Anticipating Stochastic Differential Systems with Memory

Salah Mohammed* and Tusheng Zhang

February 16, 2009

Abstract

This article establishes existence and uniqueness of solutions to two classes of stochastic systems with finite memory subject to anticipating initial conditions which are sufficiently smooth in the Malliavin sense. The two classes are semilinear stochastic functional differential equations (sfdes) and fully nonlinear sfdes with a sublinear drift term. For the semilinear case, we use Malliavin calculus techniques, existence of the stochastic semiflow and an infinite-dimensional substitution theorem. For the fully nonlinear case, we employ an anticipating version of the Itô-Ventzell formula due to Ocone and Pardoux ([O-P]). In both cases, the use of Malliavin calculus techniques is necessitated by the infinite-dimensionality of the initial condition.

1 Introduction

The purpose of this article is to study the existence and uniqueness of solutions to nonlinear Stratonovich stochastic differential systems with smooth memory (stochastic functional differential equations-sfdes) subject to anticipating initial conditions. Such sfdes take the form:

$$\left.\begin{aligned} dx(t) &= H\big(x(t-r), x(t), x_t\big)\,dt + G\big(x(t), g(x_t)\big)\circ dW(t), \quad t\geq 0\\ (x(0), x_0) &= Y \in L^0(\Omega, M_2). \end{aligned}\right\} \tag{I}$$

In the above sfde, $x_t \in L^2\big([-r,0], \mathbf{R}^d\big)$ stands for the segment $x_t(s) := x(t+s)$ for $s\in[-r,0]$ and $t>0$. The state space is the Delfour-Mitter Hilbert space

*The research of this author is supported in part by NSF Grants DMS-9975462, DMS-0203368, DMS-0705970, Alexander von Humboldt Foundation (Germany) and Institut Mittag-Leffler (Royal Swedish Academy of Sciences).

AMS 2000 subject classifications. Primary 60H07, 60H10, 60H20; secondary 60H25.

Key words and phrases. Malliavin calculus, stochastic semiflow, perfect cocycle, anticipating initial condition, substitution theorem, stochastic differential systems with memory, stochastic functional differential equation (sfde).

$M_2 := \mathbf{R}^d \times L^2\big([-r,0],\mathbf{R}^d\big)$ endowed with the Hilbert norm

$$\|(v,\eta)\|_{M_2}^2 := |v|^2 + \int_{-r}^0 |\eta(s)|^2\,ds,\ \ (v,\eta)\in M_2,$$

and the coefficients $H : \mathbf{R}^d \times M_2 \to \mathbf{R}^d, G : \mathbf{R}^d \times \mathbf{R}^p \to \mathbf{R}^{d\times m}$ satisfy appropriate regularity conditions that will be spelled out in the forthcoming sections. The memory in the diffusion coefficient in (I) is "smooth" in the sense that it is described via a quasitame functional $g : L^2\big([-r,0],\mathbf{R}^d\big) \to \mathbf{R}^p$ of the form $g(\eta) := \int_{-r}^0 \sigma\big(\eta(s)\big)\rho(s)\,ds$, again with the mappings $\sigma : \mathbf{R}^d \to \mathbf{R}^p$, $\rho : [-r,0] \to \mathbf{R}$, satisfying suitable smoothness and growth conditions. The sfde (I) is driven by m-dimensional Brownian motion W on a probability space $(\Omega,\mathcal{F},P)$. The initial condition $Y : \Omega \to M_2$ is random, possibly anticipates the Brownian motion W, and satisfies suitable integrability and Malliavin smoothness requirements ([Ma], [Nu.1], [Nu.2]).

For any metric space K, denote by $\mathcal{B}(K)$ its Borel σ-algebra. A *solution* of the initial-value problem (I) is a $(\mathcal{B}([-r,\infty))\otimes\mathcal{F},\mathcal{B}(\mathbf{R}^d))$-measurable and sample-continuous process $x : [-r,\infty)\times\Omega \to \mathbf{R}^d$ such that

$$\left.\begin{aligned} x(t) &= x(0) + \int_0^t H\big(x(u-r),x(u),x_u\big)\,du + \int_0^t G\big(x(u),g(x_u)\big)\circ dW(u), \quad t\geq 0,\\ (x(0),x_0) &= Y \in L^0(\Omega,M_2)\end{aligned}\right\}$$

a.s.. In the above equation, the stochastic integral is a Stratonovich one.

In the special case when the initial condition $Y \equiv (v,\eta)\in M_2$ is deterministic, the existence of a unique solution and indeed of stochastic semiflows for the sfde (I) is known through work by one of the authors with M. Scheutzow ([Mo.1], [M-S.1]). In this context, one should note that smoothness of the memory appears to be a requirement for the existence of the semiflow for (I); i.e. (I) is *regular* as defined in [M-S.1]. Cf. also [Mo.3] and [Mo.2].

One of the motivations for studying the sfde (I) with anticipating initial data is the existence of (local) random stable and unstable manifolds $(\mathcal{S}(\omega),\mathcal{U}(\omega))$ near stationary random points for the underlying semiflow $X : \mathbf{R}^+\times\Omega\times M_2 \to M_2$ where $X(t,\cdot,(v,\eta)) = (x(t),x_t)$, $t\geq 0$, and $x : [-r,\infty)\times\Omega\to\mathbf{R}^d$ is the unique solution of (I) with a deterministic initial condition $(v,\eta)\in M_2$. By construction (through the Ruelle-Oseledec multiplicative ergodic theorem), the local invariant submanifolds $(\mathcal{S}(\omega),\mathcal{U}(\omega))$ *anticipate* the driving Brownian motion W. So in order to study the regularity (e.g. in distribution) of the finite-dimensional manifolds $\mathcal{U}(\omega)$, one has to understand the behavior of the semiflow starting from a random point $Y(\omega)\in\mathcal{U}(\omega)$, as well as the evolution of the whole submanifold $\mathcal{U}(\omega)$ under the semiflow X. It is expected that an anticipating sfde such as (I) will be a necessary starting point for such a study.

There are a number of studies for stochastic differential equations and stochastic partial differential equations with anticipating initial conditions. See the work by Nualart and Pardoux [N-P], Ocone and Pardoux [O-P], Arnold and Imkeller [A-I], Tindel [T], Mohammed and Zhang [M-Z]. See also [G-Nu-S].

In the sfde (I), it is possible to employ a simplifying transformation that will allow us to eliminate the smooth memory from the diffusion coefficient. This works as follows:

Let $x : [-r,\infty) \times \Omega \to \mathbf{R}^d$ be a solution of the sfde (I) starting at $Y \in L^0(\Omega, M_2)$. Define $z : [-r,\infty) \times \Omega \to \mathbf{R}^p$ by

$$z(t) = \begin{cases} g(x_t), & t > 0, \\ g(\eta), & -r \leq t \leq 0. \end{cases}$$

Then the pair (x, z) is a process satisfying the coupled sfdes

$$\left.\begin{aligned} dx(t) &= H\big(x(t-r), x(t), x_t\big)\,dt + G\big(x(t), z(t)\big) \circ dW(t), \quad t > 0, \\ dz(t) &= H_0(x(t-r), x(t), x_t)\,dt, \quad t > 0, \\ \big(x(0), x_0\big) &= Y,\ z_0(s) = g(\eta), \quad \text{for all } s \in [-r, 0], \end{aligned}\right\} \qquad (\mathrm{I}')$$

where $H_0 : \mathbf{R}^d \times M_2 \to \mathbf{R}^p$ is given by

$$H_0(v_1, v, \eta) := \sigma\big(v\big)\rho(0) - \sigma\big(v_1\big)\rho(-r) - \int_{-r}^{0} \sigma\big(\eta(s)\big)\rho'(s)\,ds,\ (v,\eta) \in M_2,\ v_1 \in \mathbf{R}^d.$$

It will be apparent later that, under sufficient regularity hypotheses on the coefficients, the augmented drift and diffusion coefficients $(H, H_0), (G, 0)$ in (I$'$) will satisfy the same regularity and growth conditions as those of the original sfde (I). The main point of the above transformation is that it removes the (quasitame) memory from the diffusion coefficient in (I) and places it in the drift coefficient (H, H_0) of the coupled system (I$'$). As a result the sfde (I$'$) has *no memory* in the diffusion coefficient, while every solution pair (x, z) of (I$'$) yields a solution x of (I).

The rest of the paper is split into two sections. Section 2 looks at the existence of solutions to the sfde (I) in the semilinear case when the diffusion coefficient G is linear and the drift coefficient H has continuous and globally bounded derivatives. In this case, we employ known results on the existence of the semiflow ([M-S.1], [Mo.3]) and develop an infinite-dimensional substitution theorem in order to obtain a solution to the anticipating semilinear sfde. Because of the infinite dimensional nature of the state space M_2 in (I), it should be noted that finite-dimensional substitution theorems do not apply in this context (cf. [M-Z]). In Section 3, we study the fully nonlinear case using a nonlinear variational technique coupled with an application of the anticipating Itô-Ventzell formula due to Ocone and Pardoux ([O-P]). Here, the variational technique reduces the problem to a coupled system of random equations but requires imposing *sublinear* growth hypotheses on the drift coefficient H. Note that such a sublinear growth hypothesis is not needed in the semilinear case treated in Section 2. It is not clear to us whether the sublinear growth condition on the drift H can be relaxed within the fully nonlinear setting of Section 3.

2 Anticipating semilinear sfdes

In this section, we consider the sfde (I) of Section 1 in the special (semilinear) case when $G : \mathbf{R}^d \times \mathbf{R}^p \to \mathbf{R}^{d \times m}$ is a linear map, $H : \mathbf{R}^d \times M_2 \to \mathbf{R}^d$ is C_b^2 (i.e. with continuous globally bounded Fréchet derivatives), and the quasitame memory function

$g : L^2([-r,0],\mathbf{R}^d) \to \mathbf{R}^p$ is of the form $g(\eta) := \int_{-r}^0 \sigma(\eta(s))\rho(s)\,ds$, $\eta \in L^2([-r,0],\mathbf{R}^d)$ with $\sigma \in C_b^1$ and ρ of class C^1.

Using the memory reduction technique of Section 1, it is easy to see that the augmented drift coefficient (H, H_0) in (I′) is C_b^1 and the diffusion coefficient $(G, 0)$ is linear. Thus, without loss of generality, we need only look at the following Stratonovich semilinear sfde:

$$\left.\begin{aligned} dx(t) &= H\big(x(t-r), x(t), x_t\big)\,dt + G\big(x(t)\big) \circ dW(t), \quad t \geq 0,\\ \big(x(0), x_0\big) &= Y \in L^0(\Omega, M_2; \mathcal{F}), \end{aligned}\right\} \tag{II}$$

with no memory in the diffusion term. In (II) we assume that the drift coefficient $H : \mathbf{R}^d \times M_2 \to \mathbf{R}^d$ is C_b^2 and the diffusion coefficient is a *linear* map $G : \mathbf{R}^d \to L(\mathbf{R}^m, \mathbf{R}^d)$. The driving noise is m-dimensional Brownian motion $W : \mathbf{R}^+ \times \Omega \to \mathbf{R}^m$ on the complete filtered Wiener space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$.

Denote by $\psi : \mathbf{R} \times \mathbf{R}^d \times \Omega \to \mathbf{R}^d$ the linear stochastic flow of the Stratonovich stochastic ordinary differential equation (sode)

$$\left.\begin{aligned} d\psi(t) &= G\big(\psi(t)\big) \circ dW(t), \quad t > 0,\\ \psi(0) &= x \in \mathbf{R}^d, \end{aligned}\right\} \tag{III}$$

([Ku]). Then, for a.a. $\omega \in \Omega$ and all $t \in \mathbf{R}^+$, $\psi(t,\cdot,\omega) \in GL(\mathbf{R}^d, \mathbf{R}^d)$, the general linear group on $\mathbf{R}^d$. Let $\theta : \mathbf{R} \times \Omega \to \Omega$ denote the Brownian shift

$$\theta(t,\omega)(s) := \omega(t+s) - \omega(t),\ t, s \in \mathbf{R},\ \omega \in \Omega,$$

on the Wiener space $(\Omega, \mathcal{F}, P)$. It is well-known that the stochastic flow of the sode (III) gives a perfect cocycle (ψ, θ) on $\mathbf{R}^d$; viz.

$$\psi(t,\cdot,\omega) = \psi(t-u,\cdot,\theta(u,\omega)) \circ \psi(u,\cdot,\omega), \quad u \leq t \in \mathbf{R}^+, \omega \in \Omega.$$

We will use the following notation thoughout this article. If $p \geq 2$ is an integer, we denote by $\mathbb{D}^{1,p}(\Omega, M_2)$ the Sobolev space of all $(\mathcal{F}, \mathcal{B}(M_2))$-measurable random variables $Y : \Omega \to M_2$ which are p-integrable together with their Malliavin derivatives $\mathcal{D}Y$ ([Nu.1], [Nu.2]).

In order to investigate the existence of a solution to the anticipating semilinear sfde (II), we will adopt the following strategy:

- In the sfde (II), replace the random initial condition Y by a fixed $(v, \eta) \in M_2$; and observe that the resulting sfde

$$\left.\begin{aligned} dx(t) &= H\big(x(t-r), x(t), x_t\big)\,dt + G\big(x(t)\big) \circ dW(t), \quad t \geq 0,\\ \big(x(0), x_0\big) &= (v, \eta) \in M_2 \end{aligned}\right\} \tag{IV}$$

  is equivalent to the random variational integral equation

$$\left.\begin{aligned} x\big(t,\omega\big) &= \psi(t,v,\omega) + \int_0^t \psi\big(t-u,\cdot,\theta(u,\omega)\big) H\big(x(u-r,\omega), x(u,\omega), x_u(\omega)\big)\,du\\ x_0 &= \eta \in L^2([-r,0],\mathbf{R}^d). \end{aligned}\right\} \tag{2.1}$$

- Solutions of the above random variational equation generate a perfect cocycle $X : \mathbf{R}^+ \times M_2 \times \Omega \to M_2$ for the sfde (IV) satisfying

$$X(t, (v, \eta), \cdot) := \big({}^{(v,\eta)}x(t, \cdot), {}^{(v,\eta)}x_t(\cdot, \cdot)\big)$$

for all $t \in \mathbf{R}^+$, $(v, \eta) \in M_2$, a.s. The process ${}^{(v,\eta)}x(t, \cdot)$ denotes the unique solution of (IV) starting at $(v, \eta) \in M_2$. The existence of the cocycle X and its path regularity in $(t, (v, \eta))$ is established in [M-S.1]; cf. also [Mo.3] for the linear case.

- Using the cocycle X, rewrite the sfde (IV) in the form

$$\left.\begin{aligned} pr_1 X\big(t, (v, \eta), \cdot\big) = v &+ \int_0^t H\big(pr_2 X\big(u, (v, \eta), \cdot\big)(-r), X\big(u, (v, \eta), \cdot\big)\big)\, du \\ &+ \int_0^t G\big(pr_1 X\big(u, (v, \eta), \cdot\big)\big) \circ dW(u), \quad t \geq 0 \\ X\big(0, (v, \eta), \cdot\big) = (v, \eta) &\in M_2, \end{aligned}\right\} \qquad (\mathrm{IV}')$$

where $pr_1 : M_2 \to \mathbf{R}^d$, $pr_2 : M_2 \to L^2([-r, 0], \mathbf{R}^d)$ are the projections onto the first and second coordinates, respectively. In the above equation, the initial condition $(v, \eta) \in M_2$ is viewed as an *infinite-dimensional parameter*.

- Our main objective is to replace the deterministic initial condition (v, η) in (IV′) by a sufficiently Malliavin smooth $\mathcal{F}$-measurable random variable $Y : \Omega \to M_2$; that is $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$. This is the essence of the substitution theorem (Theorem 2.7), which gives existence of a solution to the anticipating semilinear sfde (II).

- To prove the infinite-dimensional substitution result (Theorem 2.7), we first approximate Y by a sequence of finite-dimensional random variables $Y_n : \Omega \to M_2^n$ where M_2^n is the linear subspace generated by $\{e_i : 1 \leq i \leq n\}$, with $\{e_i : i \geq 1\}$ a complete orthonormal system in M_2.

- We perform the substitution $(v, \eta) = Y_n$ for each $n \geq 1$ in (IV′), using a finite-dimensional substitution theorem in ([Nu.2]). Then we pass to the limit as $n \to \infty$ in the resulting equation, using appropriate spatial moment estimates on the cocycle X, its Fréchet and Malliavin derivatives, together with the regularity requirement that $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$. The required spatial estimates on X are developed in Lemmas 2.2 and 2.3, via the random variational equation (2.1).

We now proceed with the details of the analysis. To begin with, we borrow the following Gronwall-type lemma from [M-Z]:

Lemma 2.1. Fix $a \in (0, \infty)$. Let $f^0, g^0 : [0, a] \times \Omega \to \mathbf{R}^+$ be non-negative $(\mathcal{B}([0, a]) \otimes \mathcal{F}, \mathcal{B}(\mathbf{R}^+))$-measurable processes and $h^0 : [0, a] \times [0, a] \times \Omega \to \mathbf{R}^+$ a $(\mathcal{B}([0, a] \times [0, a]) \otimes \mathcal{F}, \mathcal{B}(\mathbf{R}^+))$-measurable random field satisfying the following hypotheses:

(i) For a.a. $\omega \in \Omega$ and all $u \in [0, a]$, the paths $f^0(\cdot, \omega), g^0(\cdot, \omega), h^0(\cdot, u, \omega)$ are continuous on $[0, a]$.

(ii) The process f^0 is $(\mathcal{F}_t)_{t\in[0,a]}$-adapted; and whenever $0 < u < t \le a$, the random variables $h^0(t-u,u,\cdot)$ are measurable with respect to the σ-algebra generated by the Brownian increments $W(u_2) - W(u_1), u \le u_1 \le u_2 \le t$.

(iii)$E\left[\sup_{0\le t\le a} g^0(t,\cdot)\right] + \sup_{0\le u\le a} E\left[\sup_{0\le t\le a} h^0(t,u,\cdot)\right] < \infty$, and

$$f^0(t,\cdot) \le g^0(t,\cdot) + \int_0^t h^0(t-u,u,\cdot)[1+f^0(u,\cdot)]\,du$$

a.s. for all $t \in [0,a]$.

Then $\sup_{0\le t\le a} f^0(t,\cdot)$ is integrable and there exist positive constants C_1, C_2 such that

$$E\left[\sup_{0\le u\le t} f^0(u,\cdot)\right] \le C_1 e^{C_2 t}$$

for all $t \in [0,a]$.

Proof. For a proof of the above lemma, see [M-Z]. □

Next, we give some spatial moment estimates on the cocycle of the semilinear sfde (IV). These will be developed in the following two lemmas.

In what follows we will denote Fréchet and Malliavin differentiation by D and $\mathcal{D}$, respectively.

Lemma 2.2. In the sfde (IV) assume that G is linear and H is C_b^1. Then the trajectories of (IV) generate a C_b^1 cocycle $X : \mathbf{R}^+ \times M_2 \times \Omega \to M_2$. Furthermore, there are positive deterministic constants C and C_R such that

$$E\|X(t,(v,\eta),\cdot) - X(t,(v_1,\eta_1),\cdot)\|_{M_2}^{2p} \le C\|(v,\eta) - (v_1,\eta_1)\|_{M_2}^{2p} \tag{2.2}$$

for all $(v,\eta),(v_1,\eta_1) \in M_2$; and

$$\left.\begin{aligned} E\|X(t,(v,\eta),\cdot) - X(t,(v_1,\eta_1),\cdot) - X(t',(v,\eta),\cdot) + X(t',(v_1,\eta_1),\cdot)\|^{2p} \\ \le C_R\|(v,\eta)-(v_1,\eta_1)'\|^{2p}\cdot|t-t'|^p \end{aligned}\right\} \tag{2.3}$$

for $p \ge 1, (v,\eta),(v_1,\eta_1) \in M_2$ with $\|(v,\eta)\|_{M_2}, \|(v_1,\eta_1)\|_{M_2} \le R,\ t,t' \in [0,a]$, and R any positive real.

Proof. Under the given hypotheses on the coefficients H and G, the existence of the C_b^1 cocycle X follows from the random variational equation (2.1) by the arguments in [M-S.1].

To prove the estimate (2.2), we first rewrite the Stratonovich sfde (IV) in Itô form. Let $\{f_k\}_{k=1}^m$ be a basis of $\mathbf{R}^m$. Express the Brownian motion W in the form

$$W(t) = \sum_{k=1}^m W^k(t) f_k, \quad t > 0,$$

where $\{W^k\}_{k=1}^m$ are independent standard one-dimensional Brownian motions. Define $G_k \in L(\mathbf{R}^d)$, $1 \leq k \leq m$, by $G_k(x) := G(x)(f_k)$, $x \in \mathbf{R}^d$, $1 \leq k \leq m$. Then (IV) is equivalent to the Itô sfde

$$\left.\begin{aligned} &dx(t) = H\big(x(t-r), x(t), x_t\big)dt + \frac{1}{2}\sum_{k=1}^m G_k^2\big(x(t)\big)dt + G\big(x(t)\big)\,dW(t),\ t>0,\\ &\big(x(0), x_0\big) = (v,\eta) \in M_2\,. \end{aligned}\right\} \tag{V}$$

Now replace (v,η) in (V) by $(v_1,\eta_1) \in M_2$, subtract the resulting integral equation from (V), take moments of $2p$-th order, use the martingale inequality, the global Lipschitz property of H and Gronwall's lemma, to obtain the estimate (2.2). Details are left to the reader.

To prove (2.3), let $0 < t' < t$, $x, x' \in \mathbf{R}^d$ be such that $|x|, |x'| \leq R$, for R a positive real. Then using the martingale inequality and the following Itô version of (III):

$$\left.\begin{aligned} &d\psi(t,x) = \frac{1}{2}\sum_{k=1}^m G_k^2\big(\psi(t,x)\big)dt + G\big(\psi(t,x)\big)\,dW(t),\ t>0,\\ &\psi(0,x) = x \in \mathbf{R}^d, \end{aligned}\right\} \tag{VI}$$

we obtain

$$E|\psi(t,x,\cdot) - \psi(t,x',\cdot) - \psi(s,x,\cdot) + \psi(s,x',\cdot)|^{2p} \leq C_R|x-x'|^{2p}\cdot|s-t|^p \tag{2.4}$$

for any $p \geq 1$. Applying a similar argument to the Itô sfde (V) gives the estimate (2.3). The reader may fill in the details. □

Lemma 2.3. Assume G is linear and H is C_b^2. Let $X : \mathbf{R}^+ \times M_2 \times \Omega \to M_2$ be the cocycle generated by the semilinear sfde (IV). Then the following estimates hold

$$E \sup_{\substack{(v,\eta)\in M_2\\ 0\leq t\leq a}} \frac{\|X\big(t,(v,\eta),\cdot\big)\|_{M_2}^{2p}}{1+\|(v,\eta)\|_{M_2}^{2p}} < \infty, \tag{2.5}$$

$$E \sup_{\substack{(v,\eta)\in M_2\\ 0\leq t\leq a}} \|DX\big(t,(v,\eta),\cdot\big)\|_{L(M_2)}^{2p} < \infty; \tag{2.6}$$

$$\sup_{0\leq u\leq a} E \sup_{\substack{(v,\eta)\in M_2\\ 0\leq t\leq a}} \frac{\|\mathcal{D}_u X\big(t,(v,\eta),\cdot\big)\|_{M_2}^{2p}}{1+\|(v,\eta)\|^{2p}} < \infty, \tag{2.7}$$

for all integers $p \geq 1$.

Proof. We first show the estimate

$$\sup_{0\leq u,v\leq a} E \sup_{u\leq t\leq a} \|\mathcal{D}_v\psi\big(t-u,\cdot,\theta(u,\cdot)\big)\|_{\mathbf{R}^{d\times d}}^{2p} < \infty. \tag{2.8}$$

To prove (2.8), note that, for fixed $u \in [0,a]$, the $\mathbf{R}^{d\times d}$-valued process

$$[u,a] \ni t \mapsto \psi(t-u, \theta(u,\cdot)) \in \mathbf{R}^{d\times d}$$

satisfies the Itô integral equation

$$\left.\begin{aligned} \psi(t-u,\cdot,\theta(u,\cdot)) = id_{\mathbf{R}^d} + \frac{1}{2}\sum_{k=1}^m \int_0^{t-u} G_k^2 \circ \psi(s,\cdot,\theta(u,\cdot))\, ds \\ + \int_0^{t-u} G \circ \psi(s,\cdot,\theta(u,\cdot))\, dW(s,\theta(u,\cdot)) \end{aligned}\right\} \tag{2.9}$$

for $t \geq u$.

Fix $v \in [0,a]$ such that $v + u \leq t$, take Malliavin derivative $(\mathcal{D}_v)$ on both sides of (2.9), use the helix property of W, and change the time variable in both integrals on the right-hand side of the resulting equation. This yields the following Itô stochastic integral equation for the process $[u,a] \ni t \mapsto \mathcal{D}_v\psi(t-u,\theta(u,\cdot)) \in \mathbf{R}^{d\times d}$:

$$\left.\begin{aligned} \mathcal{D}_v\psi(t-u,\cdot,\theta(u,\cdot)) = G\circ\psi(v,\cdot,\theta(u,\cdot)) + \frac{1}{2}\sum_{k=1}^m \int_u^t G_k^2 \circ \mathcal{D}_v\psi(s-u,\cdot,\theta(u,\cdot))\, ds \\ + \int_u^t G\circ \mathcal{D}_v\psi(s-u,\cdot,\theta(u,\cdot))\, dW(s,\cdot), \quad t \in [u,a]. \end{aligned}\right\} \tag{2.10}$$

Now apply the operation $\sup_{0\leq u,v\leq a} E \sup_{u\leq t\leq a} \|\cdot\|^{2p}_{\mathbf{R}^{d\times d}}$ to both sides of the above equation, use the martingale inequality on the Itô integral and Gronwall's lemma to get the estimate (2.8).

Next, we prove the estimate (2.5). For simplicity of computation, assume $0 < t < r$. (The case $r \leq t \leq a$ may be treated similarly). Recall that $pr_1 : M_2 \to \mathbf{R}^d$ and $pr_2 : M_2 \to L^2\big([-r,0],\mathbf{R}^d\big)$ are the projections of M_2 onto its first and second factors, respectively. It is easy to see that the following equality holds:

$$\|pr_2(X(t,(v,\eta),\omega)\|^2_{L^2} = \int_{-r}^{-t} |\eta(t+s)|^2\, ds + \int_{-t}^0 |pr_1(X(t+s,(v,\eta),\omega))|^2\, ds, \ 0 \leq t \leq r. \tag{2.11}$$

Using the cocycle X, we can rewrite the random variational equation (2.1) in the form

$$\left.\begin{aligned} pr_1X\big(t,(v,\eta),\omega\big) = \psi(t,v,\omega) + \int_0^t \psi\big(t-u,\cdot,\theta(u,\omega)\big)H\big(\eta(u-r), X(u,(v,\eta),\omega)\big)\, du, \\ X(0,(v,\eta),\omega) = (v,\eta) \in M_2. \end{aligned}\right\} \tag{2.12}$$

Using (2.12), (2.11) and the linear growth property of H, we obtain the estimate

$$\begin{aligned} \|X(t,(v,\eta),\cdot)\|^{2p} \leq \|\psi(t,\cdot)\|^{2p}_{\mathbf{R}^{d\times d}}|v|^{2p} \\ + C\int_0^t \|\psi(t-u,\theta(u,\cdot))\|^{2p}_{\mathbf{R}^{d\times d}}\big(1 + \|\eta\|^{2p} + \|X(u,(v,\eta),\cdot)\|^{2p}\big)\, du \end{aligned} \tag{2.13}$$

a.s. for $0 \le t \le a$, $(v, \eta) \in M_2$, where C is a deterministic positive constant depending only on a and p. Now, if we divide both sides of the inequality (2.13) by $(1 + \|(v, \eta)\|_{M_2}^{2p})$ and take supremum over all $(v, \eta) \in M_2$, we then get

$$\sup_{(v,\eta)\in M_2} \frac{\|X(t,(v,\eta),\cdot)\|^{2p}}{(1+\|(v,\eta)\|_{M_2}^{2p})} \le \|\psi(t,\cdot)\|_{\mathbf{R}^{d\times d}}^{2p} + C\int_0^t \|\psi(t-u,\theta(u,\cdot))\|_{\mathbf{R}^{d\times d}}^{2p}\,du$$
$$+ C\int_0^t \|\psi(t-u,\theta(u,\cdot))\|_{\mathbf{R}^{d\times d}}^{2p} \cdot \sup_{(v,\eta)\in M_2} \frac{\|X(u,(v,\eta),\cdot)\|^{2p}}{(1+\|(v,\eta)\|_{M_2}^{2p})}\,du \tag{2.14}$$

a.s. for $0 \le t \le a$. Denote

$$f^0(t,\cdot) := \sup_{(v,\eta)\in M_2} \frac{\|X(t,(v,\eta),\cdot)\|^{2p}}{(1+\|(v,\eta)\|_{M_2}^{2p})},\ g^0(t,\cdot) := \|\psi(t,\cdot)\|_{\mathbf{R}^{d\times d}}^{2p},\ h^0(t,u,\cdot) := \|\psi(t,\theta(u,\cdot))\|_{\mathbf{R}^{d\times d}}^{2p},$$

a.s. for $0 \le u \le t \le a$. Hence (2.14) may be written in the form

$$f^0(t,\cdot) \le g^0(t,\cdot) + \int_0^t h^0(t-u,u,\cdot)[1+f^0(u,\cdot)]\,du$$

a.s. for all $t \in [0,a]$. We will now apply Lemma 2.1 to the above inequality. To do this we will check that the processes f^0, g^0, h^0 fulfill all the conditions of above lemma. Observe first that the processes f^0, g^0, h^0 are finite a.s.: This follows from (2.14), a simple truncation argument and Gronwall's lemma. Secondly, an elementary argument using (2.12) and the linear growth property of H shows that the process f^0 is sample-continuous. Also it is easy to see that g^0 and $h^0(\cdot, u)$ are sample continuous for each $u \in [0,a]$. Thirdly, since $X(\cdot,(v,\eta),\cdot)$ is $(\mathcal{F}_t)_{t\in[0,a]}$-adapted, then so is f^0. Fourthly, note that for fixed $u, t, u < t$, the random variable $\psi(t-u, \theta(u,\cdot))$ is measurable with respect to the σ-algebra generated by the Brownian increments $W(u_2) - W(u_1), u \le u_1 \le u_2 \le t$; hence so is $h^0(t,u,\cdot)$. Finally, it is easy to see that hypothesis (iii) of Lemma 2.1 is satisfied because of (2.9), Gronwall's lemma and (2.14). Hence all the conditions of Lemma 2.1 are fulfilled; thus there exist positive constants C_1, C_2 (possibly dependent on a) such that

$$E \sup_{\substack{0\le t'\le t\\(v,\eta)\in M_2}} \frac{\|X(t',(v,\eta),\cdot)\|^{2p}}{(1+\|(v,\eta)\|_{M_2}^{2p})} \le C_1 e^{C_2 t}$$

for all $t \in [0,a]$. This implies (2.5).

To prove the estimate (2.6), we linearize the random variational equation (2.12) by taking Fréchet derivatives with respect to $(v,\eta) \in M_2$ on both sides of the equation, and then using a similar proof to that of (2.5). Note that the integral on the right hand side of (2.12) is C_b^1 in (v,η) because H is C_b^2. Details are left to the reader.

The last assertion (2.7) of the lemma is proved by similar arguments to the above. □

The next step in our strategy is to establish a substitution result whereby the initial condition (v,η) in the sfde (IV′) can be replaced by an $\mathcal{F}$-measurable random variable $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$. We first replace (v,η) by a finite-dimensional random variable $Y_n : \Omega \to M_2^n$ where M_2^n is the linear subspace generated by $\{e_i : 1 \le i \le n\}$, and $\{e_i : i \ge 1\}$ a complete orthonormal system in M_2. In order to perform the substitution, we will use two well-known results, Theorems 2.4, 2.5 below.

First, we recall some definitions. Let $\mathbb{D}^{k,p}$ denote the Malliavin Sobolev space associated with one-dimensional Brownian motion W_1, where k is the order of differentiation and p is the order of integrability. Denote $\mathbb{L}^{k,p} := L^p_{loc}(\mathbf{R}^+, \mathbb{D}^{k,p}; dt)$

Following Nualart ([Nu.2], Chapter 5), denote by $\mathbb{L}^{k,p}_{q^+}$ the set of all (real-valued) processes $f \in \mathbb{L}^{k,p}$ such that

$$\lim_{n\to\infty} \int_0^1 \sup_{s\le t\le (s+\frac{1}{n})\wedge 1} E\big(|\mathcal{D}_s f(t) - (\mathcal{D}^+ f)_s|^q\big) ds = 0\,, \tag{2.15}$$

where $(\mathcal{D}^+ f)_s = \lim_{\varepsilon\to 0} D_s f(s+\varepsilon)$; and denote by $\mathbb{L}^{k,p}_{q^-}$ the set of all $f \in \mathbb{L}^{k,p}$ such that

$$\lim_{n\to\infty} \int_0^1 \sup_{(s-\frac{1}{n})\vee 0\le t< s} E\big(|\mathcal{D}_s f(t) - (\mathcal{D}^- f)_s|^q\big) ds = 0, \tag{2.16}$$

where $(\mathcal{D}^- f)_s = \lim_{\varepsilon\to 0} D_s f(s-\varepsilon)$. Define $\mathbb{L}^{k,p}_q := \mathbb{L}^{k,p}_{q^+} \cap \mathbb{L}^{k,p}_{q^-}$. The proofs of the next two results are given in [Nu.2] (Theorems 5.2.3, 5.3.4).

Theorem 2.4. *Let $u \in \mathbb{L}^{1,2}_{1,loc}$. Then u is Stratonovich integrable and*

$$\int_0^1 u(t)\circ dW_1(t) = \int_0^1 u(t)\, dW_1(t) + \frac{1}{2}\int_0^1 (\nabla u)(t)dt$$

where the stochastic integral on the right hand side is a Skorohod integral, and

$$(\nabla u)(t) := (\mathcal{D}^+ u)(t) + (\mathcal{D}^- u)(t), \quad t > 0.$$

Theorem 2.5. *Let $u(t,x)$, $0 \le t \le a$, $x \in \mathbf{R}^m$, satisfy:*

(i) *For each $x \in \mathbf{R}^m$, $u(t,x)$ is $\mathcal{F}_t$-measurable for $0 \le t \le a$;*

(ii) *There exists $p \ge 2$ and $\alpha > m$ such that*

$$E|u(t,x) - u(t,x')|^p \le C_R |x - x'|^\alpha,$$

and

$$E\big(|u(t,x) - u(t,x') - u(s,x) + u(s,x')|^p\big) \le C_R |x-x'|^\alpha |s-t|^{p/2}$$

whenever $|x|$, $|x'| \le R$, $s,t \in [0,a]$, for any positive real R, and where C_R is a positive constant.

(iii) $\int_0^a E|u(t,0)|^2 dt < \infty$, $[0,a] \ni t \mapsto u(t,x) \in L^2(\Omega, \mathbf{R})$ *is continuous for each* $x \in \mathbf{R}^m$.

Then for every $F \in L^0(\Omega, \mathbf{R}^m; \mathcal{F})$, $u(t,F)$ is Stratonovich integrable and

$$\int_0^a u(t,F)\circ dW_1(t) = \int_0^a u(t,x)\circ dW_1(t)\bigg|_{x=F}\,. \tag{2.17}$$

As a consequence of Theorem 2.5, the next result allows us to substitute finite-dimensional projections of Y into (IV′).

Theorem 2.6. *For each $n \geq 1$, let $Y_n := p_n \circ Y \in L^0(\Omega, M_2^n)$ be the finite-dimensional projection of $Y \in L^0(\Omega, M_2)$. Then one can replace (v, η) by Y_n in (IV′) for each $n \geq 1$:*

$$\left.\begin{aligned} pr_1 X\big(t, Y_n, \cdot\big) &= pr_1(Y_n) + \int_0^t H\big(pr_2 X(u, Y_n)(-r), X(u, Y_n)\big)\, du \\ &\qquad + \int_0^t G\big(pr_1 X(u, Y_n)\big) \circ dW(u), \quad t \geq 0, \\ X(0, Y_n) &= Y_n. \end{aligned}\right\} \tag{2.18}$$

Proof. Fix $n \geq 1$. For simplicity of notation, we may assume that W in (IV) is one-dimensional. Apply Theorem 5.3.4 ([Nu.2]) with $u(t, x) := G\big(pr_1 X(t, x)\big)$, $x := (v, \eta) \in M_2^n$, the linear span of $\{e_1, e_2, \dots, e_n\}$, where $\{e_i\}_{i=1}^{\infty}$ is a complete orthonormal system for M_2. All conditions of Theorem 5.3.4 in ([Nu.2]) are satisfied by $u(t, x)$ because $pr_1 X(t, x)$, $(x = (v, \eta))$ satisfies these conditions (by Lemma 2.2), and G is linear (globally Lipschitz). Therefore,

$$\int_0^t G\big(pr_1 X(u, (v, \eta))\big) \circ dW(u)\bigg|_{(v,\eta)=Y_n} = \int_0^t G\big(pr_1 X(u, Y_n)\big) \circ dW(u) \tag{2.19}$$

a.s., where the Stratonovich integral on the right hand side exists. Hence (2.18) holds for each $n \geq 1$. □

The following is the main result in this section. It gives an infinite-dimensional substitution theorem for the semilinear sfde (IV).

Theorem 2.7. *Let $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$. In the semilinear sfde (IV), suppose $H : \mathbf{R}^d \times M_2 \to \mathbf{R}^d$ is C_b^2 and $G : \mathbf{R}^d \to L(\mathbf{R}^m, \mathbf{R}^d)$ is linear. Then*

$$\left.\begin{aligned} dpr_1 X(t, Y) &= H\big(pr_2(X(t, Y))(-r), X(t, Y)\big)\, dt + G\big(pr_1 X(t, Y)\big) \circ dW(t), \quad t > 0, \\ X(0, Y) &= Y. \end{aligned}\right\} \tag{2.20}$$

Indeed, the process $x(t) := pr_1 X(t, Y)$, $t > 0, x_0 := Y$, is the unique solution (in $\mathbb{L}_c^{1,4}$) of the anticipating semilinear sfde (IV).

Proof. The idea is to pass to the limit (a.s.) as $n \to \infty$ in (2.18) of Theorem 2.6.

In (2.18), the following a.s. limit holds:

$$\lim_{n \to \infty} \int_0^t H\big(pr_2 X(u, Y_n)(-r), X(u, Y_n)\big)\, du = \int_0^t H\big(pr_2 X(u, Y)(-r), X(u, Y)\big)\, du \tag{2.21}$$

for $t \geq 0$. This is because of the uniform continuity of the maps

$$M_2 \ni (v, \eta) \mapsto X\big(\cdot, (v, \eta)\big) \in L^2([0, a], M_2)$$

and the linear growth property of H.

The rest of this proof will be devoted to showing the following almost sure convergence:

$$\lim_{n\to\infty}\int_0^t G\big(pr_1X(u,Y_n)\big)\circ dW(u)=\int_0^t G\big(pr_1X(u,Y)\big)\circ dW(u),\quad t\ge 0. \tag{2.22}$$

First, we establish the following

Claim:

Let $Y\in\mathbb{D}^{1,4}(\Omega,M_2)$. Then the process $u\mapsto f(u):=G\big(pr_1X(u,Y)\big)$ belongs to $\mathbb{L}_1^{1,2}$. That is,

$$\lim_{l\to\infty}\int_0^a \sup_{s\le t\le s+\frac{1}{l}} E\big(|\mathcal{D}_s f(t)-(\mathcal{D}^+f)(s)|\big)ds=0, \tag{2.23}$$

and

$$\lim_{l\to\infty}\int_0^a \sup_{s-\frac{1}{l}\le t<s} E\big(|\mathcal{D}_s f(t)-(\mathcal{D}^-f)(s)|\big)ds=0\,. \tag{2.23'}$$

Proof of Claim:

From the definition of f, we have

$$\mathcal{D}_s f(t)=G\mathcal{D}_s pr_1X(t,Y)+GDpr_1X(t,Y)\mathcal{D}_sY,\quad s,t>0\,. \tag{2.24}$$

Express $pr_2X(u,(v,\eta))(-r)$ in the form

$$pr_2X(u,(v,\eta))(-r):=\eta(u-r)1_{[0,r]}(u)+pr_1(X(u-r,(v,\eta)))1_{[r,a]}(u)$$

for $(v,\eta)\in M_2$ and a.e. $u\in[0,a]$. Using the above expression, we can take Malliavin derivatives ($\mathcal{D}_s$) in (2.12) and then replace (v,η) by Y in the resulting integral equation. This gives the following equation, for $s<t$:

$$\begin{aligned}
\mathcal{D}_s pr_1X(t,Y)&=\mathcal{D}_s\psi\big(t,pr_1(Y),\cdot\big)+\int_0^t \mathcal{D}_s\psi\big(t-u,\theta(u)\big)H\big(pr_2X(u,Y)(-r),X(u,Y)\big)\,du\\
&+\int_s^t \psi\big(t-u,\theta(u)\big)1_{[r,a]}(u)D_1H\big(pr_1(X(u-r,Y)),X(u,Y)\big)pr_1\mathcal{D}_sX(u-r,Y)\,du\\
&+\int_s^t \psi\big(t-u,\theta(u)\big)D_2H\big(pr_2X(u,Y)(-r),X(u,Y)\big)\mathcal{D}_sX(u,Y)\,du\\
&=\mathcal{D}_s\psi\big(t,pr_1(Y)\big)+\int_0^t \mathcal{D}_s\psi\big(t-u,\theta(u)\big)H\big(pr_2X(u,Y)(-r),X(u,Y)\big)du\\
&+\int_s^t \psi\big(t-u,\theta(u)\big)1_{[r,a]}(u)D_1H\big(pr_1(X(u-r,Y)),X(u,Y)\big)\mathcal{D}_s pr_1X(u-r,Y)\,du\\
&+\int_s^t \psi\big(t-u,\theta(u)\big)D_2H\big(pr_2X(u,Y)(-r),X(u,Y)\big)\mathcal{D}_sX(u,Y)\,du\,.
\end{aligned} \tag{2.25}$$

In the above equation, D_1,D_2 denote the partial Fréchet derivatives of $H:\mathbf{R}^d\times M_2\to\mathbf{R}^d$ with respect to the first and second variable, respectively. Next, we can take Fréchet derivatives with respect to (v,η) in (2.12) and put $(v,\eta)=Y$ to obtain

$$
\begin{aligned}
&Dpr_1X(t,Y)\mathcal{D}_sY = \\
&\psi(t,\cdot)pr_1\mathcal{D}_sY + \int_0^t \psi\big(t-u,\theta(u)\big)1_{[0,r]}(u)D_1H\big(pr_2X(u,Y)(-r),X(u,Y)\big)pr_2(\mathcal{D}_sY)(u-r)\,du \\
&+\int_0^t \psi\big(t-u,\theta(u)\big)1_{[r,a]}(u)D_1H\big(pr_2X(u,Y)(-r),X(u,Y)\big)Dpr_1X(u-r,Y)\mathcal{D}_sY\,du \\
&+\int_0^t \psi\big(t-u,\theta(u)\big)D_2H\big(pr_2X(u,Y)(-r),X(u,Y)\big)DX(u,Y)\mathcal{D}_sY\,du
\end{aligned} \tag{2.26}
$$

Fix s and let $t\to s^+$ in (2.24), (2.25) and (2.26) to get

$$
\begin{aligned}
&(D^+f)(s) = G\mathcal{D}_s\psi\big(s,pr_1(Y)\big) + G\psi(s,\cdot)pr_1\mathcal{D}_sY \\
&+G\int_0^s \mathcal{D}_s\psi\big(s-u,\theta(u)\big)H\big(pr_2X(u,Y)(-r),X(u,Y)\big)du \\
&+G\int_0^s \psi\big(s-u,\theta(u)\big)1_{[0,r]}(u)D_1H\big(pr_2X(u,Y)(-r),X(u,Y)\big)pr_2(\mathcal{D}_sY)(u-r)\,du \\
&+G\int_0^s \psi\big(s-u,\theta(u)\big)1_{[r,a]}(u)D_1H\big(pr_2X(u,Y)(-r),X(u,Y)\big)Dpr_1X(u-r,Y)\mathcal{D}_sY\,du \\
&+G\int_0^s \psi\big(s-u,\theta(u)\big)D_2H\big(pr_2X(u,Y)(-r),X(u,Y)\big)DX(u,Y)\mathcal{D}_sY\,du\,.
\end{aligned} \tag{2.27}
$$

To prove (2.23), consider

$$
E|\mathcal{D}_sf(t)-(\mathcal{D}^+f)(s)| \le \|G\|E[|\mathcal{D}_s\psi\big(t,pr_1(Y)\big)-\mathcal{D}_s\psi\big(s,pr_1(Y)\big)|] + \sum_{j=1}^{10}E[|I_j(t,s)|] \tag{2.28}
$$

for $t>s$, where

$$
\begin{aligned}
|I_1(t,s)| \le \|G\|\bigg|\int_s^t \psi\big(t-u,\theta(u)\big)1_{[r,a]}(u)D_1H\big(pr_1(X(u-r,Y)),X(u,Y)\big)\cdot \\
\cdot\, pr_1\mathcal{D}_sX(u-r,Y)\,du\bigg|,
\end{aligned} \tag{2.29}
$$

$$
|I_2(t,s)| \le \|G\|\bigg|\int_s^t \psi\big(t-u,\theta(u)\big)D_2H\big(pr_2X(u,Y)(-r),X(u,Y)\big)\mathcal{D}_sX(u,Y)\,du\bigg|, \tag{2.30}
$$

$$
|I_3(t,s)| \le \|G\|\bigg|\int_0^s [\mathcal{D}_s\psi\big(t-u,\theta(u)\big)-\mathcal{D}_s\psi\big(s-u,\theta(u)\big)]H\big(pr_2X(u,Y)(-r),X(u,Y)\big)du\bigg|, \tag{2.31}
$$

$$
|I_4(t,s)| \le \|G\|\bigg|\int_s^t \mathcal{D}_s\psi\big(t-u,\theta(u)\big)H\big(pr_2X(u,Y)(-r),X(u,Y)\big)du\bigg|, \tag{2.32}
$$

$$|I_5(t,s)| \le \|G\| \bigg| \int_0^s [\psi\big(t-u,\theta(u)\big) - \psi\big(s-u,\theta(u)\big)] 1_{[0,r]}(u) D_1 H\big(pr_2 X(u,Y)(-r), X(u,Y)\big) \cdot \cdot\, pr_2(\mathcal{D}_s Y)(u-r)\, du \bigg|, \tag{2.33}$$

$$|I_6(t,s)| \le \|G\| \bigg| \int_s^t \psi\big(t-u,\theta(u)\big) 1_{[0,r]}(u) D_1 H\big(pr_2 X(u,Y)(-r), X(u,Y)\big) \cdot \cdot\, pr_2(\mathcal{D}_s Y)(u-r)\, du \bigg|, \tag{2.34}$$

$$|I_7(t,s)| \le \|G\| \bigg| \int_0^s [\psi\big(t-u,\theta(u)\big) - \psi\big(s-u,\theta(u)\big)] 1_{[r,a]}(u) D_1 H\big(pr_2 X(u,Y)(-r), X(u,Y)\big) \cdot \cdot\, Dpr_1 X(u-r,Y) \mathcal{D}_s Y\, du \bigg|, \tag{2.35}$$

$$|I_8(t,s)| \le \|G\| \bigg| \int_s^t \psi\big(t-u,\theta(u)\big) 1_{[r,a]}(u) D_1 H\big(pr_2 X(u,Y)(-r), X(u,Y)\big) \cdot \cdot\, Dpr_1 X(u-r,Y) \mathcal{D}_s Y\, du \bigg|, \tag{2.36}$$

$$|I_9(t,s)| \le \|G\| \bigg| \int_0^s [\psi\big(t-u,\theta(u)\big) - \psi\big(s-u,\theta(u)\big)] D_2 H\big(pr_2 X(u,Y)(-r), X(u,Y)\big) \cdot \cdot\, DX(u,Y) \mathcal{D}_s Y\, du \bigg|, \tag{2.37}$$

$$|I_{10}(t,s)| \le \|G\| \bigg| \int_s^t \psi\big(t-u,\theta(u)\big) D_2 H\big(pr_2 X(u,Y)(-r), X(u,Y)\big) DX(u,Y) \mathcal{D}_s Y\, du \bigg|. \tag{2.38}$$

In order to establish (2.23) of our claim, we need to estimate each of the terms on the right hand side of (2.28). We will denote all generic positive constants by the letters $C_1, C_2, C_3, \cdots$. To begin with, consider

$$\begin{aligned} E|\mathcal{D}_s\psi\big(t, pr_1(Y)\big) - &\mathcal{D}_s\psi\big(s, pr_1(Y)\big)| \\ &\le \big\{E\|\mathcal{D}_s\psi(t,\cdot) - \mathcal{D}_s\psi(s,\cdot)\|^2\big\}^{1/2} \big\{E\|Y\|^2\big\}^{1/2} \\ &\le C_1\|G\| \bigg\{(t-s)^2 + \int_s^t E\|\mathcal{D}_s\psi(u,\cdot)\|^2 du\bigg\}^{1/2} \big(E\|Y\|^2\big)^{1/2} \\ &\le C_2\|G\|(t-s)^{1/2}\big(E\|Y\|^2\big)^{1/2} \end{aligned} \tag{2.39}$$

because $Y \in L^2(\Omega, M_2)$. Next, we estimate the ten terms $I_j(t,s), 1 \le j \le 10$, in (2.28).

Using (2.7), Hölder's inequality and the fact that H is C_b^1, we have

$$
\begin{aligned}
E|I_1(t,s)| &\le C\|G\| \int_{s\vee r}^{t} \left\{ E \sup_{0\le u\le t} \|\psi(t-u,\theta(u))\|^2 \right\}^{1/2} \cdot \left\{ E\|\mathcal{D}_s X(u-r,Y)\|^2 \right\}^{1/2} du \\
&\le C\|G\| \big[1 + \big(E\|Y\|^2\big)^{1/2}\big](t-s).
\end{aligned}
\tag{2.40}
$$

Estimates for $E|I_j(t,s)|, j = 2, 3, \cdots, 10$, can be similarly derived. We will work out the estimate for $E|I_7(t,s)|$ and leave the rest of the computations to the reader. Using (2.6), the martingale and Hölder's inequalities and the fact that $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$, we get

$$
\begin{aligned}
E|I_7(t,s)| &\le \|G\| \int_r^s \left\{ E\|\psi(t-u,\theta(u)) - \psi(s-u,\theta(u))\|^2 \right\}^{1/2} \cdot \\
&\qquad\qquad \cdot \left\{ E\Big(\|Dpr_1 X(u-r,Y)\|^2 \cdot \|\mathcal{D}_s Y\|^2 \Big) \right\}^{1/2} du \\
&\le C_1 \|G\| (t-s)^{1/2} \int_r^s \{E\|Dpr_1 X(u-r,Y)\|^4\}^{1/4} \cdot \{E\|\mathcal{D}_s Y\|^4\}^{1/4}\, du \\
&\le C_2 (t-s)^{1/2} \{E\|\mathcal{D}_s Y\|^4\}^{1/4}.
\end{aligned}
\tag{2.41}
$$

In view of the above computations, and combining all the estimates on the terms on the right hand side of (2.28), it follows that (2.23) holds. The proof of (2.23′) in the claim follows along similar lines to that of (2.23). To complete the proof of the claim, it remains to check that $f \in \mathbb{L}^{1,2}$; i.e.,

$$
\int_0^a \int_0^a E|\mathcal{D}_s f(t)|^2 dt\, ds < \infty. \tag{2.42}
$$

From (2.24),

$$
\begin{aligned}
E|\mathcal{D}_s f(t)|^2 &\le 2\|G\|^2 E|\mathcal{D}_s pr_1 X(t,Y)|^2 + 2\|G\|^2 E\big\{|Dpr_1 X(t,Y)|^2 \|\mathcal{D}_s Y\|^2\big\} \\
&\le 2\|G\|^2 C\big(1 + E\|Y\|^2\big) + 2\|G\|^2 C\big\{E\|\mathcal{D}_s Y\|^4\big\}^{1/2}.
\end{aligned}
$$

Therefore,

$$
\int_0^a \int_0^a E|\mathcal{D}_s f(t)|^2 dt\, ds \le C_1 \left[1 + \int_0^a \big\{E\|\mathcal{D}_s Y\|^4\big\}^{1/2} ds \right] < \infty
$$

because $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$. Hence $f \in \mathbb{L}^{1,2}$.

Therefore, $f \in \mathbb{L}_1^{1,2}$. This completes the proof of the claim.

We now continue with the proof of (2.22). Since the processes

$$
\left.
\begin{aligned}
u &\mapsto G\big(pr_1 X(u,Y)\big) \\
u &\mapsto G\big(pr_1 X(u,Y_n)\big)
\end{aligned}
\right\}
\tag{2.43}
$$

belong to $\mathbb{L}_1^{1,2}$, it follows from Theorem 5.2.3 in ([Nu.2]) (or Theorem 2.4 of this article) that they are Stratonovich integrable. Furthermore,

$$\int_0^t G\big(pr_1 X(u, Y_n)\big) \circ dW(u)$$
$$= \int_0^t G\big(pr_1 X(u, Y_n)\big) dW(u) + \frac{1}{2}\int_0^t \big(\nabla\big\{G\big(pr_1 X(\cdot, Y_n)\big)\big\}\big)(u)\, du, \ n \geq 1, \quad (2.44)$$

and

$$\int_0^t G\big(pr_1 X(u, Y)\big) \circ dW(u)$$
$$= \int_0^t G\big(pr_1 X(u, Y)\big) dW(u) + \frac{1}{2}\int_0^t \big(\nabla\big\{G\big(pr_1 X(\cdot, Y)\big)\big\}\big)(u)\, du. \quad (2.45)$$

The stochastic integrals on the right hand sides of (2.44) and (2.45) are Skorohod integrals. Set

$$f_n(u) := G\big(pr_1 X(u, Y_n)\big), \quad u \geq 0, \ n \geq 1.$$

Recall that

$$f(u) := G\big(pr_1 X(u, Y)\big), \quad u \geq 0.$$

From the proof of the claim, $f, f_n \in \mathbb{L}_1^{1,2}$. To take limits as $n \to \infty$ in (2.44), we must prove that

$$\lim_{n\to\infty} \int_0^a \int_0^a E|\mathcal{D}_s f_n(u) - \mathcal{D}_s f(u)|^2 ds\, du = 0 \quad (2.46)$$

and

$$\lim_{n\to\infty} \int_0^a \nabla f_n(u) du = \int_0^a (\nabla f)(u) du \quad (2.47)$$

a.s. By (2.25) and a linearized version of (2.12) (in the Fréchet sense), it follows that the random fields

$$(v, \eta) \mapsto \mathcal{D}_s pr_1 X\big(t, (v, \eta)\big)$$
$$(v, \eta) \mapsto D pr_1 X\big(t, (v, \eta)\big)$$

are pathwise continuous. Hence

$$\begin{aligned}\lim_{n\to\infty} \mathcal{D}_s f_n(u) &= G \lim_{n\to\infty} \mathcal{D}_s pr_1 X(u, Y_n) + G \lim_{n\to\infty} D pr_1 X(u, Y_n) \mathcal{D}_s Y_n \\ &= G \mathcal{D}_s pr_1 X(u, Y) + G\, D pr_1 X(u, Y) \mathcal{D}_s Y \\ &= \mathcal{D}_s f(u) \end{aligned} \quad (2.48)$$

a.s. for all $s, u \in [0, a]$. Moreover,

$$\begin{aligned} E|\mathcal{D}_s f_n(u)|^2 &\leq \|G\|^2 E\|\mathcal{D}_s X(u, Y_n)\|^2 + \|G\|^2 \big(E\|D pr_1 X(u, Y_n)(0)\|^4\big)^{1/2} \big(E\|\mathcal{D}_s Y_n\|^4\big)^{1/2} \\ &\leq \|G\|^2 C_1 \big(1 + E\|Y_n\|^2\big) + \|G\|^2 C_1 \big(E\|\mathcal{D}_s Y\|^4\big)^{1/2} \\ &\leq \|G\|^2 C_1 \big(1 + E\|Y\|^2\big) + \|G\|^2 C_1 \big(E\|\mathcal{D}_s Y\|^4\big)^{1/2} \end{aligned}$$
$$(2.49)$$

for all $s, u \in [0, a]$. Since $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$, then by the dominated convergence theorem, (2.46) follows from (2.48) and (2.49).

From (2.27) and (2.24),

$$\begin{aligned}(\nabla f)(u) = G\,\mathcal{D}_u\psi\big(u, pr_1(Y)\big) + G\int_0^u \mathcal{D}_u\psi\big(u-\lambda, \theta(\lambda)\big)H\big(X(\lambda, Y)\big)d\lambda \\ + G\int_0^u \psi\big(u-\lambda, \theta(\lambda)\big)DH\big(X(\lambda, Y)\big)DX(\lambda, Y)\mathcal{D}_uY\,d\lambda \\ + G\,pr_1 DX(u, Y)(0)\mathcal{D}_uY \end{aligned} \tag{2.50}$$

for all $u \in [0, a]$. In the above equation, replace Y by Y_n, let $n \to \infty$, use the continuity of the functions involved and the dominated convergence theorem to get

$$\lim_{n\to\infty} (\nabla f_n)(u) = (\nabla f)(u)$$

a.s. for all $u \in [0, a]$. By dominated convergence and an estimate similar to (2.49), we get

$$\lim_{n\to\infty} \int_0^a (\nabla f_n)(u)du = \int_0^a (\nabla f)(u)du, \quad \text{a.s..} \tag{2.51}$$

From (2.46) and the limit property of the Skorohod integral, we have

$$\lim_{n\to\infty} \int_0^t G\big(pr_1X(u, Y_n)\big)dW(u) = \int_0^t G\big(pr_1X(u, Y)\big)dW(u) \tag{2.52}$$

a.s.. (2.22) follows from (2.44), (2.45), (2.51) and (2.52). This proves (2.20) and the existence of a solution to the semilinear sfde (IV). Uniqueness of the solution to (IV) follows by a similar argument to the one used in the proof of Theorem 3.4 in the next section. This completes the proof of Theorem 2.7. □

3 Anticipating nonlinear sfdes

In this section, we consider the fully nonlinear sfde

$$\left.\begin{aligned} dx(t) = H\big(x(t-r), x(t), x_t\big)\,dt + G\big(x(t), g(x_t)\big) \circ dW(t), \quad t > 0, \\ \big(x(0), x_0\big) = Y \in L^0(\Omega, M_2; \mathcal{F}), \end{aligned}\right\} \tag{I}$$

where W is m-dimensional Brownian motion; the coefficients $H : \mathbf{R}^d \times M_2 \to \mathbf{R}^d$, $G : \mathbf{R}^d \times \mathbf{R}^p \to \mathbf{R}^{d\times m}$ and $g : L^2\big([-r, 0], \mathbf{R}^d\big) \to \mathbf{R}^p$ satisfy the following hypotheses:

Hypotheses (A):

(i) $H : \mathbf{R}^d \times M_2 \to \mathbf{R}^d$ is C^1, Lipschitz on bounded sets and has sublinear growth: That is, there exists $\gamma \in (0, 1)$ and a positive constant C such that

$$|H(v_1, v, \eta)| \leq C\big(1 + |v_1|^\gamma + \|(v, \eta)\|_{M_2}^\gamma\big)$$

for all $(v, \eta) \in M_2$ and $v_1 \in \mathbf{R}^d$.

(ii) $G : \mathbf{R}^d \times \mathbf{R}^p \to \mathbf{R}^{d\times m}$ is $C_b^{2,\delta}$ for some $\delta \in (0,1)$.

(iii) The quasitame mapping $g : L^2\big([-r,0],\mathbf{R}^d\big) \to \mathbf{R}^p$ has the form

$$g(\eta) := \int_{-r}^{0} \sigma\big(\eta(s)\big)\rho(s)\,ds$$

for all $\eta \in L^2([-r,0],\mathbf{R}^d)$, where $\sigma \in C_b^2(\mathbf{R}^d,\mathbf{R}^p)$ is bounded, and $\rho : [-r,0] \to \mathbf{R}$ is C^1.

Under Hypotheses (A), we can reduce (I) to the memory-less diffusion case (I′) in Section 1. It is easy to see that the augmented drift coefficient (H, H_0) in (I′) satisfies Hypotheses (A)(i) above (because H_0 is C_b^1 and globally bounded), while $(G,0)$ still satisfies Hypotheses (A)(ii). Furthermore, the sfdes (I) and (I′) are equivalent in the sense that x is a solution of (I) if and only if the pair (x,z) solves (I′).

Therefore, it is sufficient to consider the fully nonlinear case

$$\left.\begin{aligned} dx(t) &= H(x(t),x_t)\,dt + G(x(t))\circ dW(t), \quad t>0 \\ \big(x(0),x_0\big) &= Y \in L^0(\Omega, M_2) \end{aligned}\right\} \tag{VII}$$

where H and G are assumed to satisfy Hypotheses (A)(i),(ii). Note that, for further simplicity of exposition, we have deliberately dropped the discrete delay in the drift coefficient H in (VII). The case of a discrete delay in H may be treated in a similar fashion to the discussion below.

Let $\psi : \mathbf{R}^+ \times \mathbf{R}^d \times \Omega \to \mathbf{R}^d$ be the $C^{2,\epsilon}, \epsilon \in (0,\delta)$, stochastic flow of the Stratonovich stochastic differential equation

$$\left.\begin{aligned} d\psi(t) &= G\big(\psi(t)\big)\circ dW(t), \quad t>0, \\ \psi(0) &= x \in \mathbf{R}^d, \end{aligned}\right\} \tag{VIII}$$

([Ku]).

In order to establish the existence of a unique solution to the anticipating sfde (VII), we outline the following strategy:

- In (VII), replace the random initial condition Y by a fixed (deterministic) $(v,\eta) \in M_2$, and let $x : [-r,a]\times\Omega \to \mathbf{R}^d$ be an $(\mathcal{F}_t)_{0\le t\le a}$-adapted solution of the resulting sfde
$$\left.\begin{aligned} dx(t) &= H\big(x(t),x_t\big)\,dt + G\big(x(t)\big)\circ dW(t), \quad t>0, \\ (x(0),x_0) &= (v,\eta) \in M_2. \end{aligned}\right\} \tag{VII′}$$
- Define the $(\mathcal{F}_t)_{0\le t\le a}$-adapted process $y : [-r,a]\times\Omega \to \mathbf{R}^d$ by
$$y(t,\omega) := \left\{\begin{matrix} \psi(t,\cdot,\omega)^{-1}(x(t,\omega)), & 0\le t\le a, \\ \eta(t) & -r\le t<0. \end{matrix}\right\} \tag{3.1}$$

- Then an application of the classical (adapted) Itô-Ventzell formula shows that the pair (x, y) satisfy the following system of random equations:

$$\left.\begin{aligned} y(t,\omega) &= v + \int_0^t \big[D\psi\big(u, y(u,\omega),\omega\big)\big]^{-1} H\big(x(u,\omega), x_u(\cdot,\omega)\big)\, du, \quad 0 \le t \le a, \\ x(t,\omega) &= \psi(t, y(t,\omega),\omega), \quad 0 \le t \le a, \\ y_0 &= \eta \in L^2\big([-r,0], \mathbf{R}^d\big),\ x_0 = \eta, \end{aligned}\right\} \tag{3.2}$$

  if and only if x is an adapted solution of the sfde (VII$'$).

- In the random system (3.2), we now replace (v, η) by a random initial condition $Y \in L^0(\Omega, M_2; \mathcal{F})$ and show by a delicate successive approximation argument that (3.2) (with the random initial condition Y) admits a pathwise unique solution (x, y). This is done in Lemma 3.1 and employs the sublinearity requirement on the drift coefficient H.

- The final step in obtaining a solution to the anticipating sfde (VII) is to apply the anticipating Itô-Ventzell formula in [O-P] (Theorem 4.1) to the relation $x(t,\omega) = \psi(t, y(t,\omega),\omega)$, $0 \le t \le a$, in (3.2) (with a random initial condition Y). This application requires that Y and (x, y) are sufficiently smooth in the Malliavin sense with Malliavin derivatives having fourth-order moments; viz. $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$ and (x, y) is locally in $\mathbb{L}_c^{1,4}$, in the sense of [O-P]. Such moment estimates are established in Lemmas 3.2 and 3.3 using a delicate localization argument in (3.2) (with the random initial condition Y), coupled with the Sobolev embedding theorem. This gives existence of a solution to the anticipating sfde (VII). Uniqueness of the solution to (VII), in the Sobolev space $\mathbb{L}_{c,\mathrm{loc}}^{1,4}$, follows from the Itô-Ventzell formula and uniqueness of the solution to the system (3.2) with a random initial condition (Theorem 3.4).

We next proceed with the details of the solution to (VII).

Define $\bar{\psi} : [-r, \infty) \times \mathbf{R}^d \times \Omega \to \mathbf{R}^d$ by

$$\bar{\psi}(t,\cdot,\omega) := \begin{cases} \psi(t,\cdot,\omega), & 0 \le t < \infty, \\ id^{\mathbf{R}^d}, & -r \le t < 0. \end{cases}$$

Lemma 3.1. Assume Hypotheses (A)(i),(ii). When $(v, \eta) \in M_2$ is non-random, the random equation (3.2) is equivalent to the Stratonovich sfde (VII$'$), by setting

$$x(t,\omega) = \bar{\psi}\big(t, y(t,\omega),\omega\big), \quad t \ge -r,\ \omega \in \Omega.$$

Furthermore, if the initial condition (v, η) in (3.2) is replaced by an $\mathcal{F}$-measurable random variable $Y : \Omega \to M_2$, then the random equation (3.2) admits a unique $\mathcal{F}$-measurable solution $(x, y) : [-r, \infty) \times \Omega \to \mathbf{R}^d \times \mathbf{R}^d$ such that $(y(0), y_0) = (x(0), x_0) = Y$.

Proof. Suppose $(v,\eta) \in M_2$ is deterministic, and let the pair (x,y) be a solution of the random equation (3.2). We will show that $x(t) = \psi\big(t, y(t)\big)$, $0 \le t \le a$, satisfies the Stratonovich sfde (VII′) for *fixed* $(v,\eta) \in M_2$. Apply the Itô-Ventzell formula (Stratonovich form) to get

$$\begin{aligned} dx(t) &= \circ d\psi\big(t, y(t)\big) + D\psi\big(t, y(t)\big) y'(t)\, dt \\ &= G\big(\psi\big(t, y(t)\big)\big) \circ dW(t) + D\psi\big(t, y(t)\big) \big[D\psi\big(t, y(t)\big)\big]^{-1} H\big(x(t), x_t\big)\, dt \\ &= H\big(x(t), x_t\big)\, dt + G\big(x(t)\big) \circ dW(t),\ 0 < t < a, \\ x(0) &= v \in \mathbf{R}^d, \quad x_0 = \eta \in L^2\big([-r,0], \mathbf{R}^d\big). \end{aligned}$$

Conversely, by a similar argument, one can show that if x is a solution of (VII′), then the pair (x,y) solves (3.2). This proves the equivalence of the sfde (VII′) and the random system (3.2).

Now consider a random initial condition $Y \in L^0(\Omega, M_2; \mathcal{F})$ in the system (3.2). We will use successive approximations to solve the random equation (3.2) in the *pair* (x,y). Define the sequence $(x^n, y^n) : [-r,a] \times \Omega \to \mathbf{R}^d \times \mathbf{R}^d$, $n \ge 1$, by

$$y^1(t) := \begin{cases} pr_1(Y), & 0 < t \le a, \\ pr_2(Y)(t), & -r < t < 0. \end{cases} \qquad \Bigg| \qquad x^1(t) := \begin{cases} \psi(t, pr_1(Y), \omega), & 0 \le t \le a, \\ pr_2(Y)(t), & -r < t < 0. \end{cases}$$

$$\left.\begin{aligned} y^{n+1}(t) &= pr_1(Y) + \int_0^t \big[D\psi\big(u, y^n(u), \omega\big)\big]^{-1} H\big(x^n(u), x^n_u\big)\, du,\ 0 \le t \le a \\ y^{n+1}(t) &= pr_2(Y)(t), \quad -r < t < 0 \end{aligned}\right\} \tag{3.3}$$

$$x^n(t) := \left.\begin{cases} \psi\big(t, y^n(t), \omega\big), & 0 \le t \le a \\ pr_2(Y)(t), & -r < t < a. \end{cases}\right\} \tag{3.4}$$

From (3.3),

$$y^{n+1}(t) - y^n(t) = \left\{\begin{array}{l} \displaystyle\int_0^t \big\{\big[D\psi(u, y^n(u))\big]^{-1} H\big(x^n(u), x^n_u\big) \\ \qquad\qquad - \big[D\psi(u, y^{n-1}(u))\big]^{-1} H\big(x^{n-1}(u), x^{n-1}_u\big)\big\}\, du, \\ \hfill 0 \le t \le a, \\ 0, \qquad -r < t < 0. \end{array}\right\} \tag{3.5}$$

For $0 \le t \le a$, consider

$$\begin{aligned} |y^{n+1}(t) - y^n(t)| \le& \int_0^t \|\big[D\psi(u, y^n(u))\big]^{-1} - \big[D\psi(u, y^{n-1}(u))\big]^{-1}\|\, |H\big(x^n(u), x^n_u\big)|\, du \\ &+ \int_0^t \|\big[D\psi\big(u, y^{n-1}(u)\big)\big]^{-1}\|\, |H\big(x^n(u), x^n_u\big) - H(x^{n-1}(u), x^{n-1}_u)|\, du. \end{aligned} \tag{3.6}$$

Now $v \mapsto \big[D\psi(u,v)\big]^{-1}$ and H are Lispchitz on bounded sets. So to estimate the right hand side of (3.6) we need to show that

$$\sup_{\substack{0 \le t \le a \\ n \ge 1}} |y^n(t)| < \infty \tag{3.7}$$

and

$$\sup_{\substack{0\le t\le a\\ n\ge 1}} \|\big(x^n(t), x^n_t\big)\|_{M_2} < \infty. \tag{3.8}$$

We prove (3.7) and (3.8) by induction on n. Define

$$y^n_*(t) := \sup_{0\le u\le t} |y^n(u)| \vee 1, \quad n \ge 1. \tag{3.9}$$

Then by (3.4),

$$|x^n(t)| \le K(\omega)\big[1 + |y^n(t)|^{1+\epsilon}\big], \quad n \ge 1,$$

where we have used the fact that for every $\varepsilon > 0$, there exists a random constant $K(\omega) = K_\varepsilon(\omega)$ such that $K \in L^p$ for all $p \ge 1$, and

$$\sup_{0\le t\le a} |\psi(t, x, \omega)| \le K(\omega)(1 + |x|^{1+\varepsilon}).$$

for all $x \in \mathbf{R}^d$. Hence

$$\sup_{0\le u\le t} |x^n(t)| \le K(\omega)\big[1 + |y^n_*(t)|^{1+\epsilon}\big] \tag{3.9$'$}$$

and

$$\begin{aligned}\|x^n_t\|^2_{L^2} &= \int_{-r}^{-t} |pr_2(Y)(t+s)|^2 ds + \int_{-t}^{0} |x^n(t+s)|^2 ds \\ &\le \|pr_2(Y)\|^2_{L^2} + K(\omega)\big[1 + |y^n_*(t)|^{2(1+\epsilon)}\big].\end{aligned}$$

Thus

$$\begin{aligned}\|(x^n(t), x^n_t)\|^\gamma_{M_2} &\le \big\{\|pr_2(Y)\|^2_{L^2} + K(\omega)\big[1 + |y^n_*(t)|^{2(1+\epsilon)}\big]\big\}^{\gamma/2}, \\ &\le \|pr_2(Y)\|^\gamma_{L^2} + K(\omega)\big[1 + |y^n_*(t)|^{\gamma(1+\epsilon)}\big],\end{aligned} \tag{3.10}$$

with $\gamma \in (0, 1)$. Recall that for every $\varepsilon > 0$, there exists a random constant $C := C_\varepsilon$ ($\in L^p$ for all $p \ge 1$), such that

$$|\big[D\psi(u, x, \omega)\big]^{-1}| \le C(1 + |x|^\varepsilon)$$

for all $x \in \mathbf{R}^d$ and $u \in [0, a]$. From (3.3), we get

$$\begin{aligned}y^{m+1}_*(t) &\le |pr_1(Y)| + \int_0^t K(\omega)\big[1 + \big(y^m_*(u)\big)^\epsilon\big] C\big(1 + \|\big(x^m(u), x^m_u\big)\|^\gamma_{M_2}\big)\, du, \\ &\le |pr_1(Y)| + C\|pr_2(Y)\|^\gamma_{L^2} + K(\omega)\big[1 + \big(y^m_*(t)\big)^\epsilon\big] \int_0^t \Big\{1 + \big(y^m_*(u)\big)^{\gamma(1+\epsilon)}\Big\}\, du, \\ &\le C\big(|pr_1(Y)| + \|pr_2(Y)\|^\gamma_{L^2}\big) + K(\omega)\big(y^m_*(t)\big)^\epsilon \int_0^t \big(y^m_*(u)\big)^{\gamma(1+\epsilon)}\, du,\end{aligned} \tag{3.11}$$

for all $t \in [0, a]$, with C and K independent of m.

Define $Y_n : [-r, a] \times \Omega \to \mathbf{R}$ by

$$Y_n(t) := \sup_{1\le m\le n} |y^m_*(t)|, \quad n \ge 1,\ 0 \le t \le a.$$

In (3.11), take $\sup\limits_{1\le m\le n}$ on both sides. Then

$$
\begin{aligned}
Y_{n+1}(t) &\le C\big(|pr_1(Y)| + \|pr_2(Y)\|_{L^2}^{\gamma}\big) + K(\omega)\big(Y_n(t)\big)^{\epsilon}\int_0^t \big(Y_n(u)\big)^{\gamma(1+\epsilon)}\,du \\
&\le C\big(|pr_1(Y)| + \|pr_2(Y)\|_{L^2}^{\gamma}\big) + K(\omega)\big(Y_{n+1}(t)\big)^{\epsilon}\int_0^t \big(Y_{n+1}(u)\big)^{\gamma(1+\epsilon)}\,du, \quad 0\le t\le a,
\end{aligned}
\tag{3.12}
$$

because $Y_n(t) \le Y_{n+1}(t)$, for all $n \ge 1$, and for all $t \in [0,a]$.

Now choose $0<\epsilon<1$ sufficiently small such that

$$\gamma(1+\epsilon) \le 1-\epsilon \Leftrightarrow \gamma + \gamma\epsilon \le 1-\epsilon \Leftrightarrow \epsilon \le \frac{1-\gamma}{1+\gamma}.$$

Divide both sides of (3.12) by $\big(Y_{n+1}(t)\big)^{\epsilon}$:

$$\big(Y_{n+1}(t)\big)^{1-\epsilon} \le C\big(|pr_1(Y)| + \|pr_2(Y)\|_{L^2}^{\gamma}\big) + K(\omega)\int_0^t \big(Y_{n+1}(u)\big)^{1-\epsilon}\,du, \quad 0\le t\le a,\ n\ge 1. \tag{3.13}$$

Applying Gronwall's lemma to (3.13) gives

$$Y_{n+1}(t) \le C\big[|pr_1(Y)|^{\frac{1}{1-\epsilon}} + \|pr_2(Y)\|^{1+\epsilon}\big]e^{\frac{K(\omega)t}{1-\epsilon}}$$

for all $n\ge 1$ and all $t\in[0,a]$. This implies that

$$\sup_{0\le t\le a}|y^n(t)| \le C\big[|pr_1(Y)|^{\frac{1}{1-\epsilon}} + \|pr_2(Y)\|^{1+\epsilon}\big]e^{\frac{K(\omega)a}{1-\epsilon}} < \infty, \tag{3.14}$$

for all $n\ge 1$, where the constants C and K are independent of n. Therefore (3.7) holds; and (3.8) follows directly from (3.7) and (3.10).

Using (3.7), (3.8) and the fact that $v \mapsto \big[D\psi(u,v)\big]^{-1}$ and H are Lipschitz on bounded sets, we get from (3.6):

$$
\begin{aligned}
|y^{n+1}(t) - y^n(t)| \le K(\omega)\int_0^t |y^n(u) - y^{n-1}(u)|\,du \\
+ K(\omega)\int_0^t |x^n(u) - x^{n-1}(u)|\,du + K(\omega)\int_0^t \|x_u^n - x_u^{n-1}\|_{L^2}\,du,
\end{aligned}
\tag{3.15}
$$

for $0\le t\le a$. Now

$$
\begin{aligned}
|x^n(u) - x^{n-1}(u)| &= |\psi\big(u, y^n(u)\big) - \psi\big(u, y^{n-1}(u)\big)| \\
&\le K(\omega)|y^n(u) - y^{n-1}(u)|
\end{aligned}
\tag{3.16}
$$

and

$$
\begin{aligned}
\|x_u^n - x_u^{n-1}\|_{L^2}^2 &= \int_{-r}^0 |x^n(u+s) - x^{n-1}(u+s)|^2\,ds \\
&= \int_{u-r}^u |x^n(s') - x^{n-1}(s')|^2\,ds' \\
&\le K(\omega)\int_0^u |y^n(s) - y^{n-1}(s)|\,ds
\end{aligned}
\tag{3.17}
$$

for all $u \in [0,a]$. By (3.15), (3.16), and (3.17), we get

$$|y^{n+1}(t) - y^n(t)| \leq K(\omega) \int_0^t |y^n(u) - y^{n-1}(u)|\, du, \quad 0 \leq t \leq a, \; n > 1. \tag{3.18}$$

By induction it follows from (3.18) that

$$|y^{n+1}(t) - y^n(t)| \leq \big(\sup_{0 \leq u \leq a} |y^2(u) - y^1(u)| \big) \frac{K(\omega)^{n-1} t^{n-1}}{(n-1)!}, \quad n \geq 1, \; 0 \leq t \leq a. \tag{3.19}$$

By comparison with the uniformly convergent exponential series

$$\sum_{n=1}^{\infty} \frac{K(\omega)^{n-1} t^{n-1}}{(n-1)!} \sup_{0 \leq u \leq a} |y^2(u) - y^1(u)|,$$

it follows from (3.19) that the series

$$y^1(t) + \sum_{n=1}^{\infty} \big[y^{n+1}(t) - y^n(t) \big], \quad 0 \leq t \leq a, \tag{3.20}$$

converges uniformly in $t \in [0,a]$. Hence the sequence of processes $y^n : [-r,a] \times \Omega \to \mathbf{R}^d$, $n \geq 1$, converges uniformly (pathwise) to a process $y : [-r,a] \times \Omega \to \mathbf{R}^d$ which is sample continuous on $[0,a]$:

$$y(t,\omega) := \begin{cases} \lim_{n \to \infty} y^n(t,\omega), & 0 \leq t \leq a, \\ pr_2(Y)(t), & -r < t < 0. \end{cases} \tag{3.21}$$

Letting $n \to \infty$ in (3.4) we get

$$x(t,\omega) = \lim_{n \to \infty} x^n(t,\omega), \quad t \in [-r,a], \tag{3.22}$$

and

$$x_t = \lim_{n \to \infty} x_t^n, \quad t \in [0,a]. \tag{3.23}$$

If we let $n \to \infty$ in (3.3) and (3.4), this shows that the pair (x,y) is a solution of the random system:

$$\left.\begin{aligned} y(t,\omega) &= pr_1(Y) + \int_0^t \big[D\psi\big(u, y(u,\omega), \omega\big) \big]^{-1} H\big(x(u,\omega), x_u(\cdot,\omega)\big)\, du, \quad 0 \leq t \leq a, \\ x(t,\omega) &= \psi(t, y(t,\omega), \omega), \quad 0 \leq t \leq a, \\ y_0 &= pr_2(Y), \; x_0 = pr_2(Y) \in L^2([-r,0], \mathbf{R}^d), \end{aligned}\right\} \tag{IX}$$

where $Y \in L^0(\Omega, M_2; \mathcal{F})$. Recall that $pr_1 : M_2 \to \mathbf{R}^d$, $pr_2 : M_2 \to L^2([-r,0], \mathbf{R}^d)$ are the projections onto the first and second factors in $M_2 = \mathbf{R}^d \times L^2([-r,0], \mathbf{R}^d)$ (respectively).

Letting $n \to \infty$ in (3.14) and (3.9′), we get the estimates

$$\sup_{0 \leq t \leq a} |y(t)| \leq C \big[|pr_1(Y)|^{\frac{1}{1-\epsilon}} + \|pr_2(Y)\|^{1+\epsilon} \big] e^{\frac{K(\omega)a}{1-\epsilon}} \tag{3.24}$$

and

$$\sup_{0\le t\le a}|x(t)| \le K(\omega)\big[1+|pr_1(Y)|^{\frac{1+\epsilon}{1-\epsilon}} + \|pr_2(Y)\|^{2(1+\epsilon)}\big]e^{\frac{K(\omega)a(1+\epsilon)}{1-\epsilon}}$$
$$= K(\omega)\big[1+|pr_1(Y)|^{\gamma} + \|pr_2(Y)\|^{2(1+\epsilon)}\big]e^{\frac{K(\omega)a}{\gamma}}. \tag{3.25}$$

To prove uniqueness of the solution (x, y) to (IX), let (x_1, y_1), (x_2, y_2) be two solutions of (IX). That is,

$$\left.\begin{aligned} y_i(t,\omega) &= pr_1(Y) + \int_0^t \big[D\psi(u, y_i(u))\big]^{-1} H\big(x_i(u), (x_i)_u(\cdot)\big)\,du,\ 0\le t\le a \\ (y_i)_0 &= pr_2(Y), \quad i=1,2, \end{aligned}\right\} \tag{3.26}$$

and

$$x_i(t) = \left\{\begin{array}{ll} \psi\big(t, y_i(t)\big), & 0\le t\le a, \\ pr_2(Y)(t), & -r<t<0, \end{array}\right\} \quad i=1,2. \tag{3.27}$$

Mimicking the argument used in deriving (3.18), we get

$$|y_1(t)-y_2(t)| \le K(\omega)\int_0^t |y_1(t)-y_2(t)|\,du, \quad 0\le t\le a. \tag{3.28}$$

Then (3.28) implies

$$|y_1(t)-y_2(t)| = 0$$

a.s. for all $t\in[0,a]$. Therefore $y_1(t)=y_2(t)$ for all $t\in[-r,a]$; and so $x_1(t)=\psi(t,y_1(t)) = \psi(t,y_2(t)) = x_2(t)$. Hence $(x_1,y_1)=(x_2,y_2)$ and (pathwise) uniqueness in (IX) holds for any $\mathcal{F}$-measurable $Y:\Omega\to M_2$. This completes the proof of Lemma 3.1. □

Our main objective in this section is to show that the application of the (anticipating) Itô-Ventzell formula is still valid if the initial condition Y in (IX) is allowed to be in $\mathbb{D}^{1,4}(\Omega, M_2)$. The Itô-Ventzell formula is developed in [O-P] and requires estimates on the moments of (x, y) in (IX) and their Malliavin derivatives $\mathcal{D}x$, $\mathcal{D}y$. Lemmas 3.2 and 3.3 below are devoted to these estimates.

In the random system (IX), let $Y\in\mathbb{D}^{1,4}(\Omega, M_2)$ and denote by (x,y) its unique solution.

We first develop moment estimates for y in (IX).

Lemma 3.2. Assume that the coefficients H, G in (IX) satisfy Hypotheses (A)(i),(ii). Let (x,y) be the solution of (IX) with initial condition $Y\in\mathbb{D}^{1,4}(\Omega, M_2)$. Then

$$E\sup_{0\le t\le a}|y(t)|^4 < \infty.$$

Proof. Throughout this proof, we will denote by $K\in L^p$ for all $p\ge 1$, a generic positive random constant which may change from line to line.

In (IX), let

$$Y(t) := \sup_{0\le s\le t}|y(s)|\vee 1, \quad 0\le t\le a. \tag{3.29}$$

Then using (IX) and a similar argument to (3.11), we get

$$\begin{aligned}
|Y(t)| &\le \|Y\| + \int_0^t K(\omega)\big[1+|Y(u)|^\epsilon\big]\big(1+\|\big(x(u),x_u\big)\|_{M_2}^\gamma\big)\,du \\
&\le \|Y\| + \int_0^t K(\omega)\big[1+|Y(u)|^\epsilon\big]\left\{1+\|pr_2(Y)\|^\gamma + K(\omega)\big[1+|Y(u)|^{\gamma(1+\epsilon)}\big]\right\}\,du \\
&\le \|Y\| + K(\omega)\|pr_2(Y)\|^\gamma\int_0^t |Y(u)|^\epsilon\,du + \int_0^t K(\omega)|Y(u)|^{\epsilon+\gamma(1+\epsilon)}\,du \\
&\le \|Y\| + K(\omega)\|Y\|^\gamma|Y(t)|^\epsilon + \int_0^t K(\omega)|Y(u)|^{\epsilon+\gamma(1+\epsilon)}\,du
\end{aligned} \tag{3.30}$$

for all $t \in [0,a]$. For each integer $N \ge 1$, let $\tau^N := \inf\{t \in [0,a] : |Y(t)| > N\}$. Define

$$Y^N(t) := Y(t\wedge\tau^N),\ 0 \le t \le a, N \ge 1.$$

Denote by c a generic positive deterministic constant which may change from line to line. From (3.30),

$$\begin{aligned}
E|Y^N(t)|^4 &\le cE\|Y\|^4 + cE\{K^4\|Y\|^{4\gamma}|Y^N(t)|^{4\epsilon}\} \\
&\qquad + c\int_0^{t\wedge\tau^N} E\big[K^4|Y(u)|^{4\{\epsilon+\gamma(1+\epsilon)\}}\big]\,du, \\
&\le cE\|Y\|^4 + cE\big\{K^4\|Y\|^{4\gamma}|Y^N(t)|^{4\epsilon}\big\} \\
&\qquad + \int_0^t E\big[K^4|Y^N(u)|^{4\{\epsilon+\gamma(1+\epsilon)\}}\big]\,du,
\end{aligned} \tag{3.31}$$

for all $t \in [0,a]$. Apply Hölder's inequality to the right hand side of (3.31) with the choices:

$$p = \frac{1}{\gamma}, \quad q = \frac{1}{\epsilon}, \quad l = (1-\gamma-\epsilon)^{-1}$$

$$q' = \frac{1}{\epsilon+\gamma(1+\epsilon)}, \quad p' = \frac{q'}{q'-1};$$

then $q' > 1$ for $\epsilon < \dfrac{1-\gamma}{1+\gamma}$. Hence,

$$\begin{aligned}
E|Y^N(t)|^4 &\le cE\|Y\|^4 + c\big(E\,K^{4l}\big)^{1/l}\big(E\|Y\|^{4\gamma/\gamma}\big)^\gamma\big(E|Y^N(t)|^{4\epsilon/\epsilon}\big)^\epsilon \\
&\qquad + c\int_0^t \big(E\,K^{4p'}\big)^{1/p'}\big(E|Y^N(u)|^4\big)^{\epsilon+\gamma(1+\epsilon)}du \\
&\le cE\|Y\|^4 + c\big(E\,K^{4l}\big)^{1/l}\big(E\|Y\|^4\big)^\gamma\big(E|Y^N(t)|^4\big)^\epsilon \\
&\qquad + cE\big(K^{4p'}\big)^{1/p'}\big(E|Y^N(t)|^4\big)^{\epsilon+\gamma(1+\epsilon)}
\end{aligned} \tag{3.32}$$

for all $t \in [0,a]$. Choose $\epsilon \in (0,1)$ sufficiently small such that $\epsilon < \dfrac{1-\gamma}{1+\gamma}$. Then $\epsilon < \epsilon+\gamma(1+\epsilon) < 1$. Divide (3.32) by $\big(E|Y^N(t)|^4\big)^{\epsilon+\gamma(1+\epsilon)}$ to get

$$\frac{E|Y^N(t)|^4}{\big(E|Y^N(t)|^4\big)^{1/q'}} \le cE\|Y\|^4 + c\big[E\,K^{4l}\big]^{1/l}\big(E\|Y\|^4\big)^\gamma + c\big(E\,K^{4p'}\big)^{1/p'}, \tag{3.33}$$

for all $t \in [0, a]$. Hence,

$$\big(E|Y^N(t)|^4\big)^{1/p'} \le cE\|Y\|^4 + c\big[E\,K^{4l}\big]^{1/l}\big(E\|Y\|^4\big)^{\gamma} + \big(E\,K^{4p'}\big)^{1/p'}$$

for all $t \in [0, a]$. Since $Y \in \mathbb{D}^{1,4}$ and $K \in L^p$ for all $p \ge 1$, then there exist $c > 0$ (deterministic) and independent of $N \ge 1$ such that

$$E|Y^N(t)|^4 \le c$$

for all $N \ge 1$ and $0 \le t \le a$. Since $Y^N(t) \uparrow Y(t)$, then $E|Y(t)|^4 \le c < \infty$. Therefore,

$$E \sup_{0\le t\le a} |y(t)|^4 \le c < \infty.$$

This completes the proof of the lemma. □

Our next lemma gives moment estimates on the Malliavin derivatives $\mathcal{D}_v y$ of y in the random system (IX).

Lemma 3.3. Assume that the coefficients H, G in (IX) satisfy Hypotheses (A)(i),(ii). Let (x, y) be the solution of (IX) with initial condition $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$. Then $(x, y) \in \mathbb{L}^{1,4}_{c,\text{loc}}$; that is (x, y) is locally in $\mathbb{L}^{1,4}_c$, in the sense of [O-P].

Proof. We will use some of the ideas in the proof of Lemma 4.1 in [O-P]. We start by localizing the random equation (IX).

$$\left.\begin{aligned} y(t,\omega) &= pr_1(Y) + \int_0^t \big[D\psi\big(u, y(u,\omega),\omega\big)\big]^{-1} H\big(x(u,\omega), x_u(\cdot,\omega)\big)\,du, \quad 0 \le t \le a,\\ x(t,\omega) &= \psi(t, y(t,\omega),\omega), \quad 0 \le t \le a,\\ y_0 &= pr_2(Y),\ x_0 = pr_2(Y) \in L^2([-r,0],\mathbf{R}^d). \end{aligned}\right\} \tag{IX}$$

where $Y \in L^0(\Omega, M_2; \mathcal{F})$.

Fix $N \ge 1$ and define smooth (C^∞) bump functions $\phi^N \in C_0^\infty(\mathbf{R}^d, \mathbf{R}^d)$, $\bar{\phi}^N \in C_0^\infty\big(L(\mathbf{R}^d), L(\mathbf{R}^d)\big)$ and $\bar{\bar{\phi}}^N \in C^\infty\big(L^2([-r,0],\mathbf{R}^d),\ L^2\big([-r,0],\mathbf{R}^d\big)\big)$ such that

$$\phi^N(v) = \begin{Bmatrix} v, & |v| \le N \\ 0, & |v| > N+1 \end{Bmatrix},\ v \in \mathbf{R}^d, \qquad \|\phi^N\|_\infty \le N$$

$$\bar{\phi}^N(Z) = \begin{Bmatrix} Z, & |Z| \le N \\ 0, & |Z| > N+1 \end{Bmatrix},\ Z \in L(\mathbf{R}^d), \qquad \|\bar{\phi}^N\|_\infty \le N$$

$$\bar{\bar{\phi}}^N(\eta) = \begin{Bmatrix} \eta, & \|\eta\|_{L_2} \le N \\ 0, & \|\eta\|_{L_2} > N+1 \end{Bmatrix},\ \eta \in L^2([-r,0],\mathbf{R}^d).$$

Note that the existence of (a smooth) $\bar{\bar{\phi}}^N$ is guaranteed by the smoothness of the L^2-norm on $L^2([-r,0],\mathbf{R}^d)$.

For fixed $N \geq 1$, let the pair (y^N, x^N) solve the following localized version of (IX):

$$\left.\begin{aligned} y^N(t) &= \phi^N\big(pr_1(Y)\big) + \int_0^t \bar{\phi}^N \left\{\big[D_2\psi(u, \phi^N\big(y^N(u)\big), \omega)\big]^{-1}\right\} \\ &\qquad\qquad\qquad H\left(\phi^N\big(x^N(u)\big), \bar{\bar{\phi}}^N(x_u^N)\right) du, \ 0 \leq t \leq a, \\ x^N(t) &= \psi\big(t, y^N(t), \omega\big), \quad 0 \leq t \leq a, \\ y_0^N &= \bar{\bar{\phi}}^N\big[pr_2(Y)\big], \ x_0^N = \bar{\bar{\phi}}^N\big(pr_2(Y)\big). \end{aligned}\right\} \qquad (\mathrm{IX}^{(N)})$$

For simplicity, define $b^N : [0,a] \times \mathbf{R}^d \times \mathbf{R}^d \times L^2\big([-r,0], \mathbf{R}^d\big) \times \Omega \to \mathbf{R}^d$

$$b^N(u, v_1, v_2, \eta, \omega) := \bar{\phi}^N \left\{\big[D_2\psi\big(u, \phi^N(v_1), \omega\big)\big]^{-1}\right\} H\big(\phi^N(v_2), \bar{\bar{\phi}}^N(\eta)\big), \tag{3.34}$$

for all $0 \leq u \leq a, v_1, v_2 \in \mathbf{R}^d, \eta \in L^2\big([-r,0], \mathbf{R}^d\big)$. Taking Malliavin derivatives $\mathcal{D}_v$ on both sides of $(\mathrm{IX}^{(N)})$, we get

$$\begin{aligned} \mathcal{D}_v y^N(t) &= D\phi^N\big(pr_1(Y)\big)\mathcal{D}_v\big(pr_1(Y)\big) \\ &\quad + \int_0^t D\bar{\phi}^N \left\{\big[D_2\psi\big(u, \phi^N\big((y^N(u))\big)\big)\big]^{-1}\right\} D_2\big[D_2\psi\big(u, \phi^N\big(y^N(u)\big)\big)\big]^{-1} D\phi^N\big(y^N(u)\big) \\ &\qquad\qquad \times \mathcal{D}_v y^N(u) H\big(\phi^N\big(x^N(u)\big), \bar{\bar{\phi}}^N(x_u^N)\big)\, du \\ &\quad - \int_0^t D\bar{\phi}^N \left\{\big[D_2\psi\big(u, \phi^N\big(y^N(u)\big), \cdot\big)\big]^{-1}\right\} \big[D_2\psi\big(u, \phi^N\big(y^N(u)\big), \cdot\big)\big]^{-1} \\ &\qquad\qquad \times \mathcal{D}_v D_2\psi\big(u, \phi^N\big(y^N(u)\big), \cdot\big)\big[D_2\psi\big(u, \phi^N\big(y^N(u)\big)\big)\big]^{-1} \\ &\qquad\qquad \times H\big(\phi^N\big(x^N(u)\big), \bar{\bar{\phi}}^N(x_u^N)\big)\, du \\ &\quad + \int_0^t \bar{\phi}^N \left\{\big[D_2\psi\big(u, \phi^N\big(y^N(u)\big)\big)\big]^{-1}\right\} D_1 H\big(\phi^N\big(x^N(u)\big), \bar{\bar{\phi}}^N(x_u^N)\big) D\phi^N\big(x^N(u)\big) \\ &\qquad\qquad \times \big\{D_2\psi\big(u, y^N(u), \cdot\big)\mathcal{D}_v y^N(u) + \mathcal{D}_v\psi\big(u, y^N(u), \cdot\big)\big\}\, du \\ &\quad + \int_0^t \bar{\phi}^N \left\{\big[D_2\psi\big(u, \phi^N\big(y^N(u)\big)\big)\big]^{-1}\right\} D_2 H\big(\phi^N\big(x^N(u)\big), \bar{\bar{\phi}}^N(x_u^N)\big) \cdot \\ &\qquad\qquad\qquad\qquad \cdot D\bar{\bar{\phi}}^N(x_u^N)\mathcal{D}_v x_u^N\, du\,. \end{aligned} \tag{3.35}$$

Consider

$$(\mathcal{D}_v x_u^N)(s) = \mathcal{D}_v x^N(u+s) = \begin{cases} \mathcal{D}_v x^N(u+s), & -u < s < 0, \\ \mathcal{D}_v pr_2(Y)(u+s), & -r \leq s < -u\,. \end{cases} \tag{3.36}$$

Hence

$$\begin{aligned} \|\mathcal{D}_v x_u^N\|^2_{L^2\left([-r,0],\mathbf{R}^d\right)} &= \int_{-u}^0 |\mathcal{D}_v x^N(u+s)|^2 ds + \int_{-r}^{-u} |\mathcal{D}_v pr_2(Y)(u+s)|^2 ds \\ &\leq \sup_{0 \leq u' \leq u} |\mathcal{D}_v x^N(u')|^2 + \|\mathcal{D}_v pr_2(Y)\|^2_{L^2\left([-r,0],\mathbf{R}^d\right)}\,. \end{aligned}$$

for $u \in [0, r]$. Thus

$$\|\mathcal{D}_v x_u^N\|_{L^2\left([-r,0],\mathbf{R}^d\right)} \leq \sup_{0 \leq u' \leq u} |\mathcal{D}_v x^N(u')| + \|\mathcal{D}_v pr_2(Y)\|_{L^2\left([-r,0],\mathbf{R}^d\right)}, \; 0 \leq u \leq a. \tag{3.37}$$

Since ϕ^N, $\bar{\phi}^N$, and $\bar{\bar{\phi}}^N$ are globally bounded by deterministic constants and since H is bounded on bounded sets, it follows from (IX$^{(N)}$) that

$$\sup_{0 \leq t \leq a} |y^N(t)| \leq M < \infty \tag{3.38}$$

where $M > 0$ is a deterministic constant (possibly depending on N).

From (IX$^{(N)}$), we have

$$\mathcal{D}_v x^N(t) = D_2\psi(t, y^N(t), \cdot)\mathcal{D}_v y^N(t) + \mathcal{D}_v\psi(t, y^N(t), \cdot), \; 0 \leq t \leq a\,. \tag{3.39}$$

Hence from (3.39) and (3.38) we get

$$|\mathcal{D}_v x^N(t)| \leq \sup_{|v_1| \leq M} \|D_2\psi(t, v_1, \cdot)\| \, |\mathcal{D}_v y^N(t)| + \sup_{|v_1| \leq M} |\mathcal{D}_v\psi(t, v_1, \cdot)|, \; 0 \leq t \leq a\,. \tag{3.40}$$

From (3.37) and (3.40),

$$\begin{aligned} \|\mathcal{D}_v x_t^N\|_{L^2\left([-r,0],\mathbf{R}^d\right)} \leq & \sup_{\substack{|v_1| \leq M \\ 0 \leq t' \leq t}} \|D_2\psi(t', v_1, \cdot)\| \sup_{0 \leq t' \leq t} |\mathcal{D}_v y^N(t')| \\ & + \sup_{\substack{|v_1| \leq M \\ 0 \leq t' \leq a}} |\mathcal{D}_v\psi(t', v_1, \cdot)| + \|\mathcal{D}_v pr_2(Y)\|_{L^2\left([-r,0],\mathbf{R}^d\right)}\,. \end{aligned} \tag{3.41}$$

Denote

$$\theta(t) := \sup_{0 \leq u \leq t} |\mathcal{D}_v y^N(u)|, \; 0 \leq t \leq a\,.$$

Then estimating both sides of (3.35), we get

$$\theta(t) \leq C|\mathcal{D}_v pr_1(Y)| + \sum_{i=1}^{7} I_i \tag{3.42}$$

where

$$I_1 := C_N \int_0^t \|D\bar{\phi}^N\|_\infty \sup_{|v_1| \leq M} \|D_2\big[D_2\psi(u, v_1, \cdot)\big]^{-1}\| \, \|D\phi^N\|_\infty \theta(u)\, du \tag{3.43}$$

$$\begin{aligned} I_2 := C_N \int_0^t \|D\bar{\phi}^N\|_\infty \sup_{|v_1| \leq M} \|\big[D_2\psi(u, v_1, \cdot)\big]^{-1}\| \sup_{|v_1| \leq M} \|\mathcal{D}_v D_2\psi(u, v_1, \cdot)\| \times \\ \times \sup_{|v_1| \leq M} \|\big[D_2\psi(u, v_1, \cdot)\big]^{-1}\|\, du \end{aligned} \tag{3.44}$$

$$I_3 := C_N \int_0^t \|\bar{\phi}^N\|_\infty \|D\bar{\phi}^N\|_\infty \sup_{|v_1| \leq M} \|D_2\psi(u, v_1, \cdot)\| \theta(u)\, du \tag{3.45}$$

$$I_4 := C_N \int_0^t \|\bar{\phi}^N\|_\infty \|D\bar{\phi}^N\|_\infty \sup_{|v_1|\le M} |D_2\psi(u,v_1,\cdot)|\, du \tag{3.46}$$

$$I_5 := C_N \int_0^t \|\bar{\phi}^N\|_\infty \|D\bar{\bar{\phi}}^N\|_\infty \sup_{\substack{|v_1|\le M\\ 0\le u'\le u}} \|D_2\psi(u',v_1,\cdot)\|\theta(u)\, du \tag{3.47}$$

$$I_6 := C_N \int_0^t \|\bar{\phi}^N\|_\infty \|D\bar{\bar{\phi}}^N\|_\infty \sup_{\substack{|v_1|\le M\\ 0\le u'\le u}} |D_2\psi(u',v_1,\cdot)|\, du \tag{3.48}$$

$$I_7 := C_N \int_0^t \|\bar{\phi}^N\|_\infty \|D\bar{\bar{\phi}}^N\|_\infty \|\mathcal{D}_v pr_2(Y)\|_{L^2\left([-r,0],\mathbf{R}^d\right)}\, du\,. \tag{3.49}$$

Now combine (3.42), (3.43)–(3.49) to get

$$\theta(t) \le C_N\|\mathcal{D}_v Y\|_{M_2} + \xi_N + C_N\int_0^t m(u)\theta(u)\, du \tag{3.50}$$

when $\xi_N \in \mathbb{L}^p(\Omega,\mathbf{R}^+)$ for all $p\ge 1$, and

$$m(u) := \sup_{|v_1|\le M} \|D_2\big[D_2\psi(u,v_1,\cdot)\big]^{-1}\| + \sup_{\substack{|v_1|\le M\\ 0\le u'\le u}} \|D_2\psi(u',v_1,\cdot)\|,\ 0\le u\le a\,. \tag{3.51}$$

Apply Gronwall's lemma to (3.50) and get

$$\theta(t) \le \big(C\|\mathcal{D}_v Y\|_{M_2} + \xi_N\big) e^{C\int_0^t m(u)\,du}, \qquad 0\le t\le a\,.$$

Thus

$$\sup_{0\le t'\le t} |\mathcal{D}_v y^N(t')| \le \big(C\|\mathcal{D}_v Y\|_{M_2} + \xi_N\big) e^{C\int_0^t m(u)\,du}, \qquad 0\le t\le a,\ \text{a.s..} \tag{3.52}$$

By the Sobolev embedding theorem, there exists an (even) integer $q_0 > d$ such that

$$\int_0^a m(u)\,du \le C\Bigg(\int_0^a\int_{|v_1|\le M+1}\Big\{\sum_{i=1}^2 \|D_2^{(i)}\big[D_2\psi(u,v_1,\cdot)\big]^{-1}\|^{q_0} + \sum_{i=1}^2 \sup_{0\le u'\le u}\|D_2^{(i)}\psi(u',v_1,\cdot)\|^{q_0}\Big\}\, dv_1\, du\Bigg)^{1/q_0}. \tag{3.53}$$

Let

$$V(\omega) := \int_0^a\int_{|v_1|\le M+1}\Big\{\sum_{i=1}^2 \|D_2^{(i)}\big[D_2\psi(u,v_1,\omega)\big]^{-1}\|^{q_0} + \sum_{i=1}^2 \sup_{0\le u'\le u}\|D_2^{(i)}\psi(u',v_1,\omega)\|^{q_0}\Big\}\, dv_1\, du\Bigg)^{1/q_0}. \tag{3.54}$$

Then

$$\int_0^a m(u)\,du \le CV^{1/q_0}. \tag{3.55}$$

By linearizing the Itô version of the sode

$$d\psi(t,x) = G\big(\psi(t,x,\cdot)\big)\circ dW(t), \quad 0<t<a,$$

it follows that $V \in \mathbb{D}^{1,p}(\Omega,\mathbf{R})$ for all $p \ge 1$. By (3.52) and (3.55), we have

$$\sup_{0\le t'\le t} |\mathcal{D}_v y^N(t')| \le \big(C_N\|\mathcal{D}_v Y\|_{M_2} + \xi_N\big)e^{CV^{1/q_0}}, \qquad \text{a.s.}. \tag{3.56}$$

For any integer $m \ge 1$, define $\beta_m \in C_0^\infty(\mathbf{R},\mathbf{R})$ such that

$$\beta_m(x) = \begin{cases} 1 & \text{if } |x| < m \\ 0 & \text{if } |x| > m+1 \end{cases}$$

and $\beta_m(x) \uparrow 1$ as $m \to \infty$ for all $x \in \mathbf{R}$. Let

$$y^{N,m}(t) := \beta_m(V) y^N(t), \qquad 0<t<a,\ m \ge 1. \tag{3.57}$$

Then $y^{N,m} \in \mathbb{L}^{1,4}$, since

$$\mathcal{D}_v y^{N,m}(t) = \beta_m'(V)\mathcal{D}_v V y^{N,m}(t) + \beta_m(V)\mathcal{D}_v y^N(t), \tag{3.58}$$

and so

$$\begin{aligned} E|\mathcal{D}_v y^{N,m}(t)|^4 &\le c_1\|\beta_m'\|_\infty \big(E|V|^8\big)^{1/2}\big(E|y^N(t)|^8\big)^{1/2} + c_1 E|\mathcal{D}_v y^N(t)|^4 \\ &\le c_2\|\beta_m'\|_\infty\big(E|V|^8\big)^{1/2} M^4 + c_2 E\big(C_N\|\mathcal{D}_v Y\|_{M_2}^4 + |\xi_N|^4\big)e^{C(m+1)^{1/q_0}} \\ &= c_2 M^4\|\beta_m'\|_\infty\big(E|\mathcal{D}_v V|^8\big)^{1/2} + c_2\big(C_N E\|\mathcal{D}_v Y\|_{M_2}^4 + E|\xi_N|^4\big)e^{C(m+1)^{1/q_0}} \\ &< \infty. \end{aligned} \tag{3.59}$$

Now $y^{N,m}(t,\omega) = y^N(t,\omega)$ for all $t \in [0,a]$ and all $\omega \in \Omega_m := \{\omega\in\Omega : |V(\omega)| < m\}$, with $\Omega_m \uparrow \Omega$ as $m\to\infty$. Thus $y^N \in \mathbb{L}^{1,4}_{\text{loc}}$ (as defined in [O-P]). By pathwise uniqueness of the solution to the system (IX), it follows that $y(t) = y^N(t)$ for all $t>0$, a.s. on the set

$$\Omega^N := \{\omega\in\Omega : \|Y(\omega)\| < N, \sup_{0\le t\le a}|y^N(t,\omega)| \le N\},$$

and $\Omega^N \uparrow \Omega$ as $N\to\infty$. Therefore, $y \in \mathbb{L}^{1,4}_{\text{loc}}$. We claim further that in fact, $y \in \mathbb{L}^{1,4}_{c,\text{loc}}$ (as defined in [O-P]). To see this, it is sufficient to prove that $y^{N,m} \in \mathbb{L}^{1,4}_c$ for all $N,m \ge 1$. First, we need to show that the family of functions

$$([0,a]\backslash\{v\}) \ni t \mapsto \mathcal{D}_v y^{N,m}(t) \in \mathbb{L}^4(\Omega,\mathbf{R}^d),\ v\in[0,a], \tag{3.60}$$

is equicontinuous. To see this, fix any $v\in[0,a]$, and consider the differences

$$|\mathcal{D}_v y^{N,m}(t_2) - \mathcal{D}_v y^{N,m}(t_1)| \leq |\beta_m'(V)||\mathcal{D}_v V||y^{N,m}(t_2) - y^{N,m}(t_1)| \\ + |\beta_m(V)||\mathcal{D}_v y^N(t_2) - \mathcal{D}_v y^N(t_1)| \tag{3.61}$$

for $t_1, t_2 \neq v$. Replacing t in (3.35) by t_2 and t_1, subtracting, taking 4-th moments, using (3.56) and (3.41), it follows from (3.61) that there is a positive deterministic constant C such that

$$E|\mathcal{D}_v y^{N,m}(t_2) - \mathcal{D}_v y^{N,m}(t_1)|^4 \leq C|t_2 - t_1|^3 \tag{3.62}$$

for $t_1, t_2 \neq v$, with C *independent* of $v \in [0, a]$. The above inequality implies the equicontinuity of the family (3.60). Secondly, we need to check that

$$\sup_{t,v\in[0,a]} E|\mathcal{D}_v y^{N,m}(t)|^4 < \infty \tag{3.63}$$

for all $N, m \geq 1$. This follows again from (3.58), (3.35), (3.41) and (3.56). Hence $y^{N,m} \in \mathbb{L}_c^{1,4}$ for all $N, m \geq 1$; and so $y \in \mathbb{L}_{c,\text{loc}}^{1,4}$. This completes the proof of Lemma 3.3. □

Since $y \in \mathbb{L}_{c,\text{loc}}^{1,4}$ (by Lemma 3.3), we can apply (a localized version of) Theorem 4.1 in [O-P] to get the main theorem in this section: Theorem 3.4 below. The theorem gives a unique solution in $\mathbb{L}_{c,\text{loc}}^{1,4}$ of the nonlinear sfde (I).

Theorem 3.4. *Let $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$ in the sfde (I) and assume that the coefficients $H : \mathbf{R}^d \times M_2 \to \mathbf{R}^d$, $G : \mathbf{R}^d \to L(\mathbf{R}^m, \mathbf{R}^d)$ and $g : L^2\big([-r, 0], \mathbf{R}^d\big) \to \mathbf{R}^p$ satisfy Hypotheses (A). Then the sfde (I) admits a unique solution x in $\mathbb{L}_{c,loc}^{1,4}$.*

Proof. Let (x, y) be the unique solution of the random equation (IX) with $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$. Since $y \in \mathbb{L}_{c,\text{loc}}^{1,4}$ (by Lemma 3.3), we can apply (a localized version of) Theorem 4.1 in [O-P] to get the following:

$$\begin{aligned} dx(t) &= \circ d\psi(t, y(t)) = G(\psi(t, y(t)) \circ dW(t) + D_2\psi(t, y(t))y'(t)\, dt \\ &= G(x(t)) \circ dW(t) + D_2\psi(t, y(t))[D_2\psi(t, y(t))]^{-1} H(x(t), x_t)\, dt \\ &= H(x(t), x_t)\, dt + G(x(t)) \circ dW(t),\ t > 0, \end{aligned}$$

and $(x(0), x_0) = (y(0), y_0) = Y \in \mathbb{D}^{1,4}(\Omega, M_2)$. This proves existence of a solution to the nonlinear anticipating sfde (VII). Therefore, the sfde (I) admits a solution.

It remains to prove uniqueness of the solution of the sfde (I) in $\mathbb{L}_{c,\text{loc}}^{1,4}$. To do this, let $x^1, x^2 \in \mathbb{L}_{c,\text{loc}}^{1,4}$ be two solutions of (VII) such that $(x^1(0), x_0^1) = (x^2(0), x_0^2) = Y \in \mathbb{D}^{1,4}(\Omega, M_2)$. Define

$$y^i(t) := \psi^{-1}(t, x^i(t)),\ t \geq 0, \quad y_0^i := pr_2(Y),\ i = 1, 2.$$

Then $x^i(t) = \psi(t, y^i(t)),\ t \geq 0, i = 1, 2$. Applying the Itô-Ventzell formula gives

$$\begin{aligned} \circ dx^i(t) &= \circ d\psi(t, y^i(t)) + D_2\psi(t, y^i(t)) \circ dy^i(t) \\ &= G(\psi(t, y^i(t)) \circ dW(t) + D_2\psi(t, y^i(t)) \circ dy^i(t) \\ &= G(x^i(t)) \circ dW(t) + H(x^i(t), x_t^i)\, dt,\ t > 0,\ i = 1, 2. \end{aligned}$$

Thus

$$dy^i(t) = [D_2\psi(t, y^i(t))]^{-1} H(x^i(t), x^i_t)\, dt, \;\; t > 0, \;\; i = 1, 2.$$

This says that the pair (x^i, y^i) is a solution of the random system (IX) for $i = 1, 2$, with the same initial data $Y \in \mathbb{D}^{1,4}(\Omega, M_2)$. By pathwise uniqueness of the solution to (IX) (Lemma 3.1), it follows that $x^1(t) = x^2(t)$ (and $y^1(t) = y^2(t)$) for all $t \geq 0$, a.s. This completes the proof of uniqueness of the solution to the nonlinear anticipating sfde

$$\left.\begin{aligned} dx(t) &= H\big(x(t), x_t\big)\, dt + G\big(x(t)\big) \circ dW(t), \quad t > 0 \\ (x(0), x_0) &= Y \in \mathbb{D}^{1,4}(\Omega, M_2). \end{aligned}\right\} \qquad \text{(VII)}$$

Hence the sfde (I) has a unique solution in $\mathbb{L}^{1,4}_{c,\text{loc}}$. □

References

[A-I] Arnold, L., and Imkeller, P., Stratonovich calculus with spatial parameters and anticipative problems in multiplicative ergodic theory, *Stochastic Processes ad their Applications,* Vol. 62 (1996), 19–54.

[G-Nu-S] Grorud, A., Nualart, D., and Sanz-Solé, M., Hilbert-valued anticipating stochastic differential equations, *Annales de l'institut Henri Poincaré (B) Probabilit et Statistiques*, 30 no. 1 (1994), 133-161.

[Ku] Kunita, H., *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, Cambridge, New York, Melbourne, Sydney (1990).

[Ma] Malliavin, P., Stochastic calculus of variations and hypoelliptic operators, *Proceedings of the International Conference on Stochastic Differential Equations, Kyoto*, Kinokuniya, 1976, 195-263.

[Mo.1] Mohammed, S.-E.A., *Stochastic Functional Differential Equations*, Research Notes in Mathematics, no. 99, Pitman Advanced Publishing Program, Boston-London-Melbourne (1984).

[Mo.2] Mohammed, S.-E. A., Non-Linear flows for linear stochastic delay equations, *Stochastics,* Vol. l7 #3, (1987), 207–212.

[Mo.3] Mohammed, S.-E. A., The Lyapunov spectrum and stable manifolds for stochastic linear delay equations, *Stochastics and Stochastic Reports*, Vol. 29 (1990), 89-131.

[M-S.1] Mohammed, S.-E. A., and Scheutzow, M. K. R., The stable manifold theorem for nonlinear stochastic systems with memory, Part I: Existence of the semiflow, *Journal of Functional Analysis*, 205, (2003), 271-305. Part II: The local stable manifold theorem, *Journal of Functional Analysis*, 206, (2004), 253-306.

[M-Z] Mohammed, S.-E. A., and Zhang, T. S., The substitution theorem for semilinear stochastic partial differential equations, *Journal of Functional Analysis*, 253, (2007), 122-157.

[Nu.1] Nualart, D., *The Malliavin Calculus and Related Topics*, Probability and its Applications, Springer-Verlag (1995).

[Nu.2] Nualart, D., Analysis on Wiener space and anticipating stochastic calculus, *Springer LNM*, 1690, Ecole d'Et'e de Probabilit'es de Saint-Flour XXV-1995, ed: P. Bernard (1995).

[N-P] Nualart, D., and Pardoux, E., Stochastic calculus with anticipating integrands, Analysis on Wiener space and anticipating stochastic calculus, *Probab. Th. Rel. Fields* , 78 (1988), 535-581.

[O-P] Ocone, D., and Pardoux, E., A generalized Itô-Ventzell formula. Application to a class of anticipating stochastic differential equations. Ann. Inst. Henri Poincaré, Probabilités et Statistiques 25:1 (1989) 39-71.

[T] Tindel, S., Stochastic parabolic equations with anticipative initial condition. Stochastics and Stochastic Reports 62 (1997), pp. 1-20.

Salah-Eldin A. Mohammed
Department of Mathematics
Southern Illinois University
Carbondale, Illinois 62901, USA.
Email: salah@sfde.math.siu.edu
Web page: http://sfde.math.siu.edu

Tusheng Zhang
School of Mathematics
Oxford Road
University of Manchester
Manchester M13 9PL
ENGLAND
Email: tzhang@maths.man.ac.uk
Web page: http://www.maths.man.ac.uk/ tzhang/