\}^{-1}(0,I)e^{\mathcal{A}(T-s)}\mathcal{A}_1
$$

is analytic and hence bounded over $[0,T]$. From (4.64), we obtain

$$Z(s)=-\mathcal{B}(s)\begin{pmatrix}\overline{X}(s)\\ \overline{Y}(s)\end{pmatrix},\qquad \text{a.e. } s\in[0,T],\ \text{a.s.} \tag{4.65}$$

Then, $(\overline{X},\overline{Y})$ is the solution of

$$\begin{cases} d\begin{pmatrix}\overline{X}\\ \overline{Y}\end{pmatrix}=\mathcal{A}\begin{pmatrix}\overline{X}\\ \overline{Y}\end{pmatrix}dt+\big[\mathcal{A}_1-\mathcal{B}(t)\big]\begin{pmatrix}\overline{X}\\ \overline{Y}\end{pmatrix}dW(t),\\ \begin{pmatrix}\overline{X}(0)\\ \overline{Y}(0)\end{pmatrix}=0. \end{cases} \tag{4.66}$$

Hence, we must have $(\overline{X},\overline{Y})=0$, which yields $Z=0$ due to (4.65). This proves (4.47).

$\square$

An important consequence of the above is the following.

Theorem 4.7. *Let $\mathcal{C}=0$. Then, linear FBSDEs (3.12) is solvable for all $g\in H$ if and only if (4.16) and (4.19) hold. In this case, the adapted solution to (3.12) is unique (for any given $g\in H$).*

Proof. Theorems 4.2 and 4.3 tell us that (4.16) and (4.19) are necessary. We now prove the sufficiency. First of all, for any $g\in H$, by (4.16), we can find $y\in\mathbb{R}^m$, such that (4.14) holds (note $\mathcal{C}=0$). Then, we have

$$g-(0,I)\Phi(T)\begin{pmatrix}0\\ I\end{pmatrix}y\in\mathcal{N}(E). \tag{4.67}$$

Next, by (4.46), there exists a $Z\in\mathcal{H}$, such that

$$g-(0,I)\Phi(T)\begin{pmatrix}0\\ I\end{pmatrix}y=\mathcal{K}Z. \tag{4.68}$$

For this pair $(y,Z)\in\mathbb{R}^m\times\mathcal{H}$, we define (X,Y) by (4.2). Then, one can easily check that $(X,Y,Z)\in\mathcal{M}[0,T]$ is an adapted solution of (3.12). The uniqueness follows easily from (4.47) and (4.16). $\square$

The above result gives a complete solution to the solvability of linear FBSDEs (3.12) with $\mathcal{C}=0$. It is interesting that in the present case, no condition is needed for $\mathcal{A}_1$. By Theorems 2.2, 3.1 and 4.7, we obtain the following result for the original linear FBSDEs (1.1).

Corollary 4.8. *Let $C = 0$, $\widehat{C} = 0$, $F = I \in \mathbb{R}^{m\times m}$ and $\ell = m$. Then (1.1) is uniquely solvable for all b, σ, $\widehat{b}$, $\widehat{\sigma}$, g and x satisfying (2.4) if and only if the following hold:*

$$
(4.69)\qquad \begin{cases} (\widehat{C}_1 - GC_1)^{-1} \quad \text{exists}, \\ \det\left\{(0,I)e^{\begin{pmatrix} A+BG & B \\ \widehat{A}-GA+(\widehat{B}-GB)G & \widehat{B}-GB \end{pmatrix}T} \begin{pmatrix} 0 \\ I \end{pmatrix}\right\} \neq 0, \\ \det\left\{(0,I)e^{\begin{pmatrix} A+BG & B \\ \widehat{A}-GA+(\widehat{B}-GB)G & \widehat{B}-GB \end{pmatrix}t} \begin{pmatrix} C_1(\widehat{C}_1-GC_1)^{-1} \\ I \end{pmatrix}\right\} > 0, \\ \qquad\qquad\qquad\qquad \forall t \in [0,T]. \end{cases}
$$

§5. A Riccati Type Equation.

In this section, we present another method. It will give a sufficient condition for the unique solvability of (3.12). Also, it is more constructive and seems to be numerically implementable. This method is inspired by the Four-Step-Scheme proposed in [11] for general nonlinear FBSDEs with deterministic coefficients and with the diffusion coefficient of the forward SDE being nondegenerate. In the present case, we do not have the non-degeneracy of the forward diffusion. Also, the drift and diffusion are all allowed to be unbounded (since they are linear). Such a case is not covered by [11]. We will obtain a Riccati type equation and a BSDE associated with (3.12). Let us now carry out a heuristic derivation.

Suppose $(X, Y, Z) \in \mathcal{M}[0,T]$ is an adapted solution of (3.12). We assume that X and Y are related by

$$
(5.1)\qquad Y(t) = P(t)X(t) + p(t), \qquad \forall t \in [0,T], \text{ a.s.}
$$

where $P : [0,T] \to \mathbb{R}^{m\times n}$ is a deterministic matrix-valued function and $p : [0,T]\times\Omega \to \mathbb{R}^m$ is an $\{\mathcal{F}_t\}$-adapted process. We are going to derive the equations for $P(\cdot)$ and $p(\cdot)$. First of all from (5.1) and the terminal condition in (3.12), we have

$$
(5.2)\qquad g = P(T)X(T) + p(T).
$$

Let us impose

$$
(5.3)\qquad P(T) = 0, \qquad p(T) = g.
$$

Since $g \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^m)$ and $p(\cdot)$ is required to be $\{\mathcal{F}_t\}$-adapted, we should assume that $p(\cdot)$ satisfies a backward stochastic differential equation:

$$
\begin{cases} dp(t) = \alpha(t)dt + q(t)dW(t), \qquad t \in [0,T], \\ p(T) = g, \end{cases} \tag{5.4}
$$

with $\alpha(\cdot), q(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$ being undetermined. Next, by Itô's formula, we have (for simplicity, we suppress t below):

$$
\begin{aligned}
dY &= \{\dot{P}X + P[AX + BY + CZ] + \alpha\}dt \\
&\qquad\qquad + \{P[A_1X + B_1Y + C_1Z] + q\}dW \\
&= \{[\dot{P} + PA + PBP]X + PCZ + PBp + \alpha\}dt \\
&\qquad\qquad + \{[PA_1 + PB_1P]X + PC_1Z + PB_1p + q\}dW,
\end{aligned} \tag{5.5}
$$

Now, compare (5.5) with the second equation in (3.12), we obtain that

$$
[\dot{P} + PA + PBP]X + PCZ + PBp + \alpha = [\widehat{A} + \widehat{B}P]X + \widehat{C}Z + \widehat{B}p, \tag{5.6}
$$

and

$$
(PA_1 + PB_1P)X + PC_1Z + PB_1p + q = Z. \tag{5.7}
$$

By assuming $I - PC_1$ to be invertible, we have from (5.7) that

$$
Z = (I - PC_1)^{-1}\big\{(PA_1 + PB_1P)X + PB_1p + q\big\}. \tag{5.8}
$$

Then, (5.6) becomes

$$
\begin{aligned}
0 =& \big[\dot{P} + PA + PBP - \widehat{A} - \widehat{B}P + (PC - \widehat{C})(I - PC_1)^{-1}(PA_1 + PB_1P)\big]X \\
&+ \big[PB - \widehat{B} + (PC - \widehat{C})(I - PC_1)^{-1}PB_1\big]p + (PC - \widehat{C})(I - PC_1)^{-1}q + \alpha.
\end{aligned} \tag{5.9}
$$

Now, we introduce the following Riccati type differential equation for $\mathbb{R}^{m\times n}$-valued function $P(\cdot)$:

$$
\begin{cases} \dot{P} + PA + PBP - \widehat{A} - \widehat{B}P \\ \qquad\qquad + (PC - \widehat{C})(I - PC_1)^{-1}(PA_1 + PB_1P) = 0, \qquad t \in [0,T], \\ P(T) = 0, \end{cases} \tag{5.10}
$$

and the following backward SDE for $\mathbb{R}^m$-valued process $p(\cdot)$:

$$
\begin{cases} dp = -\Big\{ \big[PB - \widehat{B} + (PC - \widehat{C})(I - PC_1)^{-1}PB_1\big]p \\ \qquad\qquad + (PC - \widehat{C})(I - PC_1)^{-1}q \Big\} dt + q dW, \\ p(T) = g. \end{cases} \tag{5.11}
$$

Suppose (5.10) admits a solution $P(\cdot)$ over $[0, T]$ such that

$$
[I - P(t)C_1]^{-1} \text{ is bounded for } t \in [0, T]. \tag{5.12}
$$

Then we can define the following:

$$
\begin{cases} \widetilde{A} = A + BP + C(I - PC_1)^{-1}(PA_1 + PB_1P), \\ \widetilde{A}_1 = A_1 + B_1P + C_1(I - PC_1)^{-1}(PA_1 + PB_1P), \\ \widetilde{b} = Bp + C(I - PC_1)^{-1}(PB_1p + q), \\ \widetilde{\sigma} = B_1p + C_1(I - PC_1)^{-1}(PB_1p + q). \end{cases} \tag{5.13}
$$

It is clear that $\widetilde{A}$ and $\widetilde{A}_1$ are time-dependent matrix-valued functions and $\widetilde{b}$ and $\widetilde{\sigma}$ are $\{\mathcal{F}_t\}$-adapted processes. Under (5.12), the following SDE admits a unique strong solution:

$$
\begin{cases} dX = (\widetilde{A}X + \widetilde{b})dt + (\widetilde{A}_1X + \widetilde{\sigma})dW, \qquad t \in [0, T], \\ X(0) = x. \end{cases} \tag{5.14}
$$

The following result is comparable with the main result presented in [11] (for nonlinear FBSDEs).

Theorem 5.1. *Let (5.10) admits a solution $P(\cdot)$ such that (5.12) holds. Then, (5.11) admits a unique solution $p(\cdot) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^m)$. If (X, Y, Z) is determined by (5.14), (5.1) and (5.8), then, it is the unique adapted solution of (3.12).*

Proof. First of all, a direct computation Φ shows that the process (X, Y, Z) determined by (5.14), (5.1) and (5.8) is an adapted solution of (3.12). We now prove the uniqueness. Let $(X, Y, Z) \in \mathcal{M}[0, T]$ be any adapted solution of (3.12). Set

$$
\begin{cases} \overline{Y} = PX + p, \\ \overline{Z} = (I - PC_1)^{-1}\big[(PA_1 + PB_1P)X + PB_1p + q\big], \end{cases} \tag{5.15}
$$

where P and p are solutions of (5.10) and (5.11), respectively. Denote $\widehat{Y} = Y - \overline{Y}$ and $\widehat{Z} = Z - \overline{Z}$. Then, a direct computation shows that

$$
(5.16) \qquad \begin{cases} d\widehat{Y} = \big[(PB - \widehat{B})\widehat{Y} + (PC - \widehat{C})\widehat{Z}\big]\,dt + \big[PB_1\widehat{Y} - (I - PC_1)\widehat{Z}\big]\,dW(t), \\ \widehat{Y}(T) = 0. \end{cases}
$$

By (5.12), we may set

$$
(5.17) \qquad \widetilde{Z} = PB_1\widehat{Y} - (I - PC_1)\widehat{Z},
$$

to get the following equivalent BSDE (of (5.16)):

$$
(5.18) \qquad \begin{cases} d\widehat{Y} = \big\{[PB - \widehat{B} + (PC - \widehat{C})(I - PC_1)^{-1}PB_1]\widehat{Y} \\ \qquad\qquad - (PC - \widehat{C})(I - PC_1)^{-1}\widetilde{Z}\big\}dt + \widetilde{Z}\,dW(t), \\ \widehat{Y}(T) = 0. \end{cases}
$$

It is standard that such a BSDE admits a unique adapted solution $(\widehat{Y}, \widetilde{Z}) = 0$ (see [15]). Consequently, $\widehat{Z} = 0$. Hence, by (5.15), we obtain

$$
(5.19) \qquad \begin{cases} Y = PX + p, \\ Z = (I - PC_1)^{-1}\big[(PA_1 + PB_1P)X + PB_1p + q\big], \end{cases}
$$

This means that any adapted solution (X, Y, Z) must satisfy (5.19). Then, similar to the heuristic derivation above, we have that X has to be the solution of (5.14). Hence, we obtain the uniqueness. $\square$

The following result tells us something more.

Proposition 5.2. *Let (5.10) admits a solution $P(\cdot)$ such that (5.12) holds for $t \in [T_0, T]$ (with some $T_0 \geq 0$). Then, for any $\widetilde{T} \in [0, T - T_0]$, linear FBSDEs (3.12) is uniquely solvable on $[0, \widetilde{T}]$.*

Proof. Let

$$
(5.20) \qquad \widetilde{P}(t) = P(t + T - \widetilde{T}), \qquad t \in [0, \widetilde{T}].
$$

Then, $\widetilde{P}(\cdot)$ satisfies (5.10) with $[0, T]$ replaced by $[0, \widetilde{T}]$ and

$$
(5.21) \qquad [I - \widetilde{P}(t)C_1]^{-1} \text{ is bounded for } t \in [0, \widetilde{T}].
$$

Then, Theorem 5.1 applies. $\square$

The above Proposition 5.1 tells us that if (5.10) admits a solution $P(\cdot)$ such that (5.12) holds, (3.12) is uniquely solvable over any $[0, \widetilde{T}]$ ($\widetilde{T} \leq T$). Then, in the case $\mathcal{C} = 0$, by Theorem 4.2, the corresponding two-point boundary value problem (4.17) of ODE over $[0, \widetilde{T}]$ admits a solution for all $g \in \mathbb{R}^m$. Thus, it is necessary and sufficient that

$$\det\left\{(0, I)e^{\mathcal{A}t}\begin{pmatrix} 0 \\ I \end{pmatrix}\right\} > 0, \qquad \forall t \in [0, T]. \tag{5.22}$$

Therefore, by Theorem 4.7, compare (5.22) and (4.16), we see that the solvability of Riccati type equation (5.10) is only a sufficient condition for the solvability of (3.12).

In the rest of this section, we concentrate on the case $\mathcal{C} = 0$. In this case, (5.10) becomes

$$\begin{cases} \dot{P} + PA + PBP - \widehat{A} - \widehat{B}P = 0, & t \in [0, T], \\ P(T) = 0, \end{cases} \tag{5.23}$$

and the BSDE (5.11) is reduced to

$$\begin{cases} dp = [\widehat{B} - PB]pdt + qdW(t), & t \in [0, T], \\ p(T) = g. \end{cases} \tag{5.24}$$

We have seen that (5.22) is a necessary condition for (5.23) having a solution $P(\cdot)$ satisfying (5.12). The following result gives the inverse.

Theorem 5.3. *Let $C = 0$, $\widehat{C} = 0$. Let (5.22) hold. Then, (5.23) admits a unique solution $P(\cdot)$ which has the following representation:*

$$P(t) = -\left[(0, I)e^{\mathcal{A}(T-t)}\begin{pmatrix} 0 \\ I \end{pmatrix}\right]^{-1}(0, I)e^{\mathcal{A}(T-t)}\begin{pmatrix} I \\ 0 \end{pmatrix}, \quad t \in [0, T]. \tag{5.25}$$

Moreover, it holds

$$I - P(t)C_1 = \left[(0, I)e^{\mathcal{A}(T-t)}\begin{pmatrix} 0 \\ I \end{pmatrix}\right]^{-1}\left[(0, I)e^{\mathcal{A}(T-t)}\begin{pmatrix} C_1 \\ I \end{pmatrix}\right], \quad t \in [0, T]. \tag{5.26}$$

Consequently, if in addition to (5.22), (4.19) holds, then, (5.12) holds and the linear FBS-DEs (3.12) (with $\mathcal{C} = 0$) is uniquely solvable with the representation given by (5.14), (5.1) and (5.8).

Proof. Let us first check that (5.25) is a solution of (5.23). To this end, we denote

$$\Theta(t) = (0, I)e^{\mathcal{A}(T-t)} \begin{pmatrix} I \\ 0 \end{pmatrix}, \qquad t \in [0, T]. \tag{5.27}$$

Then, we have

$$\dot{\Theta}(t) = -(0, I)e^{\mathcal{A}(T-t)} \begin{pmatrix} I \\ 0 \end{pmatrix} B - \Theta(t)\widehat{B}. \tag{5.28}$$

Hence,

$$\begin{aligned} \dot{P} &= \Theta^{-1}\dot{\Theta}\Theta^{-1}(0, I)e^{\mathcal{A}(T-t)} \begin{pmatrix} I \\ 0 \end{pmatrix} + \Theta^{-1}(0, I)e^{\mathcal{A}(T-t)}\mathcal{A} \begin{pmatrix} I \\ 0 \end{pmatrix} \\ &= \Theta^{-1}\Big\{ -(0, I)e^{\mathcal{A}(T-t)} \begin{pmatrix} I \\ 0 \end{pmatrix} B - \Theta\widehat{B}\Big\}(-P) + \Theta^{-1}(0, I)e^{\mathcal{A}(T-t)} \begin{pmatrix} A \\ \widehat{A} \end{pmatrix} \\ &= (PB - \widehat{B})(-P) + \Theta^{-1}(0, I)e^{\mathcal{A}(T-t)} \begin{pmatrix} I \\ 0 \end{pmatrix} A + \widehat{A} \\ &= -PBP + \widehat{B}P - PA + \widehat{A}. \end{aligned} \tag{5.29}$$

Thus, $P(\cdot)$ given by (5.25) is a solution of (5.23). Uniqueness is obvious since (5.23) is a terminal value problem with the right hand side of the equation being locally Lipschitz. Finally, an easy calculation shows (5.26) holds. Then, we complete the proof. $\square$

§6. Extensions and Remarks.

In this section, we first briefly look at the case with multi-dimensional Brownian motion. Let $W(t) \equiv (W^1(t), \cdots, W^d(t))$ be a d-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{P})$ with $\{\mathcal{F}_t\}_{t\geq 0}$ being the natural filtration of $W(\cdot)$ augmented by all the $\mathcal{P}$-null sets. Similar to the case of one-dimensional Brownian motion, we may also start with the most general case, by using some necessary conditions for solvability to obtain a reduced FBSDEs. For simplicity, we skip this step and directly consider the following FBSDEs:

$$\begin{cases} dX = \Big(AX + BY\Big)dt + \displaystyle\sum_{i=1}^{d} \Big(A_1^i X + B_1^i Y + C_1^i Z^i\Big)dW^i(t), \\ & t \in [0, T], \\ dY = \Big(\widehat{A}X + \widehat{B}Y\Big)dt + \displaystyle\sum_{i=1}^{d} Z^i dW^i(t), \\ X(0) = 0, \qquad Y(T) = g, \end{cases} \tag{6.1}$$

where A, B, etc. are certain matrices of proper sizes. Note that we only consider the case that Z does not appear in the drift here since we have only completely solved such a case. We keep the notation $\mathcal{A}$ as in (3.13). In the present case, we define the space $\mathcal{M}[0,T]$ as follows (compare with (2.1)):

$$\mathcal{M}[0,T] \stackrel{\Delta}{=} L^2_{\mathcal{F}}(\Omega; C([0,T]; \mathbb{R}^n)) \times L^2_{\mathcal{F}}(\Omega; C([0,T]; \mathbb{R}^m)) \times L^2_{\mathcal{F}}(0,T; \mathbb{R}^{m\times d}), \tag{6.2}$$

with the norm being defined by (2.2), where

$$|Z|^2 = \mathrm{tr}\,\{ZZ^T\}, \qquad \forall Z \in \mathbb{R}^{m\times d}. \tag{6.3}$$

If we assume $X(\cdot)$ and $Y(\cdot)$ are related by (5.1), then, we can derive a Riccati type equation, which is exactly the same as (5.23). The associated BSDE is now replaced by the following:

$$\begin{cases} dp = [\widehat{B} - PB]pdt + \sum_{i=1}^{d} q^i dW^i(t), \qquad t \in [0,T], \\ p(T) = g. \end{cases} \tag{6.4}$$

Also, (5.13), (5.14) and (5.8) are now replaced by the following:

$$\begin{cases} \widetilde{A} = A + BP, \qquad \widetilde{b} = Bp, \\ \widetilde{A}_1^i = A_1^i + B_1^i P + C_1^i (I - PC_1^i)^{-1}(PA_1^i + PB_1^i P), \\ \widetilde{\sigma}^i = B_1^i p + C_1^i (I - PC_1^i)^{-1}(PB_1^i p + q^i), \end{cases} \qquad 1 \le i \le d, \tag{6.5}$$

$$\begin{cases} dX = (\widetilde{A}X + \widetilde{b})dt + \sum_{i=1}^{d} (\widetilde{A}_1^i X + \widetilde{\sigma}^i) dW^i(t), \qquad t \in [0,T], \\ X(0) = 0, \end{cases} \tag{6.6}$$

$$Z^i = (I - PC_1^i)^{-1}\{(PA_1^i + PB_1^i P)X + PB_1^i p + q^i\}, \quad 1 \le i \le d. \tag{6.7}$$

Our main result is the following.

Theorem 6.1. *FBSEDs (6.1) admits a unique adapted solution $(X, Y, Z) \in \mathcal{M}[0,T]$ for all $g \in H$ if and only if (4.16) holds and*

$$\det\left\{(0, I)e^{\mathcal{A}t}\mathcal{C}_1^i\right\} > 0, \qquad \forall t \in [0,T],\ 1 \le i \le d. \tag{6.8}$$

If in addition, (5.22) holds, then, (5.23) admits a unique solution $P(\cdot)$ given by (5.25) such that

$$[I - P(t)C_1^i]^{-1} \text{ is bounded for } t \in [0,T],\ 1 \le i \le d, \tag{6.9}$$

and the adapted solution $(X,Y,Z) \in \mathcal{M}[0,T]$ of (6.1) can represented through (6.6), (5.1) and (6.7).

The proof can be carried out similar to the case of one-dimensional Brownian motion.

To conclude this paper, let us now make some remarks.

We have solved the case with Z not appearing in the drift. It would be nice to extend our results to the case with $\mathcal{C}$ not zero, i.e., the process Z also appears in the drift. So far, we do not know how to treat such a case.

Finally, it seems possible to extend the above results to the time-varying coefficient case. However, the most interesting case is the random coefficient case, i.e., the case with A, B, etc. being $\{\mathcal{F}_t\}$-adapted processes.

References

[1] F. Antonelli, *Backward-forward stochastic differential equations, Ann. Appl. Probab.*, 3 (1993), 777–793.

[2] A. Bensoussan, *Stochastic maximum principle for distributed parameter system, J. Franklin Inst.*, 315 (1983), 387–406.

[3] J. M. Bismut, *An introductory approach to duality in optimal stochastic control, SIAM Rev.*, 20 (1978), 62–78.

[4] R. Buckdahn and Y. Hu, *Hedging contingent claims for a large investor in an incomplete market,* preprint.

[5] J. Cvitanic and J. Ma, *Hedging Options for a Large Investor and Forward-Backward SDEs, Ann. Appl. Probab.*, to appear.

[6] J. Douglas, Jr., J. Ma and P. Protter, *Numerical methods for forward-backward stochastic differential equations,* preprint.

[7] D. Duffie, J. Ma and J. Yong, *Black's consol rate conjecture, Ann. Appl. Probab.*, 5 (1995), 356–382.

[8] N. El Karoui, S. Peng and M. C. Quenez, *Backward stochastic differential equation, finance and optimization,* preprint.

[9] Y. Hu and S. Peng, *Solution of forward-backward stochastic differential equations,* *Probab. Theory & Relat. Fields*, 103 (1995), 273–283.

[10] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1988.

[11] J. Ma, P. Protter and J. Yong, *Solving forward-backward stochastic differential equations explicitly — a four step scheme,* *Probab. Theory & Relat. Fields*, 98 (1994), 339–359.

[12] J. Ma and J. Yong, *Solvability of forward-backward SDEs and the nodal set of Hamilton-Jacobi-Bellman equations,* *Chin. Ann. Math. Ser. B*, 16 (1995), 279–298.

[13] J. Ma and J. Yong, *Adapted solution to a class of degenerate backward stochastic PDEs, with applications,* preprint.

[14] J. Ma and J. Yong, *Linear degenerate backward stochastic partial differential equations,* in preparation.

[15] E. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation,* *Sys. & Control Lett.*, 14 (1990), 55–61.

[16] E. Pardoux and S. Tang, *The study of forward-backward stochastic differential equation and its application in quasilinear PDEs,* preprint.

[17] S. Peng, *Backward stochastic differential equation and its application in optimal control,* *Appl. Math. & Optim.*, 27 (1993), 125–144.

[18] S. Peng and Z. Wu, *Fully coupled forward-backward stochastic differential equations,* preprint.

[19] W. H. Wonham, *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, Berlin, 1980.

[20] J. Yong, *Finding adapted solutions of forward-backward stochastic differential equations — method of continuation,* preprint.