

# Dynamical Low-Rank Approximation for Stochastic Differential Equations

Yoshihito Kazashi<sup>1</sup>, Fabio Nobile<sup>2</sup>, and Fabio Zoccolan<sup>3</sup>

<sup>1</sup>Department of Mathematics & Statistics, University of Strathclyde, 26 Richmond St., Glasgow, G1 1XH, UK. email: y.kazashi@strath.ac.uk

<sup>2</sup>Institut de Mathématiques, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland. email: fabio.nobile@epfl.ch

<sup>3</sup>Institut de Mathématiques, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland. email: fabio.zoccolan@epfl.ch

## Abstract

In this paper, we set the mathematical foundations of the Dynamical Low-Rank Approximation (DLRA) method for stochastic differential equations (SDEs). DLRA aims at approximating the solution as a linear combination of a small number of basis vectors with random coefficients (low rank format) with the peculiarity that both the basis vectors and the random coefficients vary in time.

While the formulation and properties of DLRA are now well understood for random/parametric equations, the same cannot be said for SDEs and this work aims to fill this gap. We start by rigorously formulating a Dynamically Orthogonal (DO) approximation (an instance of DLRA successfully used in applications) for SDEs, which we then generalize to define a parametrization independent DLRA for SDEs. We show local well-posedness of the DO equations and their equivalence with the DLRA formulation. We also characterize the explosion time of the DO solution by a loss of linear independence of the random coefficients defining the solution expansion and give sufficient conditions for global existence.

## 1 Introduction

This paper is concerned with the theoretical foundation of dynamical low-rank methods for stochastic differential equations (SDEs). Low-rank methods aim to approximate solutions of high-dimensional differential equations in a well chosen low dimensional subspace. Such methods are widely used in computational science and industrial applications [2, 24, 32, 38]. Our focus here is on low-rank methods for high-dimensional SDEs, which is a case of primary interest, for instance in finance [34, 33], or in various applications such as biology [1] or machine learning [29] where SDEs often appear as discretizations of Stochastic Partial Differential Equations (SPDEs).

Among such low-rank methods, the so-called Dynamical Low-Rank Approximation (DLRA) has gained notable attention and has been successfully applied in many applications in recent years [20, 18, 30]. Unlike traditional reduced basis methods, however, DLRA allows the low-dimensional subspace to evolve over time, enabling it to adapt to possibly rapid temporal variations of the solution. Thus DLRA evolves in a low-rank

manifold, not in a fixed linear space of low dimension. In its original form [20], such evolution in a manifold is achieved by enforcing the temporal derivative of the DLRA to be in the corresponding tangent space, which requires the differentiability of the approximate solution in time.

An equivalent formulation of the DLRA is the Dynamically Orthogonal (DO) approximation [35]. This strategy can approximate random time-dependent equations by using a sum of a small number of products between spatial and stochastic basis functions, thus leading to a low-rank approximation with explicit parametrization of the low-rank factors. In the context of random partial differential equations (RPDEs), the DO/DLRA methodology has been successfully applied in several problems, showing promising computational results [8, 6, 7, 10, 28, 26, 27, 17, 35].

The goal of this work is to set a theoretical framework of DLRA for SDEs. This is a challenging task, because, unlike the RPDEs setting, the differentiability in time of the solution cannot be exploited due to the presence of the rough diffusion term. Our approach is to first define DO dynamics for a general SDE, specifying conditions for the well-posedness of this framework. From this DO system, we derive an equation that is independent of the parametrization. We define DLRA for SDEs as the solution to this parameter-independent equation, and show that these DO and DLRA methodologies are equivalent.

The DO approximation is given by the solution of a system of equations, which provide the evolution of the deterministic and stochastic basis functions. This system is however highly nonlinear, even if the original dynamics was linear. In addition, its well-posedness is not obvious due to the presence of the inverse of a Gramian, which is not guaranteed to exist at all times. Moreover, in the specific case of SDEs, the DO solution depends on the law of the whole process at every instant of time. These features make the well-posedness study challenging and not-standard, questioning also whether the existence of solutions is global or only local and, in the latter case, what happens to the solution at the explosion time. Besides being of interest in itself, a well-posedness study of the DO equations is also important to derive consistent and stable time discretization schemes, which are eventually needed to apply these techniques in real-life problems.

For RPDEs, theoretical results on the existence and uniqueness of the solution of the DO equations [16], as well as stability and error estimates of time discretization schemes [17, 28, 42] are available. An existence and uniqueness result is also available for the DLR approximation of two-dimensional deterministic PDEs [3], where the sum involves products of basis functions for the different spatial variables. In contrast, for SDEs well-posedness results remain largely unexplored. This lack of theory for SDEs is unfortunate, since DO approximations are very appealing and, despite no results on its well-posedness, have been also used in several problems modeled by SDEs (see e.g. [35, 36, 37, 40]). The reader can find other approaches of initiating a DLRA framework starting from the Fokker-Plank equation in [35, 4].

This paper aims at filling this gap between application and theory by:

- (Re-)deriving the DO equations for SDEs. We also derive the corresponding parametrization independent DLRA equation for SDEs.
- Showing the equivalence between DO approximation and DLRA. By exploiting this equivalence, we show that the DO equations and the DLRA equation are well-posed.

- Characterising the finite explosion time in terms of the linear independence of the stochastic DO basis.
- Discussing the extension of DLRA beyond the explosion time. We also provide a sufficient condition under which the DO/DLRA solution exists globally.

More precisely, we start by revisiting the DO equations for SDEs, which were originally derived in [35] relying on the formal assumption that the DO solution is time differentiable, a property that SDE solutions do not possess. Deriving DO equations without using the time derivatives is a fundamental challenge, since the essence of the usual DLRA methodology consists in projecting the time derivative onto the tangent space of the low-rank manifold. We overcome this challenge by pushing the differentiability to the spatial basis, while the stochastic basis remains non-differentiable. It turns out that this alternative approach leads to the same DO equations as in [35], supporting their validity as a correct dynamical low-rank approximation.

In the DLRA literature [20, 18], it is well known by now that, when applied to RPDEs, the DO formulation is just a specific parametrization of the DLRA, the parameters being the deterministic and stochastic time-dependent bases defining the DO solution [28, 16]. In this paper, we derive an analogous result for SDEs. In particular, starting from the DO equations, we derive a parameter independent low-rank approximation which we name DLRA for SDEs, and show its equivalence with the DO formulation. The main result of this work is to prove local existence and uniqueness to both DO and DLRA equations. As a part of our existence result, we give a characterisation of the interval  $[0, T_e)$  on which the solution exists. More specifically, we show that, if the *explosion time*  $T_e$  is finite, then the stochastic DO basis must become linearly dependent at  $T_e$ .

Although the characterisation of the explosion time provides us a valuable insight into the DO solution, it is not satisfactory from the practical point of view; if the true solution still exists at  $T_e$  and beyond, we expect any sensible approximation to exist as well. Under some mild integrability condition on the initial datum, we show that, while the DO solution ceases to exist at the explosion time  $T_e$ , the DLRA can be continuously extended up to  $T_e$  and beyond. Thus, our findings offer insights on how to continue a DO approximation *beyond* the explosion time  $T_e$ . As a final result, we show that a sufficient condition for global existence of a DO solution (i.e.  $T_e = +\infty$ ) is that the noise in the SDE is non-degenerate.

In passing, we point out that the DO/DLRA equation derived in this work cannot be regarded as a trivial application of the standard theory of SDEs on manifolds (see e.g. [11, 12]). Indeed, the projection operator employed in the DO decomposition is not standard, as it also involves a purely stochastic term that depends on the average of all paths.

The rest of the paper is organised as follows. Section 2 introduces the DO equations, together with theoretical justifications for why they are sensible. Furthermore, it introduces a DLRA formulation that does not depend on the parametrization. Section 3 concerns the local well-posedness of the DO equations, where we also show the continuity of the solution with respect to the initial data. In Section 4, we study the behaviour of the solution up to and beyond the explosion time  $T_e$ . In Section 5 we draw some conclusions and perspectives.

## 2 The DO equations

Let  $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \geq 0})$  be a filtered complete probability space with the usual conditions; see for example [39, Remark 6.24]. Consider the stochastic differential equation (SDE)

$$X^{\text{true}}(t) = X_0^{\text{true}} + \int_0^t a(s, X^{\text{true}}(s)) ds + \int_0^t b(s, X^{\text{true}}(s)) dW_s, \quad (2.1)$$

where  $W_t = (W_t^1, \dots, W_t^m)^\top$  is a standard  $m$ -dimensional  $(\mathcal{F}_t)$ -Brownian motion. Here, we used the notation  $X^{\text{true}}(t, \omega) = (X_1^{\text{true}}(t, \omega), \dots, X_d^{\text{true}}(t, \omega))^\top \in \mathbb{R}^d$  for  $t \geq 0$  and  $\omega \in \Omega$ . Let  $|\cdot|$  and  $\|\cdot\|_F$  denote the Euclidean norm and the Frobenius norm, respectively. We work with the following general assumptions on the coefficients of the SDE (2.1).

**Assumption 1.** The drift coefficient  $a: [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  and the diffusion coefficient  $b: [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times m}$  are jointly measurable and Lipschitz continuous with respect to the second variable, uniformly in time, i.e.

$$\begin{cases} |a(s, x) - a(s, y)| \leq C_{\text{Lip}}|x - y| \\ \|b(s, x) - b(s, y)\|_F \leq C_{\text{Lip}}|x - y|, \end{cases} \quad (2.2)$$

for some constant  $C_{\text{Lip}} > 0$ .

**Assumption 2.** The drift  $a$  and the diffusion  $b$  satisfy the following linear-growth bound condition uniformly in time:

$$|a(s, x)|^2 + \|b(s, x)\|_F^2 \leq C_{\text{lgb}}(1 + |x|^2) \quad (2.3)$$

for some constant  $C_{\text{lgb}} > 0$ .

The choice of the Frobenius norm in Assumption 1 is made primarily for convenience when utilizing Ito's isometry; this is standard practice. Since we are in a finite-dimensional setting, the results presented in this work remain valid under Lipschitz continuity with respect to other norms, with a modified constant. It is also worth mentioning that in the infinite-dimensional case, Lipschitz continuity is typically considered with respect to the Hilbert–Schmidt norm between appropriate spaces; see for example [9]; hence our choice of the Frobenius norm is a first step in this direction. Furthermore, we assume that the initial condition  $X_0^{\text{true}}$  in (2.1) satisfies the following:

**Assumption 3.**

$$X_0^{\text{true}} \text{ is } \mathcal{F}_0\text{-measurable and satisfies } \mathbb{E}[|X_0^{\text{true}}|^2] < +\infty. \quad (2.4)$$

Under these assumptions, equation (2.1) has a unique strong solution; see for example [39, Theorem 21.13].

Let us consider a positive integer  $R$  such that  $R \leq d$ . To numerically approximate (2.1), in this work we consider dynamically orthogonal approximations of the form

$$X^{\text{true}} \approx X := \mathbf{U}^\top \mathbf{Y} := \sum_{j=1}^R U^j Y^j \in \mathbb{R}^d, \quad (2.5)$$

where  $\mathbf{U} = (\mathbf{U}_t)_{t \in [0, T]}$  is a deterministic absolutely continuous matrix-valued function that gives a matrix in  $\mathbb{R}^{R \times d}$  with orthogonal rows for all  $t$ , whereas  $\mathbf{Y} = (\mathbf{Y}_t)_{t \in [0, T]}$  is an

Itô process with values in  $\mathbb{R}^R$  with linearly independent components for all  $t$ . It is worth pointing out that for all  $t \in [0, T]$  and  $\omega \in \Omega$ , the approximate process  $X_t(\omega)$  belongs to an  $R$  dimensional vector space  $\text{span}\{U_t^1, \dots, U_t^R\}$  spanned by the rows  $U_t^1, \dots, U_t^R$  of  $\mathbf{U}_t$ , whereas each component  $X_t^j, j = 1, \dots, d$ , belongs to  $\text{span}\{Y_t^1, \dots, Y_t^R\}$ .

Approximations of the form (2.5) for SDEs have been considered already in [35], where the following evolution equations, hereafter called *DO equations* were derived for the factors  $(\mathbf{U}, \mathbf{Y})$  by using the same formal procedure as for ODEs via treating the process  $\mathbf{Y}$  as differentiable in time:

$$C_{\mathbf{Y}_t} \dot{\mathbf{U}}_t = \mathbb{E}[\mathbf{Y}_t a(t, \mathbf{U}_t^\top \mathbf{Y}_t)^\top] (I_{d \times d} - P_{\mathbf{U}_t}^{\text{row}}), \quad (2.6)$$

$$d\mathbf{Y}_t = \mathbf{U}_t a(t, \mathbf{U}_t^\top \mathbf{Y}_t) dt + \mathbf{U}_t b(t, \mathbf{U}_t^\top \mathbf{Y}_t) dW_t. \quad (2.7)$$

Here, we let  $C_{\mathbf{Y}_t} := \mathbb{E}[\mathbf{Y}_t \mathbf{Y}_t^\top]$  whereas  $P_{\mathbf{U}_t}^{\text{row}} \in \mathbb{R}^{d \times d}$  denotes the projection-matrix onto the row space  $\text{span}\{U_t^1, \dots, U_t^R\} \subset \mathbb{R}^d$  of  $\mathbf{U}_t$ ; when  $\mathbf{U}_t$  has orthonormal rows, one has  $P_{\mathbf{U}_t}^{\text{row}} = \mathbf{U}_t^\top \mathbf{U}_t$ . In Section 2.1, we will give a rigorous justification for (2.6) and (2.7).

Having stated these equations, we now define our notion of strong DO solution and DO approximation for an SDE problem of the type (2.1).

**Definition 2.1** (Strong DO solution). A function  $(\mathbf{U}, \mathbf{Y}) : [0, T] \rightarrow \mathbb{R}^{R \times d} \times L^2(\Omega; \mathbb{R}^R)$  is called a *strong DO solution* for (2.1) if the following conditions are satisfied:

1. the initial conditions  $(\mathbf{U}_0, \mathbf{Y}_0)$  are such that  $\mathbf{U}_0 \in \mathbb{R}^{R \times d}$  is a matrix with orthonormal rows and  $\mathbf{Y}_0 \in L^2(\Omega; \mathbb{R}^R)$  has linearly independent components;
2. the curve  $t \rightarrow \mathbf{U}_t \in \mathbb{R}^{R \times d}$  is absolutely continuous on  $[0, T]$  and  $\mathbf{U}_t \dot{\mathbf{U}}_t^\top = 0 \in \mathbb{R}^{R \times R}$  for a.e.  $t \in [0, T]$ ;
3. the curve  $t \rightarrow \mathbf{Y}_t(\omega) \in \mathbb{R}^R$  has almost surely continuous paths on  $[0, T]$  and it is  $\mathcal{F}_t$ -measurable for all  $t \in [0, T]$ . Moreover, for any  $t \in [0, T]$  the components  $Y_t^1, \dots, Y_t^R$  are linearly independent in  $L^2(\Omega)$ ;
4.  $\mathbf{U}$  satisfies equation (2.6) for a.e.  $t \in [0, T]$  and  $\mathbf{Y}$  is a strong solution of (2.7) on  $[0, T]$ .

We remark that points (1) and (2) in Definition 2.1 imply that  $\mathbf{U}$  has orthonormal rows for all  $t \in [0, T]$ , i.e.  $\mathbf{U}_t \mathbf{U}_t^\top = I \in \mathbb{R}^{R \times R}$  for all  $t \in [0, T]$ . Furthermore, the term “strong” in Definition 2.1 follows the standard terminology for the solution of SDEs: we seek a solution on the given probability space and Brownian motion. See for example [14] for more details. For convenience, given a DO solution  $(\mathbf{U}, \mathbf{Y})$  we call the product  $\mathbf{U}_t^\top \mathbf{Y}_t$  a DO approximation.

**Definition 2.2** (DO approximation). We call a process  $X : [0, T] \rightarrow L^2(\Omega; \mathbb{R}^d)$  a *DO approximation* of (2.1) if there exists a strong DO solution  $(\mathbf{U}, \mathbf{Y})$  such that  $X_t := \mathbf{U}_t^\top \mathbf{Y}_t$  for all  $t \in [0, T]$ .

Given a DO approximation  $X$ , the corresponding DO solution  $(\mathbf{U}, \mathbf{Y})$  is determined only up to a (process of) rotation matrix. Indeed, let  $(\mathbf{U}, \mathbf{Y})$  and  $(\tilde{\mathbf{U}}, \tilde{\mathbf{Y}})$  be two strong DO solutions such that  $\mathbf{U}_t^\top \mathbf{Y}_t = \tilde{\mathbf{U}}_t^\top \tilde{\mathbf{Y}}_t = X_t$  for all  $t \geq 0$ . Then the orthogonality of the rows of  $\mathbf{U}_t$  implies  $\tilde{\mathbf{Y}}_t = O_t \mathbf{Y}_t$  with  $O_t = \tilde{\mathbf{U}}_t \mathbf{U}_t^\top$ . The matrix  $O_t$  is orthogonal for every  $t \in [0, T]$ , but not necessarily an identity. See Section 2.3, in particular Proposition 2.10, for more details.

## 2.1 Consistency of the DO equations

In this section, we rigorously show the consistency of the DO equations (2.6) and (2.7) in the sense described hereafter. Assume that the exact solution of (2.1) is of the form  $X = \sum_{j=1}^R U^j Y^j$  with deterministic function  $\mathbf{U}_t = (U_t^1, \dots, U_t^R)^\top \in \mathbb{R}^{R \times d}$  and an Itô process  $\mathbf{Y}_t(\omega) = (Y_t^1(\omega), \dots, Y_t^R(\omega))^\top \in [L^2(\Omega)]^R$  for some  $R \leq d$  that fulfil the following properties:

1. the function  $[0, T] \ni t \mapsto \mathbf{U}_t \in \mathbb{R}^{R \times d}$ ,  $d \geq R$ , is absolutely continuous on  $[0, T]$  and satisfies  $\mathbf{U}_t \dot{\mathbf{U}}_t^\top = 0 \in \mathbb{R}^{R \times R}$  for almost every  $t \in [0, T]$ ; moreover,  $\mathbf{U}_t \mathbf{U}_t^\top = I \in \mathbb{R}^{R \times R}$  for almost every  $t \in [0, T]$ , where  $I$  is the identity matrix;
2.  $\mathbf{Y} = (\mathbf{Y}_t)_{t \in [0, T]}$  is an Itô process

$$d\mathbf{Y}_t = \alpha_t dt + \beta_t dW_t, \quad (2.8)$$

with coefficients  $\alpha_t \in \mathbb{R}^R$  and  $\beta_t \in \mathbb{R}^{R \times m}$ , where  $\alpha_t$  and  $\beta_t$  are progressively measurable and have a continuous path almost surely.

Then  $(\mathbf{U}_t, \mathbf{Y}_t)$  must satisfy equations (2.6) and (2.7).

Indeed, since  $\alpha_t$  and  $\beta_t$  in (2.8) are progressively measurable, by applying Itô's formula [13, 22] we have

$$dX_t = (d\mathbf{U}_t^\top) \mathbf{Y}_t + \mathbf{U}_t^\top d\mathbf{Y}_t + \sum_{j=1}^R d\langle U^j, Y^j \rangle(t) \quad (2.9)$$

$$= \dot{\mathbf{U}}_t^\top \mathbf{Y}_t dt + \mathbf{U}_t^\top (\alpha_t dt + \beta_t dW_t) + 0 \cdot \beta_t dt \quad (2.10)$$

$$= (\dot{\mathbf{U}}_t^\top \mathbf{Y}_t + \mathbf{U}_t^\top \alpha_t) dt + \mathbf{U}_t^\top \beta_t dW_t, \quad (2.11)$$

where  $\langle U^j, Y^j \rangle$ ,  $j = 1, \dots, R$  is the quadratic covariation of  $U^j$  and  $Y^j$ . Since  $X$  is assumed to satisfy (2.1), the uniqueness of the representation of Itô processes implies

$$\dot{\mathbf{U}}_t^\top \mathbf{Y}_t + \mathbf{U}_t^\top \alpha_t = a(t, \mathbf{U}_t^\top \mathbf{Y}_t) \quad (2.12)$$

$$\mathbf{U}_t^\top \beta_t = b(t, \mathbf{U}_t^\top \mathbf{Y}_t). \quad (2.13)$$

Now, using  $\mathbf{U}_t \dot{\mathbf{U}}_t^\top = 0 \in \mathbb{R}^{R \times R}$  and  $\mathbf{U}_t \mathbf{U}_t^\top = I \in \mathbb{R}^{R \times R}$  for almost every  $t \in [0, T]$ , these equalities imply

$$\alpha_t = \mathbf{U}_t a(t, \mathbf{U}_t^\top \mathbf{Y}_t) \quad \text{and} \quad \beta_t = \mathbf{U}_t b(t, \mathbf{U}_t^\top \mathbf{Y}_t). \quad (2.14)$$

Moreover,  $\mathbf{U}_t$  is absolutely continuous and by assumption  $\mathbf{Y}_t$  is a.s. continuous. This implies that  $\alpha_t$  and  $\beta_t$  have a continuous path almost surely, and hence (2.7) follows. In turn, from (2.12) we find

$$\begin{aligned} C_{\mathbf{Y}_t} \dot{\mathbf{U}}_t &= \mathbb{E}[\mathbf{Y}_t a(t, \mathbf{U}_t^\top \mathbf{Y}_t)^\top] - \mathbb{E}[\mathbf{Y}_t a(t, \mathbf{U}_t^\top \mathbf{Y}_t)^\top] \mathbf{U}_t^\top \mathbf{U}_t \\ &= \mathbb{E}[\mathbf{Y}_t a(t, \mathbf{U}_t^\top \mathbf{Y}_t)^\top] (I_{d \times d} - \mathbf{U}_t^\top \mathbf{U}_t), \end{aligned}$$

with  $C_{\mathbf{Y}_t} := \mathbb{E}[\mathbf{Y}_t \mathbf{Y}_t^\top]$ . Hence, from the orthogonality assumption for the rows of  $\mathbf{U}$  we have

$$C_{\mathbf{Y}_t} \dot{\mathbf{U}}_t = \mathbb{E}[\mathbf{Y}_t a(t, \mathbf{U}_t^\top \mathbf{Y}_t)^\top] (I_{d \times d} - P_{\mathbf{U}_t}^{\text{row}}).$$

This completes our consistency argument.

## 2.2 DO equations interpreted as a projected dynamics

The DO equations (2.6)–(2.7) are posed as a system of equations for the separate factors  $\mathbf{U}$  and  $\mathbf{Y}$ . We now discuss what equation the DO approximation  $X = \sum_{j=1}^R U^j Y^j$  should satisfy. In other words, we aim to derive an equation for  $X$  in the ambient space, independent of the parametrization  $(\mathbf{U}, \mathbf{Y})$ .

For this purpose, we substitute (2.6) and (2.7) into (2.9); we obtain

$$\begin{aligned} dX_t &= ((I_{d \times d} - P_{\mathbf{U}_t}^{\text{row}}) \mathbb{E}[a(t, \mathbf{U}_t^\top \mathbf{Y}_t) \mathbf{Y}_t^\top] C_{\mathbf{Y}_t}^{-1} \mathbf{Y}_t + P_{\mathbf{U}_t}^{\text{row}} a(t, \mathbf{U}_t^\top \mathbf{Y}_t)) dt \\ &\quad + P_{\mathbf{U}_t}^{\text{row}} b(t, \mathbf{U}_t^\top \mathbf{Y}_t) dW_t \\ &= ((I_{d \times d} - P_{\mathbf{U}_t}^{\text{row}}) [P_{\mathbf{Y}_t} a(t, \mathbf{U}_t^\top \mathbf{Y}_t)] + P_{\mathbf{U}_t}^{\text{row}} a(t, \mathbf{U}_t^\top \mathbf{Y}_t)) dt \\ &\quad + P_{\mathbf{U}_t}^{\text{row}} b(t, \mathbf{U}_t^\top \mathbf{Y}_t) dW_t, \end{aligned} \tag{2.15}$$

where  $P_{\mathbf{Y}_t} a(t, X_t) \in \mathbb{R}^d$  is the application of the  $L^2(\Omega)$ -orthogonal projection  $P_{\mathbf{Y}_t}: L^2(\Omega) \rightarrow \text{span}\{Y_t^1, \dots, Y_t^R\}$  to each component of  $a(t, X_t) \in \mathbb{R}^d$ . To derive a parameter-independent equation, we seek a parameter-free expression of the projections  $P_{\mathbf{U}_t}$  and  $P_{\mathbf{Y}_t}$ . Given  $\mathcal{X}_t \in L^2(\Omega; \mathbb{R}^d)$ ,  $t \in [0, T]$ , let  $\mathcal{P}_{\mathbf{U}(\mathcal{X}_t)}: \mathbb{R}^d \rightarrow \text{Im}(\mathbb{E}[\mathcal{X}_t \cdot]) \subset \mathbb{R}^d$  be the orthogonal projection matrix to the image of the mapping  $\mathbb{E}[\mathcal{X}_t \cdot]: L^2(\Omega) \rightarrow \mathbb{R}^d$ , and let  $\mathcal{P}_{\mathbf{Y}(\mathcal{X}_t)}: L^2(\Omega) \rightarrow \text{Im}((\mathcal{X}_t^\top \cdot)) \subset L^2(\Omega)$  be the  $L^2(\Omega)$ -orthogonal projection to the image of the mapping  $\mathcal{X}_t^\top: \mathbb{R}^d \rightarrow L^2(\Omega)$ . In Lemma 2.3 we will show  $\dim(\text{Im}((\mathcal{X}_t^\top \cdot))) = \dim(\text{Im}(\mathbb{E}[\mathcal{X}_t \cdot]))$ , i.e.,  $\text{Im}((\mathcal{X}_t^\top \cdot))$  is a finite-dimensional linear subspace of  $L^2(\Omega)$ , and thus  $\mathcal{P}_{\mathbf{Y}(\mathcal{X}_t)}$  is well defined.

In the case of DO approximation  $X_t = \mathbf{U}_t^\top \mathbf{Y}_t$ , we have  $\text{Im}(\mathbb{E}[X_t \cdot]) = \text{span}\{U_t^1, \dots, U_t^R\}$  and  $\text{Im}((X_t^\top \cdot)) = \text{span}\{Y_t^1, \dots, Y_t^R\}$ , we conclude that  $\mathcal{P}_{\mathbf{U}(X_t)} v = \mathbf{U}_t^\top \mathbf{U}_t v = P_{\mathbf{U}_t}^{\text{row}} v$ ,  $v \in \mathbb{R}^d$  and  $\mathcal{P}_{\mathbf{Y}(X_t)} w = \mathbb{E}[w \mathbf{Y}_t^\top] C_{\mathbf{Y}_t}^{-1} \mathbf{Y}_t = P_{\mathbf{Y}_t} w$ ,  $w \in L^2(\Omega)$ , so that (2.15) can be rewritten as

$$\begin{aligned} dX_t &= ((I_{d \times d} - \mathcal{P}_{\mathbf{U}(X_t)}) [\mathcal{P}_{\mathbf{Y}(X_t)} a(t, \mathbf{U}_t^\top \mathbf{Y}_t)] + \mathcal{P}_{\mathbf{U}(X_t)} a(t, \mathbf{U}_t^\top \mathbf{Y}_t)) dt \\ &\quad + \mathcal{P}_{\mathbf{U}(X_t)} b(t, \mathbf{U}_t^\top \mathbf{Y}_t) dW_t. \end{aligned} \tag{2.16}$$

This equation, derived from (2.6) and (2.7), does not depend on the parametrization of the pair  $(\mathbf{U}, \mathbf{Y})$  and could be taken as an alternative definition of DO approximation. More precisely, given any process  $\mathcal{X} = (\mathcal{X}_t)_{t \in [0, T]}$  with  $\mathcal{X}_t \in L^2(\Omega; \mathbb{R}^d)$  for  $t \in [0, T]$ , we can define the following stochastic process:

$$d\mathcal{X}_t = ((I_{d \times d} - \mathcal{P}_{\mathbf{U}(\mathcal{X}_t)}) [\mathcal{P}_{\mathbf{Y}(\mathcal{X}_t)} a(t, \mathcal{X}_t)] + \mathcal{P}_{\mathbf{U}(\mathcal{X}_t)} a(t, \mathcal{X}_t)) dt + \mathcal{P}_{\mathbf{U}(\mathcal{X}_t)} b(t, \mathcal{X}_t) dW_t. \tag{2.17}$$

It is worth noticing that (2.17) is a McKean-Vlasov-type SDE since the evolution of  $\mathcal{X}_t$  depends on the law of the process.

Finally, to speak of the *rank* of the solution  $\mathcal{X}$  to (2.17) we note the following.

Given  $\tilde{\mathcal{X}} \in L^2(\Omega; \mathbb{R}^d)$  the mapping  $K_{\tilde{\mathcal{X}}}: \mathbb{R}^d \rightarrow L^2(\Omega)$  defined by  $K_{\tilde{\mathcal{X}}} \mathbf{v} := \tilde{\mathcal{X}}^\top \mathbf{v}$  for  $\mathbf{v} \in \mathbb{R}^d$  is by definition finite rank. Moreover, the operator  $K_{\tilde{\mathcal{X}}}^* := \mathbb{E}[\tilde{\mathcal{X}} \cdot]: L^2(\Omega) \rightarrow \mathbb{R}^d$  is the adjoint of  $K_{\tilde{\mathcal{X}}}$ :

$$\mathbb{E}[y(K_{\tilde{\mathcal{X}}} \mathbf{v})] = \mathbb{E}[y(\tilde{\mathcal{X}}^\top \mathbf{v})] = \mathbb{E}[\mathbf{v}^\top \tilde{\mathcal{X}} y] = \mathbf{v}^\top K_{\tilde{\mathcal{X}}}^* y, \quad \forall y \in L^2(\Omega), \forall \mathbf{v} \in \mathbb{R}^d$$

The operator  $K_{\tilde{\mathcal{X}}}^* K_{\tilde{\mathcal{X}}}: \mathbb{R}^d \rightarrow \mathbb{R}^d$  is given by  $K_{\tilde{\mathcal{X}}}^* K_{\tilde{\mathcal{X}}} \mathbf{v} = \mathbb{E}[\tilde{\mathcal{X}} \tilde{\mathcal{X}}^\top] \mathbf{v}$  for  $\mathbf{v} \in \mathbb{R}^d$ . The following lemma characterises the rank of these operators.

**Lemma 2.3.** *Given  $\tilde{\mathcal{X}} \in L^2(\Omega; \mathbb{R}^d)$ , we have  $\text{rank}(K_{\tilde{\mathcal{X}}}^* K_{\tilde{\mathcal{X}}}) = \text{rank}(K_{\tilde{\mathcal{X}}}) = \text{rank}(K_{\tilde{\mathcal{X}}}^*)$ .*

*Proof.* First, notice that we have  $\ker(K_{\tilde{\mathcal{X}}}^*) = \text{Im}(K_{\tilde{\mathcal{X}}})^\perp$ , where  $\perp$  is the orthogonal complement with respect to the Euclidean inner product. Hence, we have  $\ker(K_{\tilde{\mathcal{X}}}^*|_{\text{Im}(K_{\tilde{\mathcal{X}}})}) = \{0\}$ . Thus, the rank-nullity theorem implies

$$\dim(\text{Im}(K_{\tilde{\mathcal{X}}}^*|_{\text{Im}(K_{\tilde{\mathcal{X}}})})) = \dim(\text{Im}(K_{\tilde{\mathcal{X}}})).$$

Therefore  $\text{rank}(K_{\tilde{\mathcal{X}}}^* K_{\tilde{\mathcal{X}}}) = \text{rank}(K_{\tilde{\mathcal{X}}})$ . Moreover, from  $\text{rank}(K_{\tilde{\mathcal{X}}}) = \text{rank}(K_{\tilde{\mathcal{X}}}^*)$  (see for example [15, Theorem III.4.13]), the proof is complete.  $\square$

In view of the lemma above, we call  $\dim(\text{Im}(\mathbb{E}[\mathcal{X}_t \cdot]))$ , equivalently  $\dim(\text{Im}((\mathcal{X}_t^\top \cdot)))$  and  $\text{rank}(\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top])$ , rank of  $\mathcal{X}_t$ .

**Definition 2.4** (DLR solution of rank  $R$ ). A process  $\mathcal{X} : [0, T] \rightarrow L^2(\Omega; \mathbb{R}^d)$  is called a *DLR solution of rank  $R$*  to (2.1) for an initial datum  $\mathcal{X}_0 \in L^2(\Omega; \mathbb{R}^d)$  if it satisfies (2.17),  $\dim(\text{Im}(\mathbb{E}[\mathcal{X}_t \cdot])) = R$  for some  $R \in \mathbb{N}$  for all  $t \in [0, T]$ , and  $\mathcal{X}$  has almost surely continuous paths.

The parameter-independent formulation (2.17) corresponds to the projected dynamics in the DLRA literature; see [20], also [17, Proposition2]. Notice however that in our formulation, only the projector  $\mathcal{P}_{\mathcal{U}(\mathcal{X}_t)}$  is applied to the diffusion term  $b(t, \mathcal{X}_t) dW_t$ , instead of the full projector

$$\mathcal{P}_{\mathcal{U}(\mathcal{X}_t)} + \mathcal{P}_{\mathcal{Y}(\mathcal{X}_t)} - \mathcal{P}_{\mathcal{U}(\mathcal{X}_t)} \mathcal{P}_{\mathcal{Y}(\mathcal{X}_t)}.$$

If we naively wrote the projected dynamics following the standard DLRA approach, we would end up with the alternative formal expression

$$d\mathcal{X}_t = (\mathcal{P}_{\mathcal{U}(\mathcal{X}_t)} + \mathcal{P}_{\mathcal{Y}(\mathcal{X}_t)} - \mathcal{P}_{\mathcal{U}(\mathcal{X}_t)} \mathcal{P}_{\mathcal{Y}(\mathcal{X}_t)}) [a(t, \mathcal{X}_t) dt + b(t, \mathcal{X}_t) dW_t], \quad (2.18)$$

which would coincide with (2.17) if one could prove that “ $\mathcal{P}_{\mathcal{Y}(\mathcal{X}_t)}[b(t, \mathcal{X}_t) dW_t] = 0$ ” for a.e.  $t > 0$ . However, it is not obvious how to give a rigorous meaning to the term  $\mathcal{P}_{\mathcal{Y}(\mathcal{X}_t)}[b(t, \mathcal{X}_t) dW_t]$  as an Itô integral and for this reason we do not pursue the formulation (2.18) further.

To confirm that (2.17) is nevertheless a sensible projected dynamics, let us ask ourselves the following question: if the solution  $\mathcal{X}$  of (2.17) satisfies

$$\dim(\text{Im}((\mathcal{X}_t^\top \cdot))) = \dim(\text{Im}(\mathbb{E}[\mathcal{X}_t \cdot])) = R \text{ around } t,$$

is the right hand side of (2.17) consistent with this structure of  $\mathcal{X}$ ? In the following, we will see that the answer is affirmative, which supports the validity of (2.17), and thus (2.6) and (2.7), as a correct DLRA formulation for SDEs.

In view of Lemma 2.3, we study the rank of the matrix  $\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top]$ , where  $\mathcal{X}_t$  is given by (2.17). To do this, we will show that  $\frac{d}{dt}(\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top])$  is in a tangent space of the rank- $R$  manifold, i.e., the manifold of  $d \times d$  matrices with rank equal to  $R$ , at  $\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top]$ . For  $j, k = 1, \dots, d$ , Itô’s formula implies

$$\begin{aligned} d(\mathcal{X}_t^j \mathcal{X}_t^k) &= \mathcal{X}_t^j [\mu_t^k dt + \sum_{\ell=1}^m \Sigma_t^{k\ell} dW_t^\ell] \\ &\quad + [\mu_t^j dt + \sum_{\ell=1}^m \Sigma_t^{j\ell} dW_t^\ell] \mathcal{X}_t^k + \sum_{\ell=1}^m \Sigma_t^{j\ell} \Sigma_t^{k\ell} dt, \end{aligned}$$



where  $\mu_t^k := [(I_{d \times d} - \mathcal{P}_{\mathcal{U}(\mathcal{X}_t)})[\mathcal{P}_{\mathcal{Y}(\mathcal{X}_t)}a(t, \mathcal{X}_t)] + \mathcal{P}_{\mathcal{U}(\mathcal{X}_t)}a(t, \mathcal{X}_t)]_k$  and  $\Sigma_t^{k\ell} := [\mathcal{P}_{\mathcal{U}(\mathcal{X}_t)}b(t, \mathcal{X}_t)]_{k\ell}$ . Thus, we have

$$\frac{d}{dt} \left( \mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top] \right) = \mathbb{E}[\mathcal{X}_t \mu_t^\top + \mu_t \mathcal{X}_t^\top + \Sigma_t \Sigma_t^\top] \quad \text{for almost every } t \in [0, T], \quad (2.19)$$

with  $\mu_t := [\mu_t^k]_{k=1, \dots, d} \in \mathbb{R}^d$  and  $\Sigma_t := [\Sigma_t^{k\ell}]_{\substack{k=1, \dots, d \\ \ell=1, \dots, m}} \in \mathbb{R}^{d \times m}$ .

On the other hand, from [41, Proposition 2.1] we know that the tangent space of a rank  $R$  manifold at  $\mathbb{E}[\mathcal{X} \mathcal{X}^\top] = Q \text{diag}(\gamma_1, \dots, \gamma_R) Q^\top$  can be characterized as

$$T_{\mathbb{E}[\mathcal{X} \mathcal{X}^\top]} \mathcal{M} := \left\{ QV_1^\top + V_2 Q^\top + QAQ^\top \left| \begin{array}{l} V_1^\top Q = 0, V_1 \in \mathbb{R}^{d \times R}, \\ V_2^\top Q = 0, V_2 \in \mathbb{R}^{d \times R}, \\ \text{and } A \in \mathbb{R}^{d \times d} \end{array} \right. \right\}.$$

To conclude  $\frac{d}{dt}(\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top]) \in T_{\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top]} \mathcal{M}$  we will use the following result.

**Lemma 2.5.** *Let the singular value decomposition  $\mathbb{E}[\mathcal{X} \mathcal{X}^\top] = Q \text{diag}(\gamma_1, \dots, \gamma_R) Q^\top$ , with  $\gamma_1, \dots, \gamma_R > 0$  be given, where  $Q \in \mathbb{R}^{d \times R}$  is a matrix consisting of  $R$  orthogonal columns  $\mathbf{q}_1, \dots, \mathbf{q}_R \in \mathbb{R}^d$ . Then, the canonical expansion of the finite rank operator  $\mathbb{E}[\mathcal{X} \cdot]: L^2(\Omega) \rightarrow \mathbb{R}^d$  is given by*

$$\mathbb{E}[\mathcal{X}y] = \sum_{k=1}^R \gamma_k^{1/2} \mathbb{E}[y \varphi_k] \mathbf{q}_k = Q \begin{bmatrix} \gamma_1^{1/2} \mathbb{E}[y \varphi_1] \\ \vdots \\ \gamma_R^{1/2} \mathbb{E}[y \varphi_R] \end{bmatrix}, \quad \text{for } y \in L^2(\Omega),$$

where  $\varphi_k := \gamma_k^{-1/2} \mathcal{X}^\top \mathbf{q}_k$ ,  $k = 1, \dots, R$  is an orthonormal basis of  $\text{Im}((\mathcal{X}^\top \cdot)) \subset L^2(\Omega)$ .

*Proof.* From

$$\mathbb{E}[\mathcal{X} \mathcal{X}^\top \gamma_k^{-1} \mathbf{q}_k] = \gamma_k^{-1} Q \text{diag}(\gamma_1, \dots, \gamma_R) Q^\top \mathbf{q}_k = \mathbf{q}_k,$$

we have  $\{\mathbf{q}_k\}_{k=1}^R \subset \text{Im}(\mathbb{E}[\mathcal{X} \cdot])$ , and thus  $\{\mathbf{q}_k\}_{k=1}^R$  is an orthonormal basis of  $\text{Im}(\mathbb{E}[\mathcal{X} \cdot]) \subset \mathbb{R}^d$ .

Thus, with some coefficients  $\{c_k\}_{k=1}^R \subset \mathbb{R}$  we have for any  $y \in L^2(\Omega)$  a representation  $\sum_{k=1}^R c_k \mathbf{q}_k = \mathbb{E}[\mathcal{X}y]$ , which implies

$$c_k = \mathbf{q}_k^\top \mathbb{E}[\mathcal{X}y] = \mathbb{E}[\mathcal{X}^\top \mathbf{q}_k y] = \gamma_k^{1/2} \mathbb{E}[\gamma_k^{-1/2} \mathcal{X}^\top \mathbf{q}_k y] = \gamma_k^{1/2} \mathbb{E}[\varphi_k y].$$

The functions  $\varphi_k = \gamma_k^{-1/2} \mathcal{X}^\top \mathbf{q}_k$ ,  $k = 1, \dots, R$  are orthonormal in  $L^2(\Omega)$ , and thus form an orthonormal basis in the  $R$ -dimensional subspace  $\text{Im}((\mathcal{X}^\top \cdot))$ .  $\square$

**Remark 2.6.** Lemmata 2.3 and 2.5 imply  $\text{rank}(\mathcal{X}) = \dim(\text{span}(\mathcal{X}))$ , where

$$\text{span}(\mathcal{X}) := \text{span}(\{\mathcal{X}(\omega) \in \mathbb{R}^d \mid \omega \in \Omega\}) = \left\{ \sum_{j=1}^n c_j \mathcal{X}(\omega_j) \in \mathbb{R}^d \mid n \in \mathbb{N}, c_j \in \mathbb{R}, \omega_j \in \Omega \right\} \subset \mathbb{R}^d,$$

i.e. the rank of  $\mathcal{X} \in L^2(\Omega; \mathbb{R}^d)$  as in Definition 2.4 gives the dimension of the space spanned by all the realizations of  $\mathcal{X}$ . To see this, take  $\mathcal{X} \in L^2(\Omega; \mathbb{R}^d)$  such that  $\dim(\text{Im}(\mathbb{E}[\mathcal{X} \cdot])) = R$ . Lemma 2.3 implies that  $\mathbb{E}[\mathcal{X} \mathcal{X}^\top] = Q \text{diag}(\gamma_1, \dots, \gamma_R) Q^\top$  holds, where  $\gamma_1, \dots, \gamma_R > 0$  and  $Q \in \mathbb{R}^{d \times R}$  is a matrix consisting of  $R$  orthogonal columns

## Dynamical low-rank approximation for stochastic differential equations

$\mathbf{q}_1, \dots, \mathbf{q}_R \in \mathbb{R}^d$ , while Lemma 2.5 tells us that  $\varphi_k := \gamma_k^{-1/2} \mathcal{X}^\top \mathbf{q}_k$ , for  $k = 1, \dots, R$ , is an orthonormal basis of  $\text{Im}((\mathcal{X}^\top \cdot)) \subset L^2(\Omega)$ . Therefore,  $\mathcal{X}$  admits the expansion

$$\mathcal{X}(\omega) = \sum_{k=1}^R \mathbb{E}[\mathcal{X} \varphi_k] \varphi_k(\omega), \quad (2.20)$$

where  $\mathbb{E}[\mathcal{X} \varphi_k] := (\mathbb{E}[\mathcal{X}_\ell \varphi_k])_{\ell=1, \dots, d}$  is the component-wise  $L^2(\Omega)$ -inner product. From  $\varphi_k = \gamma_k^{-1/2} \mathbf{q}_k^\top \mathcal{X}$  we have

$$\mathbb{E}[\mathcal{X} \varphi_k] = \gamma_k^{-1/2} \mathbb{E}[\mathcal{X} \mathcal{X}^\top \mathbf{q}_k] = \gamma_k^{-1/2} Q \text{diag}(\gamma_1, \dots, \gamma_R) Q^\top \mathbf{q}_k = \gamma_k^{1/2} \mathbf{q}_k. \quad (2.21)$$

By substituting (2.21) into (2.20), we obtain  $\mathcal{X}(\omega) = \sum_{k=1}^R \gamma_k^{1/2} \varphi_k(\omega) \mathbf{q}_k$  for any  $\omega \in \Omega$ . This implies that  $\sum_{j=1}^n c_j \mathcal{X}(\omega_j) = \sum_{k=1}^R \gamma_k^{1/2} c_j \varphi_k(\omega_j) \mathbf{q}_k$  for any  $n \in \mathbb{N}$ ,  $c_j \in \mathbb{R}$ , and  $\omega_j \in \Omega$ , i.e. the space spanned by all the realizations of  $\mathcal{X}(\omega)$  is spanned by  $\mathbf{q}_1, \dots, \mathbf{q}_R$ . Therefore,  $\dim(\text{span}(\mathcal{X})) = R$ .

The equation (2.19) together with Lemma 2.5 shows  $\frac{d}{dt}(\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top]) \in T_{\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top]} \mathcal{M}$ . Indeed, the first term of the right hand side of (2.19) can be written as

$$\begin{aligned} \mathbb{E}[\mathcal{X}_t \mu_t^\top] &= \mathbb{E}[\mathcal{X}_t (\mathcal{P}_{\mathcal{Y}(\mathcal{X}_t)} a(t, \mathcal{X}_t))^\top] (I_{d \times d} - \mathcal{P}_{\mathcal{U}(\mathcal{X}_t)}) + \mathbb{E}[\mathcal{X}_t a(t, \mathcal{X}_t)^\top] \mathcal{P}_{\mathcal{U}(\mathcal{X}_t)} \\ &= Q_t B_t (I_{d \times d} - Q_t Q_t^\top) + Q_t C_t Q_t Q_t^\top, \end{aligned}$$

where, using Lemma 2.5, we may write  $Q_t B_t := \mathbb{E}[\mathcal{X}_t (\mathcal{P}_{\mathcal{Y}(\mathcal{X}_t)} a(t, \mathcal{X}_t))^\top]$  and  $Q_t C_t := \mathbb{E}[\mathcal{X}_t a(t, \mathcal{X}_t)^\top]$  with some  $B_t, C_t \in \mathbb{R}^{d \times d}$ . Then, with  $\tilde{V}_t := (B_t (I_{d \times d} - Q_t Q_t^\top))^\top$  and  $\tilde{A}_t := C_t Q_t$  we have

$$\mathbb{E}[\mathcal{X}_t \mu_t^\top] = Q_t \tilde{V}_t^\top + Q_t \tilde{A}_t Q_t^\top.$$

Similarly, we have  $\mathbb{E}[\mu_t \mathcal{X}_t^\top] = \tilde{V}_t Q_t^\top + Q_t \tilde{A}_t^\top Q_t^\top$  and  $\mathbb{E}[\Sigma_t \Sigma_t^\top] = Q_t Q_t^\top \mathbb{E}[b(t, \mathcal{X}_t) b(t, \mathcal{X}_t)^\top] Q_t Q_t^\top$ . Hence, if  $\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top]$  is of rank  $R$  around  $t$ , then its derivative (2.19), which was derived from the projected dynamics (2.17), lies indeed in  $T_{\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top]} \mathcal{M}$ . This consistency supports the validity of the formulation (2.17), and in turn that of the DO equations (2.6) and (2.7).

### 2.3 Equivalence of DO and DLR formulations

In the previous section, we showed that if there exists a strong DO solution  $(\mathbf{U}_t, \mathbf{Y}_t)$  to (2.1), then the corresponding DO approximation  $X_t = \mathbf{U}_t^\top \mathbf{Y}_t$  satisfies (2.17). We now investigate the reverse question: if  $\mathcal{X}_t$  is a rank- $R$  solution of (2.17) (DLR solution of rank  $R$  of (2.1)) does there exist a DO solution  $(\mathbf{U}, \mathbf{Y})$  such that  $\mathcal{X} = \mathbf{U}^\top \mathbf{Y}$ ?

First, we need the following bound for the DLR solution.

**Lemma 2.7.** *Let a rank- $R$  DLR solution  $\mathcal{X}_t$ ,  $t \in [0, T]$ , to (2.1) with  $\mathcal{X}_0 \in L^2(\Omega; \mathbb{R}^d)$  be given. For all  $t \in [0, T]$ ,  $\mathcal{X}_t$  satisfies*

$$\mathbb{E}[|\mathcal{X}_t|^2] \leq 3(\mathbb{E}[|\mathcal{X}_0|^2] + C_{\text{lgb}} T(T+1)) \exp(3C_{\text{lgb}} T(T+1)) =: M(T) \quad (2.22)$$

*Proof.* Taking the squared  $L^2(\Omega)$ -norm of  $\mathcal{X}_t$  and using (2.17) in integral form, together with Itô's isometry, Jensen's inequality, and Assumption 2, we have

$$\begin{aligned}
 \mathbb{E}[|\mathcal{X}_t|^2] &= \mathbb{E}\left[|\mathcal{X}_0 + \int_0^t ((I_{d \times d} - \mathcal{P}_{U(\mathcal{X}_s)})[\mathcal{P}_{Y(\mathcal{X}_s)}a(s, \mathcal{X}_s)] + \mathcal{P}_{U(\mathcal{X}_s)}a(s, \mathcal{X}_s)) ds \right. \\
 &\quad \left. + \int_0^t \mathcal{P}_{U(\mathcal{X}_s)}b(s, \mathcal{X}_s) dW_s|^2\right] \\
 &\leq 3\mathbb{E}[|\mathcal{X}_0|^2] + 3\mathbb{E}\left[\left|\int_0^t ((I_{d \times d} - \mathcal{P}_{U(\mathcal{X}_s)})\mathcal{P}_{Y(\mathcal{X}_s)} + \mathcal{P}_{U(\mathcal{X}_s)})a(s, \mathcal{X}_s) ds\right|^2\right] \\
 &\quad + 3\mathbb{E}\left[\left|\int_0^t \mathcal{P}_{U(\mathcal{X}_s)}b(s, \mathcal{X}_s) dW_s\right|^2\right] \\
 &\leq 3\mathbb{E}[|\mathcal{X}_0|^2] + 3T \int_0^t \mathbb{E}\left[\left|\left((I_{d \times d} - \mathcal{P}_{U(\mathcal{X}_s)})\mathcal{P}_{Y(\mathcal{X}_s)} + \mathcal{P}_{U(\mathcal{X}_s)})a(s, \mathcal{X}_s)\right|^2\right] ds \\
 &\quad + 3 \int_0^t \mathbb{E}\left[\left|\mathcal{P}_{U(\mathcal{X}_s)}b(s, \mathcal{X}_s)\right|^2\right] ds \\
 &\leq 3\mathbb{E}[|\mathcal{X}_0|^2] + 3C_{\text{lgb}}(T+1) \int_0^t (1 + \mathbb{E}[|\mathcal{X}_s|^2]) ds.
 \end{aligned}$$

Hence, Gronwall's lemma implies the statement.  $\square$

The following theorem gives the uniqueness of the DLR solution.

**Theorem 2.8.** *Let  $\mathcal{X}_0 \in L^2(\Omega; \mathbb{R}^d)$  be such that  $\dim(\text{Im}(\mathbb{E}[\mathcal{X}_0 \cdot])) = R$ . Suppose that two DLR solutions  $\mathcal{X}_t$  and  $\mathcal{Z}_t$  of rank  $R$  to (2.1) with initial datum  $\mathcal{X}_0$  exist on  $[0, T]$ . Then  $\mathcal{X}$  and  $\mathcal{Z}$  are indistinguishable.*

*Proof.* The processes  $\mathcal{X}_t$  and  $\mathcal{Z}_t$  are assumed to satisfy

$$d\mathcal{X}_t = \underbrace{\left( (I_{d \times d} - \mathcal{P}_{U(\mathcal{X}_t)})[\mathcal{P}_{Y(\mathcal{X}_t)}a(t, \mathcal{X}_t)] + \mathcal{P}_{U(\mathcal{X}_t)}a(t, \mathcal{X}_t) \right)}_{=: \mathfrak{P}_{\mathcal{X}_t}a(t, \mathcal{X}_t)} dt + \mathcal{P}_{U(\mathcal{X}_t)}b(t, \mathcal{X}_t) dW_t$$

and

$$d\mathcal{Z}_t = \underbrace{\left( (I_{d \times d} - \mathcal{P}_{U(\mathcal{Z}_t)})[\mathcal{P}_{Y(\mathcal{Z}_t)}a(t, \mathcal{Z}_t)] + \mathcal{P}_{U(\mathcal{Z}_t)}a(t, \mathcal{Z}_t) \right)}_{=: \mathfrak{P}_{\mathcal{Z}_t}a(t, \mathcal{Z}_t)} dt + \mathcal{P}_{U(\mathcal{Z}_t)}b(t, \mathcal{Z}_t) dW_t,$$

respectively, for the same initial datum  $\mathcal{X}_0$ . This implies that

$$\begin{aligned}
 \mathbb{E} \left[ |\mathcal{X}_t - \mathcal{Z}_t|^2 \right] &\leq \mathbb{E} \left[ \left| \int_0^t ((I_{d \times d} - \mathcal{P}_{U(\mathcal{X}_s)}) \mathcal{P}_{Y(\mathcal{X}_s)} + \mathcal{P}_{U(\mathcal{X}_s)}) a(s, \mathcal{X}_s) \right. \right. \\
 &\quad \left. \left. - ((I_{d \times d} - \mathcal{P}_{U(\mathcal{Z}_s)}) \mathcal{P}_{Y(\mathcal{Z}_s)} + \mathcal{P}_{U(\mathcal{Z}_s)}) a(s, \mathcal{Z}_s) ds \right. \right. \\
 &\quad \left. \left. + \int_0^t \mathcal{P}_{U(\mathcal{X}_s)} b(s, \mathcal{X}_s) - \mathcal{P}_{U(\mathcal{Z}_s)} b(s, \mathcal{Z}_s) dW_s \right|^2 \right] \\
 &\leq 4\mathbb{E} \left[ \left| \int_0^t (\mathfrak{P}_{\mathcal{X}_s} - \mathfrak{P}_{\mathcal{Z}_s}) a(s, \mathcal{X}_s) ds \right|^2 \right] + 4\mathbb{E} \left[ \left| \int_0^t \mathfrak{P}_{\mathcal{Z}_s} (a(s, \mathcal{Z}_s) - a(s, \mathcal{X}_s)) ds \right|^2 \right] \\
 &\quad + 4\mathbb{E} \left[ \left| \int_0^t (\mathcal{P}_{U(\mathcal{X}_s)} - \mathcal{P}_{U(\mathcal{Z}_s)}) b(s, \mathcal{X}_s) dW_s \right|^2 \right] \\
 &\quad + 4\mathbb{E} \left[ \left| \int_0^t \mathcal{P}_{U(\mathcal{Z}_s)} (b(s, \mathcal{X}_s) - b(s, \mathcal{Z}_s)) dW_s \right|^2 \right]
 \end{aligned} \tag{2.23}$$

Denote by

$$\tilde{\gamma} := \inf_{t \in [0, T]} \max \{ \sigma_R(\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top]), \sigma_R(\mathbb{E}[\mathcal{Z}_t \mathcal{Z}_t^\top]) \}$$

the infimum over time of the maximum between the  $R$ -th singular value of the matrices  $\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top] \in \mathbb{R}^{R \times R}$  and  $\mathbb{E}[\mathcal{Z}_t \mathcal{Z}_t^\top] \in \mathbb{R}^{R \times R}$ , where for all  $t$  we are considering the singular values in a decreasing order. We have  $\tilde{\gamma} > 0$ . To see this, first note that, from Lemma 2.3 the rank of  $\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top]$  is the same as the rank of  $\mathcal{X}_t$ . But from the definition of DLR solution,  $\mathcal{X}_t$  has always rank  $R$ , and moreover, the continuity of  $\mathcal{X}_t$  implies continuity of  $\sigma_R(\mathbb{E}[\mathcal{X}_t \mathcal{X}_t^\top])$  on  $[0, T]$ . Hence  $\tilde{\gamma} > 0$  follows; see also [16, Lemma 2.1]. Then, Proposition A.3 gives us the following Lipschitz bounds for the projectors for any  $s \in [0, T]$ :

$$\begin{aligned}
 \mathbb{E}[|(\mathfrak{P}_{\mathcal{X}_s} - \mathfrak{P}_{\mathcal{Z}_s})v|^2] &\leq \left( \frac{2}{\tilde{\gamma}} \right)^2 \mathbb{E}[|\mathcal{X}_s - \mathcal{Z}_s|^2] \|v\|_{L^2(\Omega; \mathbb{R}^d)}^2 \quad \text{for any } v \in L^2(\Omega; \mathbb{R}^d), \\
 \|\mathcal{P}_{U(\mathcal{X}_s)} - \mathcal{P}_{U(\mathcal{Z}_s)}\|_2^2 &\leq \frac{\mathbb{E}[|\mathcal{X}_s - \mathcal{Z}_s|^2]}{\tilde{\gamma}^2}.
 \end{aligned} \tag{2.24}$$

Hence (2.23), Itô's isometry and Jensen's inequality lead to

$$\begin{aligned}
 \mathbb{E}[|\mathcal{X}_t - \mathcal{Z}_t|^2] &\leq 4T \int_0^t \mathbb{E}[|(\mathfrak{P}_{\mathcal{X}_s} - \mathfrak{P}_{\mathcal{Z}_s}) a(s, \mathcal{X}_s)|^2] ds \\
 &\quad + 4T \int_0^t \mathbb{E}[|\mathfrak{P}_{\mathcal{Z}_s} (a(s, \mathcal{Z}_s) - a(s, \mathcal{X}_s))|^2] ds \\
 &\quad + 4 \int_0^t \mathbb{E}[|(\mathcal{P}_{U(\mathcal{X}_s)} - \mathcal{P}_{U(\mathcal{Z}_s)}) b(s, \mathcal{X}_s)|^2] ds \\
 &\quad + 4 \int_0^t \mathbb{E}[|\mathcal{P}_{U(\mathcal{Z}_s)} (b(s, \mathcal{X}_s) - b(s, \mathcal{Z}_s))|^2] ds \\
 &\leq \bar{C} \int_0^t \mathbb{E}[|\mathcal{X}_s - \mathcal{Z}_s|^2] ds,
 \end{aligned} \tag{2.25}$$

for all  $t$  in  $[0, T]$  with  $\bar{C} := 4 \left( (4T + 1) \tilde{\gamma}^{-2} C_{\text{lgb}}(1 + M(T)) + C_{\text{Lip}}^2(T + 1) \right)$  and  $M(T)$  as in (2.22). Gronwall's lemma yields  $\mathbb{E}[|\mathcal{X}_t - \mathcal{Z}_t|^2] = 0$  for all  $t \in [0, T]$  and therefore  $(\mathcal{Z}_t)_{t \in [0, T]}$  and  $(\mathcal{X}_t)_{t \in [0, T]}$  are versions of each other. Moreover, because they both have a.s. continuous paths, they are indistinguishable; see for example [31, Theorem 2].  $\square$

Given a strong DO solution  $(\mathbf{U}, \mathbf{Y})$  to (2.1), the corresponding DO approximation  $X_t$  satisfies the equation (2.16), and thus (2.17). Hence, Theorem 2.8 gives the sought equivalence between the DLR equation (2.17) and the DO equation (2.6)–(2.7), which we state as a corollary.

**Corollary 2.9.** *Let  $\mathcal{X}_0 \in L^2(\Omega; \mathbb{R}^d)$  with  $\dim(\text{Im}(\mathbb{E}[\mathcal{X}_0 \cdot])) = R$ . Suppose that a DLR solution  $\mathcal{X}_t$  of rank  $R$  to (2.1) and a strong DO solution  $(\mathbf{U}, \mathbf{Y})$  exist on  $[0, T]$ , both with initial datum  $\mathcal{X}_0$ . Then,  $\mathcal{X}_t$  and the DO approximation  $X_t = \mathbf{U}_t^\top \mathbf{Y}_t$  are indistinguishable on  $[0, T]$ .*

Moreover, the DO solution giving the same DLR solution is unique up to a rotation matrix. See also [16, Section 2.2] for a similar result.

**Proposition 2.10.** *Assume that a DLR solution  $\mathcal{X}_t$  of rank  $R$  to (2.1) with initial datum  $\mathcal{X}_0 \in [L^2(\Omega)]^d$  exists for all  $t \in [0, T]$ . Suppose there exist two strong DO solutions  $(\mathbf{U}_t, \mathbf{Y}_t)$  and  $(\mathbf{V}_t, \mathbf{Z}_t)$ . Then there exists a unique orthogonal matrix  $\Theta \in \mathbb{R}^{R \times R}$  such that  $(\mathbf{V}_t, \mathbf{Z}_t) = (\Theta \mathbf{U}_t, \Theta \mathbf{Y}_t)$ .*

*Proof.* The proof follows closely the arguments in [16, Lemma 2.3, Corollary 2.4, and Lemma 2.5]. We will show that there exists a unique absolutely continuous curve  $t \rightarrow \Theta(t)$  with orthogonal matrix  $\Theta(t) \in \mathbb{R}^{R \times R}$  for all  $t$  such that  $(\mathbf{V}_t, \mathbf{Z}_t) = (\Theta(t)\mathbf{U}_t, \Theta(t)\mathbf{Y}_t)$ , and that such  $\Theta(t)$  is a constant in  $t$ . First, we will derive an equation that  $\Theta(t)$  must satisfy. Suppose that sought  $\Theta(t)$  exists. Note that, since both  $\mathbf{U}_t^\top \mathbf{Y}_t$  and  $\mathbf{V}_t^\top \mathbf{Z}_t$  satisfy (2.16), Theorem 2.8 implies  $\mathcal{X}_t = \mathbf{U}_t^\top \mathbf{Y}_t = \mathbf{V}_t^\top \mathbf{Z}_t$  for all  $t$  and a.s. Then since  $(\mathbf{V}_t, \mathbf{Z}_t)$  is a strong DO solution, one must have

$$\dot{\mathbf{U}}_t = (\dot{\Theta}(t)\mathbf{V}_t) = \dot{\Theta}^\top(t)\mathbf{V}_t + \Theta^\top(t)\dot{\mathbf{V}}_t \quad \text{for a.e. } t \in [0, T]. \quad (2.26)$$

As  $\mathbf{U}_t^\top \dot{\mathbf{U}}_t^\top = 0$  and  $\mathbf{U}_t \mathbf{U}_t^\top = I_{R \times R}$ , from  $\mathbf{U}_t = \Theta^\top(t)\mathbf{V}_t$  it follows

$$0 = \Theta^\top(t)\mathbf{V}_t \left( \dot{\Theta}^\top(t)\mathbf{V}_t + \Theta^\top(t)\dot{\mathbf{V}}_t \right)^\top = \Theta^\top(t) \left( \dot{\Theta}(t) + \mathbf{V}_t \dot{\mathbf{V}}_t^\top \Theta(t) \right). \quad (2.27)$$

Using orthogonality of  $\Theta(t)$  and  $\mathbf{V}_t \dot{\mathbf{V}}_t^\top = 0$ , we obtain the following differential equation that  $\Theta(t)$  has to satisfy:

$$\dot{\Theta}(t) = 0 \quad \text{for a.e. } t \in [0, T] \text{ with } \Theta(0) = \Theta_*, \quad (2.28)$$

where  $\Theta_*$  is an orthogonal matrix. But this equation has a unique solution  $\Theta(t) \equiv \Theta_*$ , which is an orthogonal matrix. Hence, going back the argument above,  $\Theta$  satisfies (2.26). The absolute continuity of  $\mathbf{U}$  and  $\Theta \mathbf{V}$  yields  $\mathbf{U}_t = \Theta(t)\mathbf{V}_t + C$  for all  $t \in [0, T]$ , with some matrix  $C \in \mathbb{R}^{R \times R}$ .

Such  $\Theta_*$  that makes  $C = 0$  can be constructed explicitly. Since  $\mathbf{U}_t, \mathbf{V}_t$  are deterministic, from  $\mathbf{U}_0^\top \mathbf{Y}_0 = \mathbf{V}_0^\top \mathbf{Z}_0$  we must have

$$\mathbf{U}_0^\top = \mathbf{V}_0^\top \mathbb{E}[\mathbf{Z}_0 \mathbf{Y}_0^\top] C_{\mathbf{Y}_0}^{-1}.$$

Take  $\Theta_* = \mathbb{E}[\mathbf{Z}_0 \mathbf{Y}_0^T] C_{\mathbf{Y}_0}^{-1}$ . Then, from the argument above  $\mathbf{U}_t = \Theta \mathbf{V}_t$  holds for all  $t \in [0, T]$ . We conclude the proof by noting the orthogonality of  $\Theta_*$ :

$$I_{R \times R} = \mathbf{U}_0 \mathbf{U}_0^\top = \Theta_*^\top \mathbf{V}_0 \mathbf{V}_0^\top \Theta_* = \Theta_*^\top \Theta_*$$

□

**Remark 2.11.** In Section 3, under Assumptions 1–3 we will show that there exists a unique strong DO solution  $(\mathbf{U}, \mathbf{Y})$  in a certain interval  $[0, T]$ . Hence, in view of Theorem 2.8 and Corollary 2.9, a unique DLR solution  $\mathcal{X}_t$  exists on  $[0, T]$  and moreover, given  $\mathcal{X}_t$ , we can always find a strong DO solution  $(\mathbf{U}, \mathbf{Y})$  such that  $\mathcal{X}_t = \mathbf{U}_t^\top \mathbf{Y}_t$ .

### 3 Local existence and uniqueness

The equations (2.6) and (2.7) define a non-standard system of stochastic differential equations. Notice that in (2.6), the matrix-valued function  $[L^2(\Omega)]^R \ni \mathbf{Y}_t \mapsto C_{\mathbf{Y}_t}^{-1} \in \mathbb{R}^{R \times R}$  is not defined everywhere in  $[L^2(\Omega)]^R$ . Moreover, the vector field  $\mathbb{E}[\mathbf{Y}_t a(t, X_t)^\top]$  requires taking the expectation, and thus depends on the knowledge of all the paths of  $\mathbf{Y}_t$ . Hence, the vector fields that define the DO equation are not defined path-wise a priori. This setting makes the existence and uniqueness result of DLR solutions non-trivial.

To establish existence and uniqueness of solutions of (2.6) and (2.7) given an initial datum  $(\varphi, \xi)$ , we follow a fixed-point argument. We will define a sequence of Picard iterates, which belongs to a specific set of functions, where the first element of the sequence is made of a pair of constant-in-time functions  $(\varphi, \xi)$ . Then we will show that this sequence converges in this set.

Let  $\mathbf{U} \in C([0, t]; \mathbb{R}^{R \times d})$  be such that the rows  $U_s^1, \dots, U_s^R \in \mathbb{R}^d$  of  $\mathbf{U}_s$  are linearly independent for every  $s \in [0, t]$  (but not necessarily orthonormal), and let  $\mathbf{Y} \in L^2(\Omega; C([0, t]; \mathbb{R}^R))$  be an  $(\mathcal{F}_t)$ -adapted process such that  $\mathbf{Y}_s = (Y^1(s), \dots, Y^R(s)) \in [L^2(\Omega)]^R$  has linearly independent components for every  $s \in [0, t]$ . For  $\varphi \in \mathbb{R}^{R \times d}$  with  $d \geq R$  and  $\xi \in [L^2(\Omega)]^R$ , define

$$F_1(\mathbf{U}, \mathbf{Y})(t) := \varphi + \int_0^t C_{\mathbf{Y}_s}^{-1} \mathbb{E}[\mathbf{Y}_s a(s, \mathbf{U}_s^\top \mathbf{Y}_s)^\top] (I_{d \times d} - P_{\mathbf{U}_s}^{\text{row}}) ds \in \mathbb{R}^{R \times d}, \quad (3.1)$$

$$F_2(\mathbf{U}, \mathbf{Y})(t) := \xi + \int_0^t \mathbf{U}_s a(s, \mathbf{U}_s^\top \mathbf{Y}_s) ds + \int_0^t \mathbf{U}_s b(s, \mathbf{U}_s^\top \mathbf{Y}_s) dW_s \in [L^2(\Omega)]^R, \quad (3.2)$$

where we recall that  $P_{\mathbf{U}_s}^{\text{row}}$  is the projection-matrix onto the row space  $\text{span}\{U_s^1, \dots, U_s^R\} \subset \mathbb{R}^d$  of  $\mathbf{U}_s$ . Note that the stochastic integral  $\int_0^t \mathbf{U}_s b(s, \mathbf{U}_s^\top \mathbf{Y}_s) dW_s$  is well defined, since  $\mathbf{Y} \in L^2(\Omega; C([0, t]; \mathbb{R}^R))$  is (indistinguishable from) a progressively measurable process.

We will construct a unique fixed point of  $F_1$  and  $F_2$ . Because of the aforementioned difficulties, defining a suitable sequence of Picard iterates requires some care. Let us consider  $\varphi \in \mathbb{R}^{R \times d}$  with  $d \geq R$  having orthogonal row vectors, and  $\xi = (\xi_1, \dots, \xi_R) \in [L^2(\Omega)]^R$ , having linearly independent components,  $\mathcal{F}_0$ -measurable with  $\rho^2 := \|\xi\|_{[L^2(\Omega)]^R}^2$  and  $\gamma := \|C_\xi^{-1}\|_{\text{F}}$ . We force iterations to belong to balls in  $\mathbb{R}^{R \times d}$  and  $[L^2(\Omega)]^R$  around  $\varphi$  and  $\xi$ , respectively, of a suitable radius  $\eta$ .

For the ball in  $\mathbb{R}^{R \times d}$ , to invoke Proposition A.1 in the appendix, we equip  $\mathbb{R}^{R \times d} = [\mathbb{R}^d]^R$  with the norm  $\|\mathbf{U}\|_{[\mathbb{R}^d]^R}^2 = \sum_{j=1}^R |U^j|^2$ . For  $\boldsymbol{\varphi} = (\varphi^1, \dots, \varphi^R)^\top$  orthogonal, we have  $\|\boldsymbol{\varphi}\|_{[\mathbb{R}^d]^R} = \sqrt{R}$ , and

$$\mathbf{Z}_\boldsymbol{\varphi} := (\boldsymbol{\varphi}^j (\boldsymbol{\varphi}^k)^\top)_{j,k=1,\dots,R} = \boldsymbol{\varphi} \boldsymbol{\varphi}^\top = I_{R \times R},$$

so with  $\eta_1 := \eta(\sqrt{R}, \sqrt{R})$  as in (A.1),  $\mathbf{v} \in B_{\eta_1}(\boldsymbol{\varphi})$  implies  $\|(\mathbf{v} \mathbf{v}^\top)^{-1}\|_{\mathbb{F}} \leq 2\sqrt{R}$ . Hence, the projection  $(I_{d \times d} - P_{\mathbf{U}_t^{\text{row}}})$  is Lipschitz continuous on  $B_{\eta_1}(\boldsymbol{\varphi})$ ; see Lemma A.4.

For the ball in  $[L^2(\Omega)]^R$ , first note that, since  $[L^2(\Omega)]^R \ni \mathbf{Z} \mapsto \mathbb{E}[\mathbf{Z} \mathbf{Z}^\top]^{-1} \in \mathbb{R}^{R \times R}$  is continuous in the open set  $\Gamma = \{\mathbf{Z} \in [L^2(\Omega)]^R \mid \det(\mathbb{E}[\mathbf{Z} \mathbf{Z}^\top]) \neq 0\}$  (cf. [16, Proof of Lemma 3.5]), and  $\boldsymbol{\xi} \in \Gamma$ , via Proposition A.1 there exists a ball  $B_{\eta_2}(\boldsymbol{\xi})$  in  $[L^2(\Omega)]^R$  around  $\boldsymbol{\xi}$  with radius  $\eta_2 := \eta(\rho, \gamma) > 0$  such that any  $\mathbf{w} \in B_{\eta_2}(\boldsymbol{\xi})$  has linearly independent components and  $\|C_{\mathbf{w}}^{-1}\|_{\mathbb{F}} \leq 2\gamma$ . Abusing the notation slightly, we let

$$\eta = \eta(R, \rho, \gamma) := \min\{\eta(\sqrt{R}, \sqrt{R}), \eta(\rho, \gamma)\}. \quad (3.3)$$

Proposition A.1 tells us that  $\eta$  is non-increasing in both variables  $\rho$  and  $\gamma$ .

We want to define sequences  $\mathbf{Y}_t^{(n)} \in B_\eta(\boldsymbol{\xi})$  and  $\mathbf{U}_t^{(n)} \in B_\eta(\boldsymbol{\varphi})$  for  $t \in [0, \delta]$  with a suitable  $\delta = \delta(\boldsymbol{\varphi}, \boldsymbol{\xi}) > 0$ . To this extent, let  $(\mathbf{U}_t^{(0)}, \mathbf{Y}_t^{(0)}) := (\boldsymbol{\varphi}, \boldsymbol{\xi})$ ,  $\mathbf{U}_t^{(n+1)} := F_1(\mathbf{U}^{(n)}, \mathbf{Y}^{(n)})(t)$ ,  $\mathbf{Y}_t^{(n+1)} := F_2(\mathbf{U}^{(n)}, \mathbf{Y}^{(n)})(t)$ , and  $X_t^{(n)} := (\mathbf{U}_t^{(n)})^\top \mathbf{Y}_t^{(n)}$ ,  $n = 0, 1, \dots$  for  $t \in [0, \delta]$  with

$$\delta := \min\left\{1, \frac{\min\{1, \eta^2\}}{36RC_{\text{lgb}}(1 + 3R(3\rho^2 + 1))}, \frac{\min\{\eta^2, R\}}{8\gamma^2(3\rho^2 + 1)C_{\text{lgb}}(1 + 3R(3\rho^2 + 1))}\right\}. \quad (3.4)$$

Moreover, let

$$\mathbb{D}_{\text{det}} := \left\{ \mathbf{V} \in C([0, \delta]; \mathbb{R}^{R \times d}) \left| \begin{array}{l} \sup_{0 \leq t \leq \delta} \|\mathbf{V}_t\|_{\mathbb{F}}^2 \leq 3R, \text{ and} \\ \mathbf{V}_t \in B_\eta(\boldsymbol{\varphi}) \text{ for } t \in [0, \delta] \end{array} \right. \right\}$$

and

$$\mathbb{D}_{\text{sto}} := \left\{ \mathbf{Z} \in L^2(\Omega; C([0, \delta]; \mathbb{R}^R)) \left| \begin{array}{l} \mathbf{Z} \text{ is } \mathcal{F}_t\text{-adapted,} \\ \mathbb{E}\left[\sup_{0 \leq t \leq \delta} |\mathbf{Z}_t|^2\right] \leq 3\rho^2 + 1, \text{ and} \\ \mathbf{Z}_t \in B_\eta(\boldsymbol{\xi}) \text{ for } t \in [0, \delta] \end{array} \right. \right\}.$$

The following lemma shows that our Picard sequence takes value in  $\mathbb{D}_{\text{det}} \times \mathbb{D}_{\text{sto}}$ .

**Lemma 3.1.** *Under Assumptions 2–3, the sequence  $((\mathbf{U}^{(n)}, \mathbf{Y}^{(n)}))_{n \geq 0}$  defined above satisfies*

$$(\mathbf{U}^{(n)}, \mathbf{Y}^{(n)}) \in \mathbb{D}_{\text{det}} \times \mathbb{D}_{\text{sto}} \quad \text{for all } n \in \mathbb{N},$$

where  $\delta = \delta(C_{\text{lgb}}, d, \eta, R, \rho)$  is defined in (3.4).

*Proof.* We have  $\|\boldsymbol{\varphi}\|_{\mathbb{F}}^2 = R$ ,  $\|\boldsymbol{\xi}\|_{[L^2(\Omega)]^R}^2 = \rho^2$ ,  $\|C_{\boldsymbol{\xi}}^{-1}\|_{\mathbb{F}} = \gamma$ , so, trivially,  $\boldsymbol{\xi} \in B_\eta(\boldsymbol{\xi})$  and  $\boldsymbol{\varphi} \in B_\eta(\boldsymbol{\varphi})$ , where  $\eta = \eta(R, \rho, \gamma)$  is built as in (3.3). Moreover,  $\boldsymbol{\xi}$  is  $\mathcal{F}_t$ -adapted thanks to Assumption 3. Thus,  $(\mathbf{U}^{(0)}, \mathbf{Y}^{(0)}) \in \mathbb{D}_{\text{det}} \times \mathbb{D}_{\text{sto}}$ . Assume  $(\mathbf{U}^{(n)}, \mathbf{Y}^{(n)}) \in \mathbb{D}_{\text{det}} \times \mathbb{D}_{\text{sto}}$  for  $n \in \mathbb{N}$ . Then, from Assumption 2, we see that  $\mathbf{U}^{(n+1)}$  and  $\mathbf{Y}^{(n+1)}$  are well defined,

and that  $\mathbf{Y}^{(n+1)}$  is  $\mathcal{F}_t$ -adapted. Moreover,  $\mathbb{E}\left[\sup_{0 \leq t \leq \delta} |\mathbf{Y}_t^{(n)}|^2\right] < \infty$  and the inequality  $\|AB\|_F \leq \|A\|_2 \|B\|_F$  for  $A \in \mathbb{R}^{n \times m}$ ,  $B \in \mathbb{R}^{m \times p}$  implies

$$\begin{aligned}
 & \mathbb{E}\left[\sup_{0 \leq t \leq \delta} |\mathbf{Y}_t^{(n+1)}|^2\right] \\
 & \leq 3\mathbb{E}[|\boldsymbol{\xi}|^2] + 3\mathbb{E}\left[\sup_{0 \leq t \leq \delta} t \int_0^t \|\mathbf{U}_s^{(n)}\|_2^2 |a(s, X_s^{(n)})|^2 ds + \sup_{0 \leq t \leq \delta} \left|\int_0^t \mathbf{U}_s^{(n)} b(s, X_s^{(n)}) dW_s\right|^2\right] \\
 & \leq 3\mathbb{E}[|\boldsymbol{\xi}|^2] + 3\mathbb{E}\left[\sup_{0 \leq t \leq \delta} t \int_0^t \|\mathbf{U}_s^{(n)}\|_2^2 |a(s, X_s^{(n)})|^2 ds\right] + 12\mathbb{E}\left[\left|\int_0^\delta \mathbf{U}_s^{(n)} b(s, X_s^{(n)}) dW_s\right|^2\right] \\
 & \leq 3\rho^2 + 3\mathbb{E}\left[3\delta R \int_0^\delta |a(s, X_s^{(n)})|^2 ds\right] + 12\mathbb{E}\left[\int_0^\delta \|\mathbf{U}_s^{(n)} b(s, X_s^{(n)})\|_F^2 ds\right] \\
 & \leq 3\rho^2 + 36RC_{\text{lgb}}(1 + \mathbb{E}[\sup_{0 \leq s \leq \delta} |X_s^{(n)}|^2])\delta \\
 & \leq 3\rho^2 + 36RC_{\text{lgb}}(1 + 3R(3\rho^2 + 1))\delta \leq 3\rho^2 + 1,
 \end{aligned}$$

where in the second and in the third lines we employ the Doob's martingale inequality and Itô's isometry, respectively, in the penultimate line we used  $\delta \leq 1$ , and in the last line  $|a(s, x)|^2 + \|b(s, x)\|_F^2 \leq C_{\text{lgb}}(1 + |x|^2)$  together with the definition of  $\delta$ . Similarly, we have

$$\begin{aligned}
 \mathbb{E}\left[\sup_{0 \leq t \leq \delta} |\mathbf{Y}_t^{(n+1)} - \boldsymbol{\xi}|^2\right] & \leq 2\mathbb{E}\left[\sup_{0 \leq t \leq \delta} t \int_0^t \|\mathbf{U}_s^{(n)}\|_2^2 |a(s, X_s^{(n)})|^2 ds + \sup_{0 \leq t \leq \delta} \left|\int_0^t \mathbf{U}_s^{(n)} b(s, X_s^{(n)}) dW_s\right|^2\right] \\
 & \leq 24C_{\text{lgb}}R(1 + 3R(3\rho^2 + 1))\delta \leq \eta^2.
 \end{aligned}$$

We readily have  $\mathbf{Y}^{(n+1)} \in L^2(\Omega; C([0, \delta]; \mathbb{R}^R))$  and hence  $\mathbf{Y}^{(n+1)} \in \mathbb{D}_{\text{sto}}$ . Likewise, we have  $\mathbf{U}^{(n+1)} \in \mathbb{D}_{\text{det}}$ , since

$$\begin{aligned}
 \sup_{0 \leq t \leq \delta} \|\mathbf{U}_t^{(n+1)}\|_F^2 & \leq 2R + 2 \sup_{0 \leq t \leq \delta} t \int_0^t \|C_{\mathbf{Y}_s^{(n)}}^{-1} \mathbb{E}[\mathbf{Y}_s^{(n)} a(s, X_s^{(n)})^\top](I_{d \times d} - P_{\mathbf{U}_s^{(n)}}^{\text{row}})\|_F^2 ds \\
 & \leq 2R + 8\delta\gamma^2 \int_0^\delta \mathbb{E}\left[\sup_{0 \leq s \leq \delta} |\mathbf{Y}_s^{(n)}|^2\right] \mathbb{E}\left[\sup_{0 \leq s \leq \delta} |a(s, X_s^{(n)})|^2\right] ds \\
 & \leq 2R + 8\delta\gamma^2(3\rho^2 + 1)C_{\text{lgb}}(1 + 3R(3\rho^2 + 1)) \\
 & \leq 3R,
 \end{aligned}$$

where in the second inequality we have used the fact that  $P_{\mathbf{U}_s^{(n)}}^{\text{row}}$  is an orthogonal projector. Finally,

$$\sup_{0 \leq t \leq \delta} \|\mathbf{U}_t^{(n+1)} - \boldsymbol{\varphi}\|_F^2 \leq 4\gamma^2(3\rho^2 + 1)C_{\text{lgb}}(1 + 3R(3\rho^2 + 1))\delta \leq \eta^2.$$

Thus, by induction we conclude  $(\mathbf{U}^{(n)}, \mathbf{Y}^{(n)}) \in \mathbb{D}_{\text{det}} \times \mathbb{D}_{\text{sto}}$  for  $n \in \mathbb{N}$ .  $\square$

We now establish a Lipschitz continuity for  $F_1$  and  $F_2$  on  $\mathbb{D}_{\text{det}} \times \mathbb{D}_{\text{sto}}$ .



**Lemma 3.2.** *Take  $\delta > 0$  as in (3.4). There exists a constant  $\tilde{C} := \tilde{C}_{a,b,R,\rho,\delta,\gamma} > 0$  such that for any  $(\mathbf{V}, \mathbf{Z}), (\tilde{\mathbf{V}}, \tilde{\mathbf{Z}}) \in \mathbb{D}_{\det} \times \mathbb{D}_{\text{sto}}$  it holds*

$$\begin{aligned} & \sup_{t \in [0, \delta]} \|F_1(\mathbf{V}, \mathbf{Z})(t) - F_1(\tilde{\mathbf{V}}, \tilde{\mathbf{Z}})(t)\|_{\mathbb{F}}^2 + \mathbb{E} \left[ \sup_{t \in [0, \delta]} |F_2(\mathbf{V}, \mathbf{Z})(t) - F_2(\tilde{\mathbf{V}}, \tilde{\mathbf{Z}})(t)|^2 \right] \\ & \leq \tilde{C} \int_0^\delta \left( \sup_{t \in [0, s]} \|\mathbf{V}_t - \tilde{\mathbf{V}}_t\|_{\mathbb{F}}^2 + \mathbb{E} \left[ \sup_{t \in [0, s]} |\mathbf{Z}_t - \tilde{\mathbf{Z}}_t|^2 \right] \right) ds \end{aligned} \quad (3.5)$$

*Proof.* For any  $(\mathbf{V}, \mathbf{Z}), (\tilde{\mathbf{V}}, \tilde{\mathbf{Z}}) \in \mathbb{D}_{\det} \times \mathbb{D}_{\text{sto}}$ , from Doob's martingale inequality and Itô's isometry applied similarly as in the proof of Lemma 3.1, we have

$$\begin{aligned} & \mathbb{E} \left[ \sup_{t \in [0, \delta]} |F_2(\mathbf{V}, \mathbf{Z})(t) - F_2(\tilde{\mathbf{V}}, \tilde{\mathbf{Z}})(t)|^2 \right] \\ & \leq 2\mathbb{E} \left[ \sup_{t \in [0, \delta]} \left| \int_0^t \mathbf{V}_s a(s, \mathbf{V}_s^\top \mathbf{Z}_s) ds - \int_0^t \tilde{\mathbf{V}}_s a(s, \tilde{\mathbf{V}}_s^\top \tilde{\mathbf{Z}}_s) ds \right|^2 \right] \\ & \quad + 2\mathbb{E} \left[ \sup_{t \in [0, \delta]} \left| \int_0^t (\mathbf{V}_s b(s, \mathbf{V}_s^\top \mathbf{Z}_s) - \tilde{\mathbf{V}}_s b(s, \tilde{\mathbf{V}}_s^\top \tilde{\mathbf{Z}}_s)) dW_s \right|^2 \right] \\ & \leq 2\mathbb{E} \left[ \sup_{t \in [0, \delta]} \int_0^t |\mathbf{V}_s a(s, \mathbf{V}_s^\top \mathbf{Z}_s) - \tilde{\mathbf{V}}_s a(s, \tilde{\mathbf{V}}_s^\top \tilde{\mathbf{Z}}_s)|^2 ds \right] \\ & \quad + 8\mathbb{E} \left[ \int_0^\delta \|\mathbf{V}_s b(s, \mathbf{V}_s^\top \mathbf{Z}_s) - \tilde{\mathbf{V}}_s b(s, \tilde{\mathbf{V}}_s^\top \tilde{\mathbf{Z}}_s)\|_{\mathbb{F}}^2 ds \right] \\ & \leq C_{a,b,R,\rho,\delta} \int_0^\delta (\|\mathbf{V}_s - \tilde{\mathbf{V}}_s\|_{\mathbb{F}}^2 + \mathbb{E}[|\mathbf{Z}_s - \tilde{\mathbf{Z}}_s|^2]) ds \\ & \leq C_{a,b,R,\rho,\delta} \int_0^\delta \left( \sup_{r \in [0, s]} \|\mathbf{V}_r - \tilde{\mathbf{V}}_r\|_{\mathbb{F}}^2 + \mathbb{E} \left[ \sup_{r \in [0, s]} |\mathbf{Z}_r - \tilde{\mathbf{Z}}_r|^2 \right] \right) ds, \end{aligned}$$

where  $C_{a,b,R,\rho,\delta}$  is a positive constant. Similarly, we have for a constant  $C_{a,b,R,\rho,\delta,\gamma} > 0$  that

$$\sup_{t \in [0, \delta]} \|F_1(\mathbf{V}, \mathbf{Z})(t) - F_1(\tilde{\mathbf{V}}, \tilde{\mathbf{Z}})(t)\|_{\mathbb{F}}^2 \leq C_{a,b,R,\rho,\delta,\gamma} \int_0^\delta \left( \sup_{r \in [0, s]} \|\mathbf{V}_r - \tilde{\mathbf{V}}_r\|_{\mathbb{F}}^2 + \mathbb{E} \left[ \sup_{r \in [0, s]} |\mathbf{Z}_r - \tilde{\mathbf{Z}}_r|^2 \right] \right) ds,$$

where we used the Lipschitz continuity of  $C_{\mathbf{Y}_s}^{-1}$  and  $P_{\mathbf{U}_s}^{\text{row}}$ ; see [16, Lemma 3.5] and Lemma A.4.  $\square$

Thanks to the previous results, we have that sequences  $(\mathbf{U}^{(n)})_n$  and  $(\mathbf{Y}^{(n)})_n$  not only live in  $\mathbb{D}_{\det}$  and  $\mathbb{D}_{\text{sto}}$ , respectively, but also converge therein.

**Lemma 3.3.** *The sequence  $(\mathbf{U}^{(n)})_n$  admits a limit  $\mathbf{U} \in \mathbb{D}_{\det} \subset C([0, \delta]; \mathbb{R}^{R \times d})$ , and the sequence  $(\mathbf{Y}^{(n)})_n$  admits a limit  $\mathbf{Y} \in \mathbb{D}_{\text{sto}} \subset C([0, \delta]; \mathbb{R}^R)$  almost surely. Moreover,  $\mathbf{Y}$  is also an  $L^2(\Omega; C([0, \delta]; \mathbb{R}^R))$ -limit.*

*Proof.* From Lemma 3.1 we have  $(\mathbf{U}^{(n)}, \mathbf{Y}^{(n)}) \in \mathbb{D}_{\det} \times \mathbb{D}_{\text{sto}}$  for all  $n \in \mathbb{N}$ . Let  $\Delta_U^{(n)}(s) := \sup_{0 \leq r \leq s} \|\mathbf{U}_r^{(n)} - \mathbf{U}_r^{(n-1)}\|_{\mathbb{F}}^2$ . Then, since  $\|\mathbf{U}^{(n)} - \mathbf{U}^{(n-1)}\|_{\mathbb{F}}^2$  is continuous on  $[0, \delta]$ , so is  $\Delta_U^{(n)}$ , and thus  $\Delta_U^{(n)}$  is measurable. Similarly,  $\Delta_Y^{(n)}(s) := \sup_{0 \leq r \leq s} |\mathbf{Y}_r^{(n)} - \mathbf{Y}_r^{(n-1)}|^2$

is a.s. continuous on  $[0, \delta]$ . Noting that  $\mathbf{Y}^{(n)} \in \mathbb{D}_{\text{sto}}$  implies  $\|C_{\mathbf{Y}_t^{(n)}}^{-1}\|_{\mathbb{F}}^2 \leq 2\gamma$ , from Lemma 3.2 we have

$$\begin{aligned} \Delta_U^{(n)}(\delta) + \mathbb{E}[\Delta_Y^{(n)}(\delta)] &\leq \tilde{C} \int_0^\delta (\Delta_U^{(n-1)}(s) + \mathbb{E}[\Delta_Y^{(n-1)}(s)]) ds \\ &\leq \tilde{C}^{n-1} \int_0^\delta \int_0^{s_{n-1}} \cdots \int_0^{s_2} (\Delta_U^{(1)}(s_1) + \mathbb{E}[\Delta_Y^{(1)}(s_1)]) ds_1 \cdots ds_{n-1} \\ &= \frac{(\tilde{C}\delta)^{n-1}}{(n-1)!} (\Delta_U^{(1)}(\delta) + \mathbb{E}[\Delta_Y^{(1)}(\delta)]). \end{aligned} \quad (3.6)$$

Chebyshev's inequality then implies

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\Delta_U^{(n)}(\delta) + \Delta_Y^{(n)}(\delta) \geq \frac{1}{2^n}\right) \leq \left(\Delta_U^{(1)}(\delta) + \mathbb{E}[\Delta_Y^{(1)}(\delta)]\right) 2 \sum_{n=1}^{\infty} \frac{(2\tilde{C}\delta)^{n-1}}{(n-1)!} < \infty,$$

and thus from the Borel-Cantelli lemma we have

$$\mathbb{P}\left(\left\{\exists k = k(\omega) \text{ s.t. } n \geq k \implies \Delta_U^{(n)}(\delta) + \Delta_Y^{(n)}(\delta) < \frac{1}{2^n}\right\}\right) = 1.$$

Hence,  $(\mathbf{Y}^{(n)}(\omega))_n$  has a limit  $\mathbf{Y}(\omega) \in C([0, \delta]; \mathbb{R}^R)$ , where the convergence is uniformly in  $t$  on  $[0, \delta]$ , a.s. Moreover, from the completeness of the underlying probability space,  $\mathbf{Y}$  is  $(\mathcal{F}_t)$ -adapted. Also, from  $\mathbf{Y}^{(n)} \in \mathbb{D}_{\text{sto}}$ ,  $n \in \mathbb{N}$ , Fatou's lemma implies

$$\mathbb{E}\left[\sup_{0 \leq t \leq \delta} |\mathbf{Y}_t|^2\right] \leq \liminf_{n \rightarrow \infty} \mathbb{E}\left[\sup_{0 \leq t \leq \delta} |\mathbf{Y}_t^{(n)}|^2\right] \leq 4\rho, \quad (3.7)$$

$$\mathbb{E}\left[\sup_{0 \leq t \leq \delta} |\mathbf{Y}_t - \boldsymbol{\xi}|^2\right] \leq \eta^2, \quad (3.8)$$

and thus  $\mathbf{Y} \in \mathbb{D}_{\text{sto}}$ . An analogous argument applies to see that  $(\mathbf{U}^{(n)})_n$  has a limit  $\mathbf{U}$  in  $C([0, \delta]; \mathbb{R}^{R \times d})$  with  $\mathbf{U} \in \mathbb{D}_{\text{det}}$ .

The sequences  $\mathbf{Y}^{(n)}$  converge in  $L^2(\Omega; C([0, \delta]; \mathbb{R}^R))$  as well. Indeed, from (3.6), for  $j > n$  we have

$$\begin{aligned} \sqrt{\mathbb{E}\left[\sup_{0 \leq t \leq \delta} |\mathbf{Y}_t^{(j)} - \mathbf{Y}_t^{(n)}|^2\right]} &\leq \sum_{k=n}^{j-1} \sqrt{\mathbb{E}\left[\sup_{0 \leq t \leq \delta} |\mathbf{Y}_t^{(k+1)} - \mathbf{Y}_t^{(k)}|^2\right]} \\ &\leq \sqrt{\Delta_U^{(1)}(\delta) + \mathbb{E}[\Delta_Y^{(1)}(\delta)]} \sum_{k=n}^{j-1} \sqrt{\frac{(\tilde{C}\delta)^{k-1}}{(k-1)!}} \end{aligned}$$

and thus Fatou's lemma implies

$$\mathbb{E}\left[\sup_{0 \leq t \leq \delta} |\mathbf{Y}_t - \mathbf{Y}_t^{(n)}|^2\right] \leq (\Delta_U^{(1)}(\delta) + \mathbb{E}[\Delta_Y^{(1)}(\delta)]) \left(\sum_{k=n}^{\infty} \sqrt{\frac{(\tilde{C}\delta)^{k-1}}{(k-1)!}}\right)^2 < \infty,$$

hence  $\lim_{n \rightarrow \infty} \mathbb{E}\left[\sup_{0 \leq t \leq \delta} |\mathbf{Y}_t - \mathbf{Y}_t^{(n)}|^2\right] = 0$ . □

The argument used in Lemmata 3.1-3.3 to show convergence of the Picard iterates resembles the standard one used in establishing well-posedness of SDEs, however containing important differences. First, the map (3.1) to update  $\mathbf{U}$  involves a “stochastic projection” introducing a dependence on the law of the  $\mathbf{Y}$  process. Second, the same projection involves the term  $C_{\mathbf{Y}}^{-1}$  (inverse of the Gramian of the stochastic basis), which might not be defined for all iterates or it may fail to be Lipschitz continuous with respect to  $\mathbf{Y}$ . We thus have to restrict to a well-designed set, containing the initial state, in which the inverse of the Gramian is well-defined and Lipschitz continuous. Lemma 3.1 guarantees that the Picard iterates  $(\mathbf{U}^{(n)}, \mathbf{Y}^{(n)})$  are contained in such set for all  $n$ . Lemma 3.2 shows the Lipschitz continuity of the mapping that defines the iterates on this set, and Lemma 3.3 assures that  $(\mathbf{U}^{(n)}, \mathbf{Y}^{(n)})$  admit a limit in the same set. We are now prepared to establish our existence result for the DO system, employing an argument that closely resembles the conventional approach used for SDEs.

**Theorem 3.4** (Existence of a DO solution). *For any  $(\varphi, \xi)$ , where  $\varphi \in \mathbb{R}^{d \times R}$  has orthonormal rows and  $\xi \in L^2(\Omega, \mathbb{R}^R)$  satisfies  $\|\xi\|_{[L^2(\Omega)]^R} = \rho > 0$  and  $\|C_{\xi}^{-1}\|_{\mathbb{F}} = \gamma > 0$ , the DO equations (2.6) and (2.7) with the initial condition  $(\mathbf{U}_0, \mathbf{Y}_0) = (\varphi, \xi)$  have a local (in time) strong DO solution  $(\mathbf{U}, \mathbf{Y}) \in \mathbb{D}_{\det} \times \mathbb{D}_{\text{sto}}$  with  $\delta$  given by (3.4).*

*Proof.* We show that the limit  $(\mathbf{U}, \mathbf{Y})$  in Lemma 3.3 satisfies  $\mathbf{U} = F_1(\mathbf{U}, \mathbf{Y})$  and  $\mathbf{Y} = F_2(\mathbf{U}, \mathbf{Y})$ , hence it is a DO solution. From Lemma 3.2,  $F_1$  and  $F_2$  are (Lipschitz) continuous on  $\mathbb{D}_{\det} \times \mathbb{D}_{\text{sto}}$ :

$$\begin{aligned} & \max \left\{ \sup_{t \in [0, \delta]} \|F_1(\mathbf{V}, \mathbf{Z})(t) - F_1(\tilde{\mathbf{V}}, \tilde{\mathbf{Z}})(t)\|_{\mathbb{F}}^2, \mathbb{E} \left[ \sup_{t \in [0, \delta]} |F_2(\mathbf{V}, \mathbf{Z})(t) - F_2(\tilde{\mathbf{V}}, \tilde{\mathbf{Z}})(t)|^2 \right] \right\} \\ & \leq \tilde{C} \int_0^\delta \left( \sup_{t \in [0, s]} \|\mathbf{V}_t - \tilde{\mathbf{V}}_t\|_{\mathbb{F}}^2 + \mathbb{E} \left[ \sup_{t \in [0, s]} |\mathbf{Z}_t - \tilde{\mathbf{Z}}_t|^2 \right] \right) ds \\ & \leq \tilde{C} \delta \left( \sup_{t \in [0, \delta]} \|\mathbf{V}_t - \tilde{\mathbf{V}}_t\|_{\mathbb{F}}^2 + \mathbb{E} \left[ \sup_{t \in [0, \delta]} |\mathbf{Z}_t - \tilde{\mathbf{Z}}_t|^2 \right] \right). \end{aligned}$$

Thus, we have

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \in [0, \delta]} |\mathbf{Y}_t - F_2(\mathbf{U}, \mathbf{Y})(t)|^2 \right] &= \lim_{n \rightarrow \infty} \mathbb{E} \left[ \sup_{t \in [0, \delta]} |\mathbf{Y}_t^{(n)} - F_2(\mathbf{U}^{(n)}, \mathbf{Y}^{(n)})(t)|^2 \right] \\ &= \lim_{n \rightarrow \infty} \mathbb{E} \left[ \sup_{t \in [0, \delta]} |\mathbf{Y}_t^{(n)} - \mathbf{Y}_t^{(n+1)}|^2 \right] = 0, \end{aligned}$$

and hence  $\mathbf{Y}_t = F_2(\mathbf{U}, \mathbf{Y})(t)$  for  $t \in [0, \delta]$ , a.s. We see  $\mathbf{U}_t = F_1(\mathbf{U}, \mathbf{Y})(t)$  analogously.  $\square$

To establish uniqueness, first we will show the following norm bound analogous to Lemma 2.7.

**Lemma 3.5.** *For  $T > 0$ , suppose that  $\mathbf{U} \in C([0, T]; \mathbb{R}^{R \times d})$ , with  $\mathbf{U}_0$  having orthonormal rows, and  $\mathbf{Y} \in L^2(\Omega; C([0, T]; \mathbb{R}^R))$ , with  $\mathbf{Y}_0$  having linearly independent components, satisfy  $\mathbf{U}_t = F_1(\mathbf{U}, \mathbf{Y})(t)$  and  $\mathbf{Y}_t = F_2(\mathbf{U}, \mathbf{Y})(t)$  for all  $t \in [0, T]$ . Then, for all  $t \in [0, T]$  we have*

$$\|\mathbf{U}_t\|_{\mathbb{F}} = \sqrt{R}; \quad (3.9)$$

$$\mathbb{E} [|\mathbf{Y}_t|^2] \leq 3(\mathbb{E}[|\mathbf{Y}_0|^2] + (1+T)TC_{\text{lgb}}) \exp(3(1+T)TC_{\text{lgb}}) =: M(T). \quad (3.10)$$

*Proof.* First, from  $\mathbf{U}_t = F_1(\mathbf{U}, \mathbf{Y})(t)$ , the function  $\mathbf{U}$  is absolutely continuous on  $[0, T]$ , and thus differentiable almost everywhere. The derivative  $\dot{\mathbf{U}}_t$  satisfies

$$\dot{\mathbf{U}}_t \mathbf{U}_t^\top = C_{\mathbf{Y}_t}^{-1} \mathbb{E}[\mathbf{Y}_t a(t, \mathbf{U}_t^\top \mathbf{Y}_t)^\top] (I_{d \times d} - P_{\mathbf{U}_t}^{\text{row}}) \mathbf{U}_t^\top = 0,$$

and thus  $\frac{d}{dt}(\mathbf{U}_t \mathbf{U}_t^\top) = 0$  a.e. on  $[0, T]$ . Therefore, from the orthonormality of the initial condition  $\boldsymbol{\varphi} \boldsymbol{\varphi}^\top = I_{R \times R}$ , for all  $t \in [0, T]$  we have

$$(\mathbf{U}_t \mathbf{U}_t^\top)_{jk} = \epsilon_{jk} + \int_0^t 0 \, ds = \epsilon_{jk},$$

where  $\epsilon_{jk} = 1$  only if  $j = k$ , and 0 otherwise. This shows the identity (3.9).

For  $\mathbf{Y}$ , Itô's isometry implies

$$\begin{aligned} \mathbb{E}[|\mathbf{Y}_t|^2] &\leq 3\mathbb{E}\left[|\mathbf{Y}_0|^2 + t \int_0^t \|\mathbf{U}_s\|_2^2 |a(s, \mathbf{U}_s^\top \mathbf{Y}_s)|^2 \, ds\right] + 3\mathbb{E}\left[\int_0^t \|\mathbf{U}_s b(s, \mathbf{U}_s^\top \mathbf{Y}_s)\|_{\mathbb{F}}^2 \, ds\right] \\ &\leq 3\mathbb{E}[|\mathbf{Y}_0|^2] + 3(1+T)C_{\text{lgb}} \int_0^t (1 + \mathbb{E}[|\mathbf{Y}_s|^2]) \, ds, \end{aligned}$$

where we used Assumption 2  $|a(s, x)|^2 + \|b(s, x)\|_{\mathbb{F}}^2 \leq C_{\text{lgb}}(1 + |x|^2)$ . Thus, Gronwall's lemma implies (3.10).  $\square$

Now the following uniqueness result follows.

**Theorem 3.6** (Uniqueness of DO solutions). *Let the assumptions of Lemma 3.5 hold. For  $T > 0$ , suppose that  $\mathbf{U}, \tilde{\mathbf{U}} \in C([0, T]; \mathbb{R}^{R \times d})$  and  $\mathbf{Y}, \tilde{\mathbf{Y}} \in L^2(\Omega; C([0, T]; \mathbb{R}^R))$  satisfy  $\mathbf{U}_t = F_1(\mathbf{U}, \mathbf{Y})(t)$  and  $\tilde{\mathbf{U}}_t = F_1(\tilde{\mathbf{U}}, \tilde{\mathbf{Y}})(t)$ ;  $\mathbf{Y}_t = F_2(\mathbf{U}, \mathbf{Y})(t)$  and  $\tilde{\mathbf{Y}}_t = F_2(\tilde{\mathbf{U}}, \tilde{\mathbf{Y}})(t)$  for  $t \in [0, T]$ . Then, we have*

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} \|\mathbf{U}_t - \tilde{\mathbf{U}}_t\|_{\mathbb{F}}^2 + \sup_{0 \leq t \leq T} |\mathbf{Y}_t - \tilde{\mathbf{Y}}_t| > 0\right) = 0.$$

*Proof.* By hypothesis, the solutions  $\mathbf{U}, \tilde{\mathbf{U}}$  and  $\mathbf{Y}, \tilde{\mathbf{Y}}$  satisfy the stability estimates shown in Lemma 3.5. Moreover, from the continuity of  $t \rightarrow \mathbb{E}[\mathbf{Y}_t \mathbf{Y}_t^\top]^{-1}$ , we have

$$\max\left\{\max_{s \in [0, T]} \|\mathbb{E}[\mathbf{Y}_s \mathbf{Y}_s^\top]^{-1}\|_{\mathbb{F}}, \max_{s \in [0, T]} \|\mathbb{E}[\tilde{\mathbf{Y}}_s \tilde{\mathbf{Y}}_s^\top]^{-1}\|_{\mathbb{F}}\right\} = \tilde{\gamma} < \infty$$

for some  $\tilde{\gamma} > 0$ . Then, noting the norm bounds (3.9) and (3.10), by an argument similar to the proof of Lemma 3.2 (see also [16, Lemma 3.5] and Lemma A.4), with a constant  $\tilde{C} = \tilde{C}(\tilde{\gamma}) > 0$  we have

$$\sup_{0 \leq s \leq t} \|\mathbf{U}_s - \mathbf{U}'_s\|_{\mathbb{F}}^2 + \mathbb{E}\left[\sup_{0 \leq s \leq t} |\mathbf{Y}_s - \mathbf{Y}'_s|^2\right] \leq \tilde{C} \int_0^t \left(\sup_{0 \leq s \leq r} \|\mathbf{U}_s - \mathbf{U}'_s\|_{\mathbb{F}}^2 + \mathbb{E}\left[\sup_{0 \leq s \leq r} |\mathbf{Y}_s - \mathbf{Y}'_s|^2\right]\right) \, dr \quad (3.11)$$

for  $t \in [0, T]$ . Thus, applying the Gronwall's lemma yields

$$\sup_{0 \leq s \leq t} \|\mathbf{U}_s - \mathbf{U}'_s\|_{\mathbb{F}}^2 + \mathbb{E}\left[\sup_{0 \leq s \leq t} |\mathbf{Y}_s - \mathbf{Y}'_s|^2\right] = 0.$$

Now the proof is complete.  $\square$

**Remark 3.7.** The uniqueness of the DO solution can be also deduced from the proof of Proposition 2.10. To see this, let  $(\mathbf{U}, \mathbf{Y})$  and  $(\tilde{\mathbf{U}}, \tilde{\mathbf{Y}})$  be two strong DO solutions with the same initial datum  $(\mathbf{U}_0, \mathbf{Y}_0)$ . Then, following the proof of Proposition 2.10, we have  $(\mathbf{U}, \mathbf{Y}) = (\Theta \mathbf{U}_t, \Theta \mathbf{Y}_t)$  with  $\Theta = I_{R \times R}$ .

We conclude this section by showing the continuity of the solution with respect to the initial datum.

**Lemma 3.8.** *Let  $(\mathbf{U}_t, \mathbf{Y}_t)$  and  $(\tilde{\mathbf{U}}_t, \tilde{\mathbf{Y}}_t)$  be DO solutions on  $[0, T]$  with initial data  $(\mathbf{U}_0, \mathbf{Y}_0)$  and  $(\tilde{\mathbf{U}}_0, \tilde{\mathbf{Y}}_0)$ , respectively. Define  $\rho = \max\{\|\mathbf{Y}_0\|_{[L^2(\Omega)]^R}, \|\tilde{\mathbf{Y}}_0\|_{[L^2(\Omega)]^R}\}$  and  $\gamma = \max\{\|C_{\mathbf{Y}_0}^{-1}\|_F, \|C_{\tilde{\mathbf{Y}}_0}^{-1}\|_F\}$ . Then, under Assumptions 1 and 2,*

$$\mathbb{E}\left[\sup_{t \in [0, T]} |\mathbf{Y}_t - \tilde{\mathbf{Y}}_t|^2\right] + \sup_{t \in [0, T]} \|\mathbf{U}_t - \tilde{\mathbf{U}}_t\|_F^2 \leq C_{a,b,R,\rho,T,\gamma} \left( \mathbb{E}[|\mathbf{Y}_0 - \tilde{\mathbf{Y}}_0|^2] + \|\mathbf{U}_0 - \tilde{\mathbf{U}}_0\|_F^2 \right)$$

holds. Moreover, let  $X_t$  and  $\tilde{X}_t$  be DO approximations on  $[0, T]$  with initial data  $X_0$  and  $\tilde{X}_0$ , respectively, and let

$$\tilde{\gamma} := \inf_{t \in [0, T]} \max\{\sigma_R(\mathbb{E}[X_t X_t^\top]), \sigma_R(\mathbb{E}[\tilde{X}_t \tilde{X}_t^\top])\}.$$

Then, there exists a positive constant  $C > 0$  such that

$$\sup_{t \in [0, T]} \mathbb{E}[|X_t - \tilde{X}_t|^2] \leq C \mathbb{E}[|X_0 - \tilde{X}_0|^2]. \quad (3.12)$$

*Proof.* Using an argument similar to the proof of Lemma 3.2 and the Lipschitz continuity of the projections as in Lemma A.4, we obtain the following relation for the stochastic basis:

$$\begin{aligned} \mathbb{E}\left[\sup_{t \in [0, T]} |\mathbf{Y}_t - \tilde{\mathbf{Y}}_t|^2\right] &\leq 2\mathbb{E}[|\mathbf{Y}_0 - \tilde{\mathbf{Y}}_0|^2] \\ &+ C_{a,b,R,\rho,T} \int_0^T \left( \sup_{r \in [0, t]} \|\mathbf{U}_r - \tilde{\mathbf{U}}_r\|_F^2 + \mathbb{E}\left[\sup_{r \in [0, t]} |\mathbf{Y}_r - \tilde{\mathbf{Y}}_r|^2\right] \right) dt. \end{aligned}$$

Similarly, noting the Lipschitz continuity of the inverse Gram matrix [16, Lemma 3.5], for the deterministic basis we have

$$\|\mathbf{U}_t - \tilde{\mathbf{U}}_t\|_F^2 \leq 2\|\mathbf{U}_0 - \tilde{\mathbf{U}}_0\|_F^2 + C_{a,b,R,\rho,T,\gamma} \int_0^T \left( \sup_{r \in [0, t]} \|\mathbf{U}_r - \tilde{\mathbf{U}}_r\|_F^2 + \mathbb{E}\left[\sup_{r \in [0, t]} |\mathbf{Y}_r - \tilde{\mathbf{Y}}_r|^2\right] \right) dt.$$

Then, the Gronwall's lemma yields the first part of the statement. Finally, to prove (3.12), we proceed as above and as done in Theorem 2.8. Therefore, for all  $t \in [0, T]$  there exists a positive constant  $C$  such that it holds

$$\begin{aligned} \mathbb{E}[|X_t - \tilde{X}_t|^2] &\leq 2\mathbb{E}[|X_0 - \tilde{X}_0|^2] + 2\bar{C}_{a,b,R,T,\tilde{\gamma}} \int_0^T \mathbb{E}[|X_s - \tilde{X}_s|^2] ds, \\ &\leq C \mathbb{E}[|X_0 - \tilde{X}_0|^2], \end{aligned}$$

where in the last line we use the Gronwall's lemma.  $\square$

## 4 Maximality

In the previous section, we established the existence and uniqueness of strong DO solutions locally in time on an interval  $[0, T]$ . In this section, we investigate how much such an interval can be extended.

We give a characterisation of the maximal interval of existence of the strong DO solution in terms of  $\|C_{\mathbf{Y}_t}^{-1}\|_{\mathbb{F}}$ . It turns out that the DO solution can be extended until  $\|C_{\mathbf{Y}_t}^{-1}\|_{\mathbb{F}}$  explodes. If  $\|C_{\mathbf{Y}_t}^{-1}\|_{\mathbb{F}}$  stays bounded for all  $t > 0$ , then the DO solution exists globally. If  $\|C_{\mathbf{Y}_t}^{-1}\|_{\mathbb{F}}$  explodes at a finite *explosion time*  $T_e$ , the couple  $(\mathbf{U}, \mathbf{Y})$  inevitably ceases to exist at  $T_e$ ; we will show nevertheless that the DO approximation  $X = \mathbf{U}^\top \mathbf{Y}$  can be extended beyond  $T_e$ .

### 4.1 Explosion time of the DO solution

Theorem 3.4 guarantees the (unique) existence of a strong DO solution, albeit up to possibly a short time  $T$ . In this section, we show that the solution can be extended until  $C_{\mathbf{Y}_t}^{-1}$  becomes singular.

Analogous maximality results have been considered in the DLRA literature for deterministic and random PDEs [21, 16]. For the SDE case of this paper, we need to proceed with caution. Indeed, adhering to the definition of SDEs, we need the DO solution to be path-wise continuous a.s. As such, we consider extension with countable number of operations, which we describe in the following.

Let  $[0, T]$  be the interval on which the DO solution  $(\mathbf{U}, \mathbf{Y})$  exists; such existence is guaranteed by Theorem 3.4. Let us choose  $n \in \mathbb{N}$  such that the following two bounds are satisfied

$$\begin{aligned} \mathbb{E}[\|\mathbf{Y}_T\|^2] &\leq \mathbb{E}[\|\mathbf{Y}_0\|^2] + n := \rho_n^2, \\ \|C_{\mathbf{Y}_T}^{-1}\|_{\mathbb{F}}^2 &\leq \|C_{\mathbf{Y}_0}^{-1}\|_{\mathbb{F}}^2 + n := \gamma_n^2. \end{aligned} \quad (4.1)$$

We will show that the solution can be extended at least until  $\mathbb{E}[\|\mathbf{Y}_t\|^2]$  or  $\|C_{\mathbf{Y}_t}^{-1}\|_{\mathbb{F}}^2$  hits the bound  $\rho_n^2$  or  $\gamma_n^2$ . Define  $\delta(n)$  by

$$\delta(n) := \min \left\{ 1, \frac{\min\{1, \eta_n^2\}}{36RC_{\text{lgb}}(1 + 3R(3\rho_n^2 + 1))}, \frac{\min\{\eta_n^2, R\}}{8\gamma_n^2(3\rho_n^2 + 1)C_{\text{lgb}}(1 + 3R(3\rho_n^2 + 1))} \right\}, \quad (4.2)$$

where  $\eta_n := \min\{\eta(\rho_n, \gamma_n), \eta(\sqrt{R}, \sqrt{R})\}$ , with  $\eta(\cdot, \cdot)$  defined in (A.1). Then from the proofs of Lemmata 3.1 and 3.3, and Theorem 3.4, we can construct a convergent Picard-iteration for the interval  $[T, T + \delta(n)]$ , which yields the extension of the DO solution up to  $[0, T + \delta(n)]$ . Set  $T := T + \delta(n)$ . By construction  $\delta(n)$  is independent of  $T$ , and with the same  $\delta(n)$  we can repeat the same argument as long as the bound (4.1) is satisfied. Hence, the solution can be extended until either  $\mathbb{E}[\|\mathbf{Y}_{T+\delta(n)}\|^2] \leq \rho_n^2$  or  $\|C_{\mathbf{Y}_{T+\delta(n)}}^{-1}\|_{\mathbb{F}}^2 \leq \gamma_n^2$  gets violated.

From the argument above, the following quantities are well defined for any  $n \in \mathbb{N}$ :

$$\begin{aligned} \tau_n^1 &:= \inf\{t > 0 \mid \|C_{\mathbf{Y}_t}^{-1}\|_{\mathbb{F}} = \|C_{\mathbf{Y}_0}^{-1}\|_{\mathbb{F}} + n\}, \\ \tau_n^2 &:= \inf\{t > 0 \mid \|\mathbf{Y}_t\|_{[L^2(\Omega)]^R} = \|\mathbf{Y}_0\|_{[L^2(\Omega)]^R} + n\} \end{aligned}$$

with a convention  $\inf \emptyset = \infty$ . With these, we define the sequence

$$\tau_n := \tau_n(\mathbf{U}_0, \mathbf{Y}_0) := \min\{\tau_n^1, \tau_n^2\}, \quad n \in \mathbb{N}, \quad (4.3)$$

which is a sequence of stopping times. By continuity of the paths,  $(\tau_n)_{n \in \mathbb{N}}$  is a non-decreasing sequence, which allows us to define  $T_e := \lim_{n \rightarrow \infty} \tau_n$ .

Now we will show that if  $T_e < \infty$ , then the norm of the inverse of the Gram matrix must blow up.

**Proposition 4.1.** *We have either  $T_e = \lim_{n \rightarrow \infty} \tau_n = \infty$  or  $T_e < \infty$ , where  $\tau_n$  is defined in (4.3). If  $T_e < \infty$ , we necessarily have  $\lim_{t \uparrow T_e} \|C_{\mathbf{Y}_t}^{-1}\|_F = \infty$ .*

*Proof.* From the norm bound (3.10), for sufficiently large  $n$  we must have  $\|\mathbf{Y}_t\|_{[L^2(\Omega)]^R} < \|\mathbf{Y}_0\|_{[L^2(\Omega)]^R} + n$  for any  $t \in (0, T_e)$ , and thus without loss of generality we assume

$$\tau_n = \tau_n^1 = \inf\{t > 0 \mid \|C_{\mathbf{Y}_t}^{-1}\|_F = \|C_{\mathbf{Y}_0}^{-1}\|_F + n\}, \quad n \in \mathbb{N}.$$

We will first show that  $T_e < \infty$  implies  $\limsup_{t \uparrow T_e} \|C_{\mathbf{Y}_t}^{-1}\|_F = \infty$ . We argue by contradiction, and assume  $T_e < \infty$  and  $\limsup_{t \uparrow T_e} \|C_{\mathbf{Y}_t}^{-1}\|_F < \infty$ . Then, we have  $\sup_{t \in [T_e - \delta, T_e]} \|C_{\mathbf{Y}_t}^{-1}\|_F < K < \infty$  for some  $K > 0$  and  $\delta > 0$ . But then since  $T_e < \infty$ , the continuity of  $t \mapsto \mathbf{Y}_t$  implies  $\|C_{\mathbf{Y}_{\tau_n}}^{-1}\|_F = \|C_{\mathbf{Y}_0}^{-1}\|_F + n$  and thus for any  $n$  sufficiently large we have

$$\|C_{\mathbf{Y}_{\tau_n}}^{-1}\|_F = \|C_{\mathbf{Y}_0}^{-1}\|_F + n < K,$$

which is absurd. Hence,  $T_e < \infty$  implies  $\limsup_{t \uparrow T_e} \|C_{\mathbf{Y}_t}^{-1}\|_F = \infty$ .

To conclude the proof we will show

$$\lim_{t \uparrow T_e} \|C_{\mathbf{Y}_t}^{-1}\|_F = \infty.$$

If this is false, then there exist a sequence  $t_m \uparrow T_e$  and  $\gamma > 0$  such that  $\|C_{\mathbf{Y}_{t_m}}^{-1}\|_F \leq \gamma$  for all  $m \geq 0$ . But since  $\limsup_{t \uparrow T_e} \|C_{\mathbf{Y}_t}^{-1}\|_F = \infty$  there is a sequence  $s_k \uparrow T_e$  such that  $\|C_{\mathbf{Y}_{s_k}}^{-1}\|_F \geq \gamma + 1$  for all  $k \geq 0$ . We take a sub-sequence  $(s_{k_m})_m$  so that  $t_m < s_{k_m}$  for all  $m$ . From the continuity of  $t \mapsto \|C_{\mathbf{Y}_t}^{-1}\|_F$  on  $[t_m, s_{k_m}]$ , there exists  $h_m \in [0, s_{k_m} - t_m]$  such that  $\|C_{\mathbf{Y}_{t_m+h_m}}^{-1}\|_F = \gamma + 1$ . Now, from (3.10) and [16, Lemma 3.5] we have for any  $m \geq 0$

$$1 \leq \|C_{\mathbf{Y}_{t_m+h_m}}^{-1}\|_F - \|C_{\mathbf{Y}_{t_m}}^{-1}\|_F \leq C_{T_e, R, \gamma} \|\mathbf{Y}_{t_m+h_m} - \mathbf{Y}_{t_m}\|_{[L^2(\Omega)]^R},$$

which is impossible since  $h_m \rightarrow 0$  as  $m \rightarrow \infty$  and  $\mathbf{Y}$  is continuous on  $[0, T_e)$ . Hence, the proof is complete.  $\square$

## 4.2 Extension up to the explosion time

Even when the explosion time  $T_e$  for the DO solution is finite, and thus  $\mathbf{U}$  and  $\mathbf{Y}$  cease to exist at  $T_e$ , we will show that the product  $X = \mathbf{U}^\top \mathbf{Y}$  nevertheless admits a continuous extension up to  $[0, T_e]$ , and beyond  $T_e$ , under suitable assumptions.

For any  $t' < t < T_e$ , from (2.17) we have

$$X_t - X_{t'} = \int_{t'}^t (\mathcal{P}_U(X_s) + \mathcal{P}_Y(X_s) - \mathcal{P}_U(X_s)\mathcal{P}_Y(X_s))a(s, X_s) ds + \int_{t'}^t \mathcal{P}_U(X_s)b(s, X_s) dW_s, \quad (4.4)$$

and, hence, the Itô's isometry implies

$$\mathbb{E}[|X_t - X_{t'}|^2] \leq 2T_e \int_{t'}^t |a(s, X_s)|^2 ds + 2 \int_{t'}^t \|b(s, X_s)\|_{\mathbb{F}}^2 ds.$$

From Assumption 2, the orthogonality of the rows of  $\mathbf{U}_t$  and the norm bound (3.10) of  $\mathbf{Y}_t$  it follows that

$$\mathbb{E}[|X_t - X_{t'}|^2] \leq C(t - t')$$

with  $C := 4 \max\{1, T_e\} C_{\text{lgb}}(1 + M(T_e))$ . Therefore,  $(X_t)_{0 \leq t < T_e}$  admits a unique extension  $X_{T_e} := \lim_{t \uparrow T_e} X_t \in L^2(\Omega; \mathbb{R}^d)$ .

Thus obtained  $(X_t)_{0 \leq t \leq T_e}$  is continuous from  $[0, T_e]$  to  $L^2(\Omega; \mathbb{R}^d)$ , but not necessarily path-wise a.s. continuous on  $[0, T_e]$ . It turns out that for initial data with suitable integrability, the DO approximation  $X_t$  actually admits a.s. Hölder continuous paths. Namely, we assume the following  $\mathbb{P}$ -integrability condition.

**Assumption 4.** The initial condition  $\mathbf{X}_0$  satisfies

$$\mathbb{E}[|\mathbf{X}_0|^{2k}] < +\infty, \quad \text{for some } k \in \mathbb{N}. \quad (4.5)$$

Notice that this condition is equivalent to

$$\mathbb{E}[|\mathbf{Y}_0|^{2k}] < \infty, \quad \text{for some } k \in \mathbb{N}, \quad (4.6)$$

where  $\mathbf{Y}_0 \in [L^2(\Omega)]^R$  is arbitrary such that  $X_0 = \mathbf{U}_0^\top \mathbf{Y}_0$  for  $\mathbf{U}_0 \in \mathbb{R}^{R \times d}$  orthonormal. Indeed, for any DO initial condition  $\mathbf{U}_0^\top \mathbf{Y}_0 = X_0$ , the orthogonality of the rows of  $\mathbf{U}_0$  implies

$$\mathbb{E}[|\mathbf{Y}_0|^{2k}] \leq \mathbb{E}\left[\left(\|\mathbf{U}_0^\top\|_2^2 |\mathbf{U}_0^\top \mathbf{Y}_0|^2\right)^k\right] \leq \mathbb{E}[|X_0|^{2k}] \leq \mathbb{E}\left[\left(\|\mathbf{U}_0^\top\|_2^2 |\mathbf{Y}_0|^2\right)^k\right] = \mathbb{E}[|\mathbf{Y}_0|^{2k}] < +\infty,$$

provided that the  $2k$ -th moment of either  $\mathbf{Y}_0$  or  $X_0$  (hence both) exists. Thus,  $\mathbb{E}[|X_0|^{2k}] = \mathbb{E}[|\mathbf{Y}_0|^{2k}]$ . Similarly, for any other DO initial condition  $\tilde{\mathbf{U}}_0^\top \tilde{\mathbf{Y}}_0 = X_0$  we have

$$\mathbb{E}[|\tilde{\mathbf{Y}}_0|^{2k}] = \mathbb{E}[|X_0|^{2k}] = \mathbb{E}[|\mathbf{Y}_0|^{2k}].$$

Hence, Assumption 4 is equivalent to (4.6), with an arbitrarily fixed stochastic basis.

Analogously, for all  $t > 0$  we have

$$\mathbb{E}[|X_t|^{2k}] = \mathbb{E}[|\mathbf{Y}_t|^{2k}] < +\infty. \quad (4.7)$$

The DO solution preserves the  $\mathbb{P}$ -integrability of the initial datum.

**Lemma 4.2** (Even order moments of the solution). *Let  $a$  and  $b$  satisfy Assumptions 1 and 2. Suppose that the DO solution for (2.1) exists on  $[0, T]$ . Suppose that Assumption 4 is fulfilled for some  $k \in \mathbb{N}$ . Then, for any  $t \in [0, T]$  we have*

$$\mathbb{E}[|\mathbf{Y}_t|^{2k}] \leq (\mathbb{E}[|\mathbf{Y}_0|^{2k}] + K_1(T)) K_2(T),$$

with  $K_1(T) := \left(3k^2 \frac{C_{\text{lgb}} T}{(1 + 1/C_{\text{lgb}})^{k-1}}\right)$  and  $K_2(T) := \exp\{6k^2 C_{\text{lgb}}(1 + 1/C_{\text{lgb}})T\}$ .



*Proof.* From the Itô formula,  $|\mathbf{Y}_t|^{2k}$  satisfies the following SDE (see also [19, Theorem 4.5.4]):

$$\begin{aligned} |\mathbf{Y}_t|^{2k} &= |\mathbf{Y}_0|^{2k} + \int_0^t 2k|\mathbf{Y}_s|^{2k-2} \left( \mathbf{U}_s a(s, \mathbf{U}_s^\top \mathbf{Y}_s) \right)^\top \mathbf{Y}_s + \\ &\quad + k|\mathbf{Y}_s|^{2k-2} \text{Tr} \left( \left( \mathbf{U}_s b(s, \mathbf{U}_s^\top \mathbf{Y}_s) \right) \left( \mathbf{U}_s b(s, \mathbf{U}_s^\top \mathbf{Y}_s) \right)^\top \right) \\ &\quad + k(2k-1)|\mathbf{Y}_s|^{2k-4} |\mathbf{Y}_s^\top \mathbf{U}_s b(s, \mathbf{U}_s^\top \mathbf{Y}_s)|^2 ds \\ &\quad + \int_0^t 2k|\mathbf{Y}_s|^{2k-2} (\mathbf{Y}_s)^\top \mathbf{U}_s b(s, \mathbf{U}_s^\top \mathbf{Y}_s) dW_s. \end{aligned}$$

We take the expectation of both sides, and, noting the progressive measurability of  $\mathbf{Y}_t$ ,  $a(t, \mathbf{U}_t^\top \mathbf{Y}_t)$ , and  $b(t, \mathbf{U}_t^\top \mathbf{Y}_t)$ , use Fubini's theorem to obtain

$$\begin{aligned} \mathbb{E}[|\mathbf{Y}_t|^{2k}] &\leq \mathbb{E}[|\mathbf{Y}_0|^{2k}] + \int_0^t \mathbb{E}[2k|\mathbf{Y}_s|^{2k-2} \left( \mathbf{U}_s a(s, \mathbf{U}_s^\top \mathbf{Y}_s) \right)^\top \mathbf{Y}_s] ds + \\ &\quad + \int_0^t \mathbb{E}[2k^2|\mathbf{Y}_s|^{2k-2} \text{Tr} \left( \left( \mathbf{U}_s b(s, \mathbf{U}_s^\top \mathbf{Y}_s) \right) \left( \mathbf{U}_s b(s, \mathbf{U}_s^\top \mathbf{Y}_s) \right)^\top \right)] ds \\ &\leq \mathbb{E}[|\mathbf{Y}_0|^{2k}] + \int_0^t \mathbb{E}[k|\mathbf{Y}_s|^{2k-2} (|a(s, \mathbf{U}_s^\top \mathbf{Y}_s)|^2 + |\mathbf{Y}_s|^2)] ds \\ &\quad + \int_0^t \mathbb{E}[2k^2|\mathbf{Y}_s|^{2k-2} \|b(s, \mathbf{U}_s^\top \mathbf{Y}_s)\|_{\mathbb{F}}^2] ds \\ &\leq \mathbb{E}[|\mathbf{Y}_0|^{2k}] + \int_0^t \mathbb{E}[k|\mathbf{Y}_s|^{2k-2} C_{\text{lgb}} (1 + (1 + 1/C_{\text{lgb}})|\mathbf{Y}_s|^2)] ds \\ &\quad + \int_0^t \mathbb{E}[2k^2|\mathbf{Y}_s|^{2k-2} C_{\text{lgb}} (1 + |\mathbf{Y}_s|^2)] ds \\ &\leq \mathbb{E}[|\mathbf{Y}_0|^{2k}] + \int_0^t \mathbb{E}[k|\mathbf{Y}_s|^{2k-2} C_{\text{lgb}} (1 + (1 + 1/C_{\text{lgb}})|\mathbf{Y}_s|^2)] ds \\ &\quad + \int_0^t \mathbb{E}[2k^2|\mathbf{Y}_s|^{2k-2} C_{\text{lgb}} (1 + (1 + 1/C_{\text{lgb}})|\mathbf{Y}_s|^2)] ds \\ &\leq \mathbb{E}[|\mathbf{Y}_0|^{2k}] + \int_0^t \mathbb{E}[3k^2|\mathbf{Y}_s|^{2k-2} C_{\text{lgb}} (1 + (1 + 1/C_{\text{lgb}})|\mathbf{Y}_s|^2)] ds \\ &\leq \mathbb{E}[|\mathbf{Y}_0|^{2k}] + 3k^2 \frac{C_{\text{lgb}}}{(\sqrt{(1 + 1/C_{\text{lgb}})})^{2k-2}} \\ &\quad \cdot \int_0^t \mathbb{E} \left[ \left( \sqrt{(1 + 1/C_{\text{lgb}})} \right)^{2k-2} |\mathbf{Y}_s|^{2k-2} \left( 1 + \left( \sqrt{(1 + 1/C_{\text{lgb}})} \right)^2 |\mathbf{Y}_s|^2 \right) \right] ds \end{aligned}$$

Then, by using the relation  $(1 + r^2)r^{2k-2} \leq 1 + 2r^{2k}$  for  $r \in \mathbb{R}_+$ , the statement follows by Gronwall's lemma:

$$\begin{aligned} \mathbb{E}[|\mathbf{Y}_t|^{2k}] &\leq \mathbb{E}[|\mathbf{Y}_0|^{2k}] + 3k^2 \frac{C_{\text{lgb}}}{(\sqrt{(1 + 1/C_{\text{lgb}})})^{2k-2}} \int_0^t \mathbb{E}[(1 + 2(\sqrt{(1 + 1/C_{\text{lgb}})})^{2k} |\mathbf{Y}_s|^{2k})] ds \\ &\leq \mathbb{E}[|\mathbf{Y}_0|^{2k}] + 3k^2 \frac{C_{\text{lgb}}}{(\sqrt{(1 + 1/C_{\text{lgb}})})^{2k-2}} T + 6k^2 C_{\text{lgb}} (1 + 1/C_{\text{lgb}}) \int_0^t \mathbb{E}[|\mathbf{Y}_s|^{2k}] ds \end{aligned}$$

□

With Lemma 4.2, we can now establish a Hölder-type bound on  $\mathbb{E}[|X_t - X_{t'}|^{2k}]$  for  $0 \leq t' < t < T_e$ , which by Kolmogorov-Chenstov theorem, implies the existence of an a.s. continuous version of  $X_t$ .

**Proposition 4.3.** *Suppose that the DO approximation  $X_t$  of (2.1) exists on  $[0, T_e)$ . Suppose that Assumption 4 holds for some  $k \in \mathbb{N}$ . If  $a$  and  $b$  satisfy Assumption 2, then there exists a constant  $C := C(X_0, R, k, T_e, C_{\text{lgb}}) > 0$  such that*

$$\mathbb{E}[|X_t - X_{t'}|^{2k}] \leq C(t - t')^k, \quad 0 \leq t' < t < T_e. \quad (4.8)$$

*Proof.* The proof follows closely the one in [19, Theorem 4.5.4]. First, notice that, in the view of (4.7), Lemma 4.2 gives an estimate of the  $2k$ -moments of the DO approximation  $X$  for some  $k \in \mathbb{N}$ . Now we use the inequality  $|r_1 + \dots + r_m|^p \leq m^{p-1} \sum_{i=1}^m |r_i|^p$  for  $r \in \mathbb{R}$  twice and Jensen's inequality to (4.4) to obtain

$$\begin{aligned} \mathbb{E}[|X_t - X_{t'}|^{2k}] &= \mathbb{E}\left[\left|\int_{t'}^t (\mathcal{P}_{U(X_s)} + \mathcal{P}_{Y(X_s)} - \mathcal{P}_{U(X_s)}\mathcal{P}_{Y(X_s)})a(s, X_s) ds \right. \right. \\ &\quad \left. \left. + \int_{t'}^t \mathcal{P}_{U(X_s)}b(s, X_s)dW_s\right|^{2k}\right] \\ &\leq 2^{2k-1}\mathbb{E}\left[\left|\int_{t'}^t (\mathcal{P}_{U(X_s)} + \mathcal{P}_{Y(X_s)} - \mathcal{P}_{U(X_s)}\mathcal{P}_{Y(X_s)})a(s, X_s) ds\right|^{2k}\right] \\ &\quad + 2^{2k-1}\mathbb{E}\left[\left|\int_{t'}^t \mathcal{P}_{U(X_s)}b(s, X_s)dW_s\right|^{2k}\right] \\ &\leq 2^{2k-1}(t - t')^{2k-1}\mathbb{E}\left[\int_{t'}^t (\|\mathcal{P}_{U(X_s)} + \mathcal{P}_{Y(X_s)} - \mathcal{P}_{U(X_s)}\mathcal{P}_{Y(X_s)}\|_2^2 |a(s, X_s)|^2)^k ds\right] \\ &\quad + 2^{2k-1}\mathbb{E}\left[\left|\int_{t'}^t \mathcal{P}_{U(X_s)}b(s, X_s)dW_s\right|^{2k}\right] \\ &\leq 2^{2k-1}(t - t')^{2k-1}\mathbb{E}\left[\int_{t'}^t |a(s, X_s)|^{2k} ds\right] + 2^{2k-1}\mathbb{E}\left[\left|\int_{t'}^t \mathcal{P}_{U(X_s)}b(s, X_s)dW_s\right|^{2k}\right] \end{aligned}$$

For the first term in the last inequality above, using the inequality  $(1 + r^2)^k \leq 2^{k-1}(1 + r^{2k})$  for  $r \in \mathbb{R}_+$ , Assumption 2 and Lemma 4.2, we get

$$\begin{aligned} \mathbb{E}\left[\int_{t'}^t |a(s, X_s)|^{2k} ds\right] &\leq \mathbb{E}\left[\int_{t'}^t C_{\text{lgb}}^k (1 + |X_s|^2)^k ds\right] \\ &\leq C_{\text{lgb}}^k 2^{k-1} \int_{t'}^t (1 + \mathbb{E}[|U_s^\top Y_s|^{2k}]) ds \\ &\leq C_{\text{lgb}}^k 2^{k-1} \int_{t'}^t (1 + \mathbb{E}[|Y_s|^{2k}]) ds \\ &\leq C_{\text{lgb}}^k 2^{k-1} \int_{t'}^t \left(1 + \left(\mathbb{E}[|Y_0|^{2k}] + K_1(T_e)\right) K_2(T_e)\right) ds \\ &= C_{\text{lgb}}^k 2^{k-1} \left(1 + \left(\mathbb{E}[|X_0|^{2k}] + K_1(T_e)\right) K_2(T_e)\right) (t - t'). \end{aligned} \quad (4.9)$$

For the second term, we will show

$$\mathbb{E}\left[\left|\int_{t'}^t \mathcal{P}_{U(X_s)}b(s, X_s)dW_s\right|^{2k}\right] \leq (t - t')^{k-1} [k(2k - 1)]^k \int_{t'}^t \mathbb{E}[\|b(s, X_s)\|_{\mathbb{F}}^{2k}] ds. \quad (4.10)$$

To see this, let  $I_t := \int_{t'}^t \mathcal{P}_{U(X_s)} b(s, X_s) dW_s$ . Then the Itô formula implies

$$d|I_t|^{2k} = 2k|I_t|^{2k-1} \text{sgn}(I_t) \mathcal{P}_{U(X_t)} b(t, X_t) dW_t + \frac{1}{2} 2k(2k-1) |I_t|^{2k-2} \|\mathcal{P}_{U(X_t)} b(t, X_t)\|_F^2 dt. \quad (4.11)$$

Taking the expectation of both sides, by Hölder's inequality we get

$$\begin{aligned} \mathbb{E} \left[ \left| \int_{t'}^t \mathcal{P}_{U(X_s)} b(s, X_s) dW_s \right|^{2k} \right] &= \mathbb{E}[|I_t|^{2k}] = 0 + k(2k-1) \int_{t'}^t \mathbb{E}[|I_s|^{2k-2} \|\mathcal{P}_{U(X_s)} b(s, X_s)\|_F^2] ds \\ &\leq k(2k-1) \int_{t'}^t \mathbb{E}[|I_s|^{2k}]^{1-\frac{1}{k}} \mathbb{E}[\|b(s, X_s)\|_F^{2k}]^{\frac{1}{k}} ds \end{aligned}$$

By Lemma 4.2, we have  $\int_{t'}^t \mathbb{E}|X_s|^{2k} ds = \int_{t'}^t \mathbb{E}|\mathbf{Y}_s|^{2k} ds < \infty$  and, hence, by  $\|AB\|_F \leq \|A\|_2 \|B\|_F$  and (2.3), it holds  $\mathbb{E} \left[ \left| \int_{t'}^t \mathcal{P}_{U(X_s)} b(s, X_s) dW_s \right|^{2k} \right] < \infty$ .

This implies that  $|I_t|^{2k} = \left| \int_{t'}^t \mathcal{P}_{U(X_s)} b(s, X_s) dW_s \right|^{2k}$  is a submartingale. Therefore, it follows

$$\mathbb{E} \left[ \left| \int_{t'}^t \mathcal{P}_{U(X_s)} b(s, X_s) dW_s \right|^{2k} \right] \leq k(2k-1) \mathbb{E}[|I_t|^{2k}]^{1-\frac{1}{k}} \int_{t'}^t \mathbb{E}[\|b(s, X_s)\|_F^{2k}]^{\frac{1}{k}} ds.$$

Raising both sides to the  $k$ -th power and dividing them by  $\mathbb{E}[|I_t|^{2k}]^{k-1}$ , we get

$$\mathbb{E} \left[ \left| \int_{t'}^t \mathcal{P}_{U(X_s)} b(s, X_s) dW_s \right|^{2k} \right] \leq [k(2k-1)]^k \left( \int_{t'}^t \mathbb{E}[\|b(s, X_s)\|_F^{2k}]^{\frac{1}{k}} ds \right)^k,$$

and thus (4.10) follows by Jensen's inequality. Therefore, bounding  $\int_{t'}^t \mathbb{E}[\|b(s, X_s)\|_F^{2k}] ds$  similarly to (4.9) we obtain

$$\mathbb{E} \left[ \left| \int_{t'}^t \mathcal{P}_{U(X_s)} b(s, X_s) dW_s \right|^{2k} \right] \leq (t-t')^k [k(2k-1)]^k C_{\text{lgb}}^k 2^{2k-1} \left( 1 + \left( \mathbb{E}[|X_0|^{2k}] + K_1(T_e) \right) K_2(T_e) \right).$$

Putting all together, (4.8) follows:

$$\begin{aligned} \mathbb{E}[|X_t - X_{t'}|^{2k}] &\leq 2^{2k-1} (t-t')^{2k-1} \mathbb{E} \left[ \int_{t'}^t |a(s, X_s)|^{2k} ds \right] + 2^{2k-1} \mathbb{E} \left[ \left| \int_{t'}^t \mathcal{P}_{U(X_s)} b(s, X_s) dW_s \right|^{2k} \right] \\ &\leq 2^{4k-2} (t-t')^{2k} C_{\text{lgb}}^k \left( 1 + \left( \mathbb{E}[|X_0|^{2k}] + K_1(T_e) \right) K_2(T_e) \right) \\ &\quad + 2^{4k-2} (t-t')^k [k(2k-1)]^k C_{\text{lgb}}^k \left( 1 + \left( \mathbb{E}[|X_0|^{2k}] + K_1(T_e) \right) K_2(T_e) \right) \\ &\leq (t-t')^k \left( T_e^k + [k(2k-1)]^k \right) \left[ 2^{4k-2} C_{\text{lgb}}^k \left( 1 + \left( \mathbb{E}[|X_0|^{2k}] + K_1(T_e) \right) K_2(T_e) \right) \right]. \end{aligned} \quad (4.12)$$

□

Having proved these results concerning boundedness of  $2k$ -moments of the DO solution, we are now ready to extend it up to the explosion time  $T_e$  with a.s. continuous paths.

**Theorem 4.4.** *Let  $a$  and  $b$  satisfy Assumptions 1 and 2. Let an DO approximation  $X$  on  $[0, T_e)$  be given. If  $X_0$  satisfies Assumption 4 for some integer  $k > 1$ , then  $X$  admits a unique continuous extension to  $[0, T_e]$ , almost surely. This extension is Hölder continuous on  $[0, T_e]$ , and satisfies  $\mathbb{E}[|X_t|^{2k}] < \infty$  for all  $t \in [0, T_e]$ .*

*Proof.* By Assumption 4, Proposition 4.3 holds and  $(X_t)_{0 \leq t < T_e}$  admits an extension  $X_{T_e} := \lim_{t \uparrow T_e} X_t \in L^{2k}(\Omega; \mathbb{R}^d)$  that is unique in  $L^{2k}(\Omega; \mathbb{R}^d)$ . From  $\|X_t - X_{t'}\|_{L^2(\Omega; \mathbb{R}^d)} \leq \|X_t - X_{t'}\|_{L^{2k}(\Omega; \mathbb{R}^d)}$  for  $0 \leq t' < t < T_e$ , the extension is unique in  $L^p(\Omega; \mathbb{R}^d)$  for  $p \in [2, 2k]$ . By construction of  $X_{T_e}$ , we have

$$\mathbb{E}[|X_t - X_{t'}|^{2k}] \leq C(t - t')^k, \quad 0 \leq t' < t \leq T_e. \quad (4.13)$$

Therefore, by Kolmogorov-Chenstov Theorem (see for example [14]) there exists a version  $(\tilde{X}_t)_{t \in [0, T_e]}$  of the DO approximation that is  $q$ -Hölder continuous for all  $0 < q < \frac{k-1}{2k}$ . By construction,  $(\tilde{X}_t)_{t \in [0, T_e]}$  is a process with bounded  $2k$ -moments.

Finally, let us see that  $(X_t)_{t \in [0, T_e]}$  and  $(\tilde{X}_t)_{t \in [0, T_e]}$  are indistinguishable. Indeed, by construction of  $\tilde{X}$  we have  $X_{T_e} = \tilde{X}_{T_e}$  a.s. Moreover,  $(X_t)_{t \in [0, T_e]}$  and  $(\tilde{X}_t)_{t \in [0, T_e]}$  have a.s. continuous paths and are versions of each other, and thus indistinguishable. Hence,  $(X_t)_{t \in [0, T_e]}$  and  $(\tilde{X}_t)_{t \in [0, T_e]}$  are indistinguishable. This completes the proof.  $\square$

Together with Proposition 4.1, Theorem 4.4 gives us an insight into how to continue a DO approximation *beyond* the explosion time  $T_e$ . Let a DO approximation  $X$  of (2.1) with explosion time  $T_e$  that satisfies (4.5) be given. Suppose  $\dim(\text{Im}(\mathbb{E}[\mathcal{X}_t \cdot])) = R$  for all  $t \in [0, T_e)$ , with a positive integer  $R$ . By Theorem 4.4,  $X$  can be continuously extended up to  $T_e$ , while from Proposition 4.1 we know  $\lim_{t \uparrow T_e} \|C_{\mathbf{Y}_t}^{-1}\|_{\text{F}} = \infty$ . This implies that  $\dim(\text{Im}(\mathbb{E}[\mathcal{X}_{T_e} \cdot])) = R' < R$ . Therefore, one can extend our DO approximation  $X$  continuously in  $t$  beyond  $T_e$  by considering the DO system (2.6)–(2.7) with initial datum  $X_{T_e}$ . The corresponding DO solution satisfies  $\mathbf{U}_t \in \mathbb{R}^{R' \times d}$  and  $\mathbf{Y}_t \in [L^2(\Omega)]^{R'}$  for  $t \in [T_e, T_e + T_e')$ , where  $T_e'$  is the new explosion time for  $(\mathbf{U}_t, \mathbf{Y}_t)$ .

### 4.3 The case of uniformly positive-definite diffusion

In the previous section, we have shown that by assuming the boundedness of the  $2k$  moments of the initial condition of the DO approximation for  $k > 1$ , the DO approximation can be extended up to the explosion time  $T_e$ . It turns out that, under the condition that we will introduce in the following, the explosion time  $T_e$  is never finite; as a result, rank- $R$  DO solution exists globally.

A sufficient condition is that the diffusion matrix  $b(t, x)b(t, x)^\top$  is positive definite with a lower bound on the smallest eigenvalue, where the bound is uniform in  $t$  and  $x$ . This condition turns out to ensure that the smallest eigenvalue of the Gram matrix  $\mathbb{E}[\mathbf{Y}_t \mathbf{Y}_t^\top]$  remains bounded below uniformly in  $t$  and  $x$ , which in turn guarantees the global existence of the DO solution. In the following, we use the notation  $A \succ B$  (respectively  $A \succeq B$ ) with  $A, B$  square matrices to indicate that  $A - B$  is positive definite (respectively positive semidefinite).

**Proposition 4.5.** *Suppose the rank  $R$  DO solution  $(\mathbf{U}, \mathbf{Y})$  exists on  $[0, T]$ . Assume moreover that there exist  $\sigma_{\mathbf{Y}_0}, \sigma_B > 0$  such that*

$$C_{\mathbf{Y}_0} := \mathbb{E}[\mathbf{Y}_0 \mathbf{Y}_0^\top] \succeq \sigma_{\mathbf{Y}_0} I_{R \times R}; \quad (4.14)$$

$$b(t, x)b(t, x)^\top \succeq \sigma_B I_{d \times d}, \quad \text{for any } t \in [0, T] \text{ and for any } x \in \mathbb{R}^d. \quad (4.15)$$

Then for all  $t \in [0, T]$  we have

$$C_{\mathbf{Y}_t} \succeq \min\left\{\sigma_{\mathbf{Y}_0}; \frac{\sigma_B^2}{4C_{\text{lgb}}(1 + M)}\right\} I_{R \times R}, \quad (4.16)$$

where  $M = M(T)$  is defined in Lemma 3.5.

*Proof.* By introducing the shorthand notation  $a_t = a(t, X_t) \in \mathbb{R}^d$  and  $b_t = b(t, X_t) \in \mathbb{R}^{d \times m}$ , the  $k$ -th component  $Y_t^k$  of  $\mathbf{Y}_t \in [L^2(\Omega)]^R$  satisfies

$$dY_t^k = \sum_{i=1}^d U_t^{ki} a_t^i dt + \sum_{l=1}^m \sum_{r=1}^d U_t^{kr} b_t^{rl} dW_t^l$$

Hence, using Itô's formula for  $1 \leq j, k \leq R$ , we have

$$\begin{aligned} d(Y_t^j Y_t^k) &= d(Y_t^j) Y_t^k + Y_t^j d(Y_t^k) + \sum_{l=1}^m \sum_{i=1}^d U_t^{ji} b_t^{il} \sum_{r=1}^d U_t^{kr} b_t^{rl} dt \\ &= \left( \sum_{i=1}^d U_t^{ji} a_t^i dt + \sum_{l=1}^m \sum_{i=1}^d U_t^{ji} b_t^{il} dW_t^l \right) Y_t^k \\ &\quad + Y_t^j \left( \sum_{i=1}^d U_t^{ki} a_t^i dt + \sum_{l=1}^m \sum_{i=1}^d U_t^{kr} b_t^{rl} dW_t^l \right) \\ &\quad + \sum_{l=1}^m \sum_{i=1}^d U_t^{ji} b_t^{il} \sum_{r=1}^d U_t^{kr} b_t^{rl} dt \end{aligned}$$

Hence,

$$d(\mathbf{Y}_t \mathbf{Y}_t^\top) = (d\mathbf{Y}_t) \mathbf{Y}_t^\top + \mathbf{Y}_t (d\mathbf{Y}_t)^\top + \mathbf{U}_t b_t (\mathbf{U}_t b_t)^\top dt$$

and taking the expectation of both sides yields

$$\frac{d\mathbb{E}[(\mathbf{Y}_t \mathbf{Y}_t^\top)]}{dt} = \mathbb{E}[\mathbf{U}_t a_t (\mathbf{Y}_t^\top)] + \mathbb{E}[\mathbf{Y}_t (\mathbf{U}_t a_t)^\top] + \mathbb{E}[\mathbf{U}_t b_t b_t^\top \mathbf{U}_t^\top].$$

We now aim at analyzing the smallest eigenvalue of  $C_{\mathbf{Y}_t} := \mathbb{E}[(\mathbf{Y}_t \mathbf{Y}_t^\top)]$  through the Rayleigh quotient. For any unit vector  $v \in \mathbb{R}^R$ , we have

$$v^\top \frac{d\mathbb{E}[(\mathbf{Y}_t \mathbf{Y}_t^\top)]}{dt} v = 2v^\top \mathbb{E}[\mathbf{U}_t a_t (\mathbf{Y}_t^\top)] v + v^\top \mathbb{E}[\mathbf{U}_t b_t b_t^\top \mathbf{U}_t^\top] v. \quad (4.17)$$

Thanks to (4.15), the last term can be bounded below as

$$v^\top \mathbb{E}[\mathbf{U}_t b_t b_t^\top \mathbf{U}_t^\top] v = v^\top \mathbf{U}_t \mathbb{E}[b_t b_t^\top] \mathbf{U}_t^\top v \geq v^\top \mathbf{U}_t \sigma_B I_{R \times R} \mathbf{U}_t^\top v \geq \sigma_B.$$

The first term can be bounded above as

$$\begin{aligned} |v^\top \mathbb{E}[\mathbf{Y}_t (\mathbf{U}_t a_t)^\top] v| &= |\mathbb{E}[v^\top \mathbf{Y}_t a_t^\top] \mathbf{U}_t^\top v| \leq |\mathbb{E}[v^\top \mathbf{Y}_t a_t^\top]| \\ &\leq \frac{1}{2\varepsilon} \mathbb{E}[|v^\top \mathbf{Y}_t|^2] + \frac{\varepsilon}{2} \mathbb{E}[|a_t|^2] \\ &\leq \frac{1}{2\varepsilon} \mathbb{E}[v^\top \mathbf{Y}_t \mathbf{Y}_t^\top v] + \frac{\varepsilon}{2} C_{\text{lgb}} (1 + \mathbb{E}[|\mathbf{Y}_t|^2]) \\ &\leq \frac{1}{2\varepsilon} v^\top \mathbb{E}[\mathbf{Y}_t \mathbf{Y}_t^\top] v + \frac{\varepsilon}{2} C_{\text{lgb}} (1 + M) \quad \text{for any } \varepsilon > 0. \end{aligned}$$

Taking  $\varepsilon = \frac{\sigma_B}{2C_{\text{lgb}}(1+M)}$  leads to the following estimate on the derivative of  $A_t := v^\top C_{\mathbf{Y}_t} v = v^\top \mathbb{E}[(\mathbf{Y}_t \mathbf{Y}_t^\top)] v$ :

$$\frac{d}{dt}A_t \geq -\frac{1}{\varepsilon}A_t + \frac{\sigma_B}{2}, \quad (4.18)$$

from which we deduce

$$A_t \geq \frac{\varepsilon\sigma_B}{2}(e^{\frac{t}{\varepsilon}} - 1)e^{-\frac{t}{\varepsilon}} + A_0e^{-\frac{t}{\varepsilon}}.$$

Noting that (4.14) implies  $A_0 \geq \sigma_{\mathbf{Y}_0}$ , we conclude

$$\begin{aligned} v^\top \mathbb{E}[\mathbf{Y}_t \mathbf{Y}_t^\top] v &\geq \frac{\varepsilon\sigma_B}{2} + (\sigma_{\mathbf{Y}_0} - \frac{\varepsilon\sigma_B}{2})e^{-\frac{t}{\varepsilon}} \\ &\geq \min\{\sigma_{\mathbf{Y}_0}, \frac{\sigma_B^2}{4C_{\text{lgb}}(1+M)}\} > 0 \quad \text{for any } t \in [0, T]. \end{aligned}$$

□

As a consequence of the proposition above, the following global existence result is obtained.

**Theorem 4.6** (Global Existence of DO solution). *Let Assumptions 1–3 hold. Suppose that the assumptions of Proposition 4.5 hold. Then, the DO solution exists for all  $t \geq 0$ .*

*Proof.* From Theorems 3.4 and 3.6, the rank  $R$  solution uniquely exists up to a certain time  $T > 0$ . Denote by  $T_e$  its explosion time and suppose, to obtain a contradiction,  $T_e < +\infty$ . Then, from Proposition 4.1 we have  $\lim_{t \uparrow T_e} \|C_{\mathbf{Y}_t}^{-1}\|_{\text{F}} = +\infty$ . But by Proposition 4.5 the Rayleigh quotient  $A_t = v^\top C_{\mathbf{Y}_t} v$  with  $v \in \mathbb{R}^R$  satisfies

$$A_t \geq \min\{\sigma_{\mathbf{Y}_0}; \frac{\sigma_B^2}{4C_{\text{lgb}}(1+M(T_e))}\} > 0,$$

and, hence, for some constant  $\bar{k} > 0$  we have  $\|C_{\mathbf{Y}_t}^{-1}\|_{\text{F}} \leq \bar{k}$  for all  $t \in [0, T_e)$ . This contradicts  $\lim_{t \uparrow T_e} \|C_{\mathbf{Y}_t}^{-1}\|_{\text{F}} = +\infty$ . Therefore,  $T_e = +\infty$ . □

**Remark 4.7.** It is worth noticing that (4.15) is satisfied in the case of additive non-degenerate noise.

## 5 Conclusion

In this work, we achieved to set a rigorous DO setting for SDEs under the conditions that the studied drift and diffusion satisfy a Lipschitz condition and a linear-growth bound. First, we (re-)derived the equations which characterize the evolution of the deterministic and stochastic modes, in a DO formulation, and showed how these can be re-interpreted as a projected dynamics leading to a DLRA formulation. Our derivation makes use of the Itô's formula and avoids the direct use of time derivatives. We proved local-existence and uniqueness of the DO formulation and analyzed the possibility of extending the solution up to and beyond the explosion time. Finally, we gave a sufficient condition that assures the global existence of the DLR approximation.

One natural development of this work would be to extend this DO framework and well-posedness results to accommodate weaker assumptions on the drift and diffusion (e.g. local Lipschitzianity and/or weak monotonicity).

Furthermore, it would be interesting to build a DLR formulation as a fully projecting dynamics as in (2.18) by giving a rigorous meaning to the term  $\mathcal{P}_{\mathcal{Y}(x_t)}[b(t, \mathcal{X}_t) dW_t]$ , possibly in a distributional sense. In case one manages to achieve this goal, the direct connection to the standard DLRA formulation for deterministic or random equations would allow us to apply numerical projector splitting schemes, which have been shown to perform very well in those contexts [5, 25].

## Acknowledgements

Yoshihito Kazashi acknowledges the financial support of the University of Strathclyde through a Faculty of Science Starter Grant. This work has also been supported by the Swiss National Science Foundation under the Project n. 200518 “Dynamical low rank methods for uncertainty quantification and data assimilation”.

## A Appendix

For a Hilbert space  $(H, \langle \cdot, \cdot \rangle)$ , denote by  $[H]^R$  the product Hilbert space equipped with the norm  $\|\mathbf{Y}\|_{[H]^R} = \sqrt{\sum_{j=1}^R \|Y^j\|_H^2}$  for  $\mathbf{Y} = (Y^j) \in [H]^R$ . For  $\mathbf{Y} \in [H]^R$  let  $Z_{\mathbf{Y}}$  be the Gram matrix

$$Z_{\mathbf{Y}} := \left( \langle Y^j, Y^k \rangle \right)_{j,k=1,\dots,R} \in \mathbb{R}^{R \times R}.$$

If  $Z_{\mathbf{Y}_0}$  is invertible, then for  $\mathbf{Y}$  close to  $\mathbf{Y}_0$ ,  $Z_{\mathbf{Y}}$  is also invertible. The following result makes this intuition precise.

**Proposition A.1.** *Suppose that  $\mathbf{Y}_0 \in [H]^R$  has linearly independent components  $Y_0^j$ ,  $j = 1, \dots, R$  in  $H$ , and that  $\|\mathbf{Y}_0\|_H \leq \rho$  and  $\|Z_{\mathbf{Y}_0}^{-1}\|_F \leq \gamma$  hold for  $\rho, \gamma > 0$ . Then, there exists  $\eta := \eta(\rho, \gamma) > 0$  defined as*

$$\eta(\rho, \gamma) := -\rho + \sqrt{\rho^2 + \frac{1}{2\gamma}} \tag{A.1}$$

such that we have

$$\|Z_{\mathbf{Y}}^{-1}\|_F \leq 2\gamma, \quad \text{for any } \mathbf{Y} \in B_\eta(\mathbf{Y}_0),$$

and  $\eta(\rho, \gamma)$  is decreasing in both  $\rho$  and  $\gamma$ . Here,  $B_\eta(\mathbf{Y}_0)$  is the open ball in  $[H]^R$  of radius  $\eta$  around  $\mathbf{Y}_0$ .

*Proof.* Any  $\mathbf{Y} \in B_\eta(\mathbf{Y}_0)$  may be written as  $\mathbf{Y} = \mathbf{Y}_0 + r\boldsymbol{\delta}$  with  $r < \eta$  and  $\|\boldsymbol{\delta}\|_{[H]^R} = 1$ . We will derive an upper bound  $\eta$  on  $r$  that guarantees  $\|Z_{\mathbf{Y}}^{-1}\|_F \leq 2\gamma$ .

Define  $\mathbf{Y}(s) := \mathbf{Y}_0 + s\boldsymbol{\delta}$  and let  $f(s) := Z_{\mathbf{Y}(s)}^{-1}$ . Then, we have

$$\frac{d}{ds} f(s) = -f(s) \left( \underbrace{\left( \langle \delta^j, Y(s)^k \rangle \right)_{j,k=1,\dots,R}}_{:= Z_{\boldsymbol{\delta}, \mathbf{Y}(s)}} + \underbrace{\left( \langle Y(s)^j, \delta^k \rangle \right)_{j,k=1,\dots,R}}_{:= Z_{\mathbf{Y}(s), \boldsymbol{\delta}}} \right) f(s).$$

Therefore, it holds

$$\begin{aligned}
 \frac{d}{ds} \|f(s)\|_{\mathbb{F}}^2 &= \frac{d}{ds} \text{Tr}(f(s)f^\top(s)) \\
 &= -\text{Tr}(f(s)(Z_{\delta, \mathbf{Y}(s)} + Z_{\mathbf{Y}(s), \delta})f(s)f^\top(s)) - \text{Tr}(f(s)f^\top(s)(Z_{\mathbf{Y}(s), \delta} + Z_{\delta, \mathbf{Y}(s)})f^\top(s)) \\
 &\leq 2\|Z_{\delta, \mathbf{Y}(s)} + Z_{\mathbf{Y}(s), \delta}\|_{\mathbb{F}} \|f^3(s)\|_{\mathbb{F}} \\
 &\leq 2(\|Z_{\delta, \mathbf{Y}(s)}\|_{\mathbb{F}} + \|Z_{\mathbf{Y}(s), \delta}\|_{\mathbb{F}}) \|f^3(s)\|_{\mathbb{F}} \\
 &\leq 4(\|\delta\|_{[H]^R} \|\mathbf{Y}(s)\|_{[H]^R}) \|f(s)\|_{\mathbb{F}}^3 \\
 &\leq 4(\rho + s) \|f(s)\|_{\mathbb{F}}^3.
 \end{aligned}$$

Let  $E_s = \|f(s)\|_{\mathbb{F}}^2$  so that  $E_s^{-\frac{3}{2}} \frac{dE_s}{ds} \leq 4(\rho + s)$  and

$$\int_0^r 2 \frac{d(-E_s^{-1/2})}{ds} ds = \int_0^r E_s^{-\frac{3}{2}} \frac{dE_s}{ds} ds \leq 4\rho r + 2r^2,$$

which implies  $-2E_r^{-\frac{1}{2}} + 2E_0^{-\frac{1}{2}} \leq 4\rho r + 2r^2$ . Now we use  $E_0 \leq \gamma^2$  to obtain the bound  $\gamma^{-1} - 2\rho r - r^2 \leq E_r^{-\frac{1}{2}}$ . For any  $0 < r < -\rho + \sqrt{\rho^2 + \frac{1}{\gamma}}$ , we have  $\frac{1}{\gamma} - 2\rho r - r^2 > 0$  and thus this bound yields

$$\|Z_{\mathbf{Y}_r}^{-1}\|_{\mathbb{F}} \leq \frac{1}{\frac{1}{\gamma} - 2\rho r - r^2} = \frac{1}{\frac{1}{\gamma} + \rho^2 - (\rho + r)^2}.$$

Then,  $\|Z_{\mathbf{Y}_r}^{-1}\|_{\mathbb{F}}^2 \leq 4\gamma^2$  is guaranteed by the condition  $0 < r \leq -\rho + \sqrt{\rho^2 + \frac{1}{2\gamma}}$ , which also implies  $\frac{1}{\gamma} - 2\rho r - r^2 > 0$  above. Finally, the function  $\eta(\rho, \gamma)$ , defined as in (A.1), is decreasing in  $\rho$  and  $\gamma$ :

$$\begin{aligned}
 \frac{\partial \eta}{\partial \rho} &= \frac{\rho}{\sqrt{\rho^2 + \frac{1}{2\gamma}}} - 1 < 0, \quad \forall \rho, \gamma > 0; \\
 \frac{\partial \eta}{\partial \gamma} &= -\frac{1}{4\gamma^2} \frac{1}{\sqrt{\rho^2 + \frac{1}{2\gamma}}} < 0, \quad \forall \rho, \gamma > 0.
 \end{aligned}$$

□

Recall that a finite rank function  $X \in L^2(\Omega; H)$  admits a representation reminiscent of the singular value decomposition as in (A.2) below; see [16, Lemma 2.1], as well as Lemma 2.5 and Remark 2.6. For this decomposition, in Proposition A.3 we show that projector-valued mappings, where the projection is onto the space spanned by the spatial basis of this decomposition, are Lipschitz continuous. In order to achieve this goal, we need the following Lemma A.2. The proofs of Lemma A.2 and Proposition A.3 follow closely [3, Proof of Lemma A.2]; see also [43, Lemmata 4.1 and 4.2] for a finite dimensional version. However, unlike [3, Lemma A.2] we do not assume  $L^2(\Omega; H)$  to be separable.

**Lemma A.2.** *Consider  $X, \hat{X} \in L^2(\Omega; H)$  that have the following representations*

$$X = \sum_{j=1}^R \sigma_j U_j V_j, \quad \hat{X} = \sum_{j=1}^R \hat{\sigma}_j \hat{U}_j \hat{V}_j, \tag{A.2}$$



with  $\sigma_j, \hat{\sigma}_j > 0$ , and  $U_j, \hat{U}_j \in H$ ,  $V_j, \hat{V}_j \in L^2(\Omega; \mathbb{R})$  all orthonormal in their respective Hilbert spaces, for  $j = 1, \dots, R$ . Here,  $\sigma_j$  and  $\hat{\sigma}_j$ ,  $j = 1, \dots, R$ , are ordered in the descending order. Define the projections  $P_U = \sum_{j=1}^R \langle U_j, \cdot \rangle_H U_j : L^2(\Omega; H) \rightarrow L^2(\Omega; H)$ ,  $P_V = \sum_{j=1}^R \mathbb{E}[V_j \cdot] V_j : L^2(\Omega; H) \rightarrow L^2(\Omega; H)$ ,  $P_{U^\perp} := \text{id} - P_U$ , and  $P_{V^\perp} := \text{id} - P_V$ ; and similarly for  $P_{\hat{U}}$ ,  $P_{\hat{U}^\perp}$ ,  $P_{\hat{V}}$ , and  $P_{\hat{V}^\perp}$ . Then, we have

$$\begin{aligned} \|P_U - P_{\hat{U}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} &= \|P_{\hat{U}^\perp} P_U\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} = \|P_{\hat{U}} P_{U^\perp}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)}; \\ \|P_V - P_{\hat{V}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} &= \|P_{\hat{V}^\perp} P_V\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} = \|P_{\hat{V}} P_{V^\perp}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)}. \end{aligned} \quad (\text{A.3})$$

*Proof.* The mappings  $P_U$  and  $P_V$  are orthogonal projections. Indeed, the idempotency of  $P_U$  and  $P_V$  can be easily checked and the  $L^2(\Omega; H)$ -orthogonality of  $P_U$  follows from its  $H$ -orthogonality. To show the orthogonality of  $P_V$ , first consider  $z, w \in L^2(\Omega; H)$ . Noting that  $\langle \mathbb{E}[V_k w], \cdot \rangle_H : H \rightarrow H$  is a continuous linear functional on  $H$  for  $k = 1, \dots, R$ , from [23, Theorem 8.13] we have:

$$\begin{aligned} \langle z, P_V w \rangle_{L^2(\Omega; H)} &= \mathbb{E}[\langle z, \sum_{k=1}^R \mathbb{E}[w V_k] V_k \rangle_H] = \sum_{k=1}^R \mathbb{E}[\langle z, \mathbb{E}[w V_k] V_k \rangle_H] \\ &= \sum_{k=1}^R \langle \mathbb{E}[z V_k], \mathbb{E}[w V_k] \rangle_H = \sum_{k=1}^R \mathbb{E}[\langle \mathbb{E}[z V_k], w V_k \rangle_H] = \langle P_V z, w \rangle_{L^2(\Omega; H)}. \end{aligned}$$

Similarly,  $P_{\hat{U}}$  and  $P_{\hat{V}}$  are orthogonal projections.

Let us prove the second identity in (A.3). For  $f \in L^2(\Omega; H)$ , we have

$$(P_V - P_{\hat{V}})f = (P_{\hat{V}} + P_{\hat{V}^\perp})(P_V - P_{\hat{V}})(P_V + P_{V^\perp})f = (P_{\hat{V}^\perp} P_V - P_{\hat{V}} P_{V^\perp})f, \quad (\text{A.4})$$

As  $P_{\hat{V}}$  is an orthogonal projector,  $\mathbb{E}[\langle P_{\hat{V}^\perp} f, P_{\hat{V}} g \rangle_H] = 0$  for any  $f, g \in L^2(\Omega; H)$ . Hence, together with (A.4), for all  $f \in L^2(\Omega; H)$  one has

$$\|(P_V - P_{\hat{V}})f\|_{L^2(\Omega; H)}^2 = \|P_{\hat{V}^\perp} P_V f\|_{L^2(\Omega; H)}^2 + \|P_{\hat{V}} P_{V^\perp} f\|_{L^2(\Omega; H)}^2. \quad (\text{A.5})$$

Moreover, for any  $f \in L^2(\Omega; H)$  it also holds

$$\|f\|_{L^2(\Omega; H)}^2 = \|P_V f\|_{L^2(\Omega; H)}^2 + \|P_{V^\perp} f\|_{L^2(\Omega; H)}^2, \quad (\text{A.6})$$

and

$$f = \|P_V f\|_{L^2(\Omega; H)} \tilde{f}_1 + \|P_{V^\perp} f\|_{L^2(\Omega; H)} \tilde{f}_2 \quad (\text{A.7})$$

with unit vectors  $\tilde{f}_1 := P_V f / \|P_V f\|_{L^2(\Omega; H)}$  and  $\tilde{f}_2 := P_{V^\perp} f / \|P_{V^\perp} f\|_{L^2(\Omega; H)}$ . From (A.5), one has

$$\begin{aligned} \|(P_V - P_{\hat{V}})f\|_{L^2(\Omega; H)}^2 &= \|P_V f\|_{L^2(\Omega; H)}^2 \|P_{\hat{V}^\perp} P_V \tilde{f}_1\|_{L^2(\Omega; H)}^2 + \|P_{V^\perp} f\|_{L^2(\Omega; H)}^2 \|P_{\hat{V}} P_{V^\perp} \tilde{f}_2\|_{L^2(\Omega; H)}^2 \\ &\leq (\|P_V f\|_{L^2(\Omega; H)}^2 + \|P_{V^\perp} f\|_{L^2(\Omega; H)}^2) \max \{ \|P_{\hat{V}^\perp} P_V \tilde{f}_1\|_{L^2(\Omega; H)}^2, \|P_{\hat{V}} P_{V^\perp} \tilde{f}_2\|_{L^2(\Omega; H)}^2 \}. \end{aligned}$$

Using (A.6) and  $\|\tilde{f}_1\|_{L^2(\Omega; H)} = \|\tilde{f}_2\|_{L^2(\Omega; H)} = 1$ , we obtain

$$\|P_V - P_{\hat{V}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} \leq \max \{ \|P_{\hat{V}^\perp} P_V\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)}, \|P_{\hat{V}} P_{V^\perp}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} \}.$$

To show the opposite inequality we first notice that

$$\begin{aligned} \|P_{\hat{V}^\perp} P_V\|_{L^2(\Omega;H) \rightarrow L^2(\Omega;H)}^2 &= \sup_{\|f\|_{L^2(\Omega;H)}=1} \|P_{\hat{V}^\perp} P_V f\|_{L^2(\Omega;H)}^2 \\ &= \sup_{\|f\|_{L^2(\Omega;H)}=1, P_{V^\perp} f=0} \|P_{\hat{V}^\perp} P_V f\|_{L^2(\Omega;H)}^2. \end{aligned}$$

Therefore, from (A.5) one has

$$\begin{aligned} \|P_V - P_{\hat{V}}\|_{L^2(\Omega;H) \rightarrow L^2(\Omega;H)}^2 &\geq \sup_{\|f\|_{L^2(\Omega;H)}=1, P_{V^\perp} f=0} \|P_{\hat{V}^\perp} P_V f\|_{L^2(\Omega;H)}^2 \\ &= \sup_{\|f\|_{L^2(\Omega;H)}=1} \|P_{\hat{V}^\perp} P_V f\|_{L^2(\Omega;H)}^2 = \|P_{\hat{V}^\perp} P_V\|_{L^2(\Omega;H) \rightarrow L^2(\Omega;H)}^2. \end{aligned}$$

Similarly,  $\|P_V - P_{\hat{V}}\|_{L^2(\Omega;H) \rightarrow L^2(\Omega;H)} \geq \|P_{\hat{V}} P_{V^\perp}\|_{L^2(\Omega;H) \rightarrow L^2(\Omega;H)}$ . Hence,

$$\|P_V - P_{\hat{V}}\|_{L^2(\Omega;H) \rightarrow L^2(\Omega;H)} = \max \left\{ \|P_{\hat{V}^\perp} P_V\|_{L^2(\Omega;H) \rightarrow L^2(\Omega;H)}, \|P_{\hat{V}} P_{V^\perp}\|_{L^2(\Omega;H) \rightarrow L^2(\Omega;H)} \right\}.$$

Finally, let us prove that  $\|P_{\hat{V}^\perp} P_V\|_{L^2(\Omega;H) \rightarrow L^2(\Omega;H)} = \|P_{\hat{V}} P_{V^\perp}\|_{L^2(\Omega;H) \rightarrow L^2(\Omega;H)}$ . First define the vector space  $H^R := [H]^R$  and endow it with the norm  $\|z\|_{H^R} = \sqrt{\sum_{j=1}^R \|z_j\|_H^2}$  for  $z = (z_j)_{j=1, \dots, R} \in H^R$ . Thanks to the orthogonality of  $(V_j)_j$  and the properties of orthogonal projectors, the following chain of equalities holds:

$$\begin{aligned} \|P_{\hat{V}^\perp} P_V\|_{L^2(\Omega;H) \rightarrow L^2(\Omega;H)}^2 &= \sup_{\|f\|_{L^2(\Omega;H)}=1} \left\| P_{\hat{V}^\perp} \sum_{j=1}^R \mathbb{E}[V_j f] V_j \right\|_{L^2(\Omega;H)}^2 \\ &= \sup_{\|f\|_{L^2(\Omega;H)}=1, P_{V^\perp} f=0} \left\| P_{\hat{V}^\perp} \sum_{j=1}^R \mathbb{E}[V_j f] V_j \right\|_{L^2(\Omega;H)}^2 \\ &= \sup_{x \in H^R, \|x\|_{H^R}=1} \left\| P_{\hat{V}^\perp} \sum_{j=1}^R x_j V_j \right\|_{L^2(\Omega;H)}^2 \\ &= \sup_{x \in H^R, \|x\|_{H^R}=1} \left( \left\| \sum_{j=1}^R x_j V_j \right\|_{L^2(\Omega;H)}^2 - \left\| P_{\hat{V}} \left( \sum_{j=1}^R x_j V_j \right) \right\|_{L^2(\Omega;H)}^2 \right) \\ &= 1 - \inf_{x \in H^R, \|x\|_{H^R}=1} \left\| \sum_{i=1}^R \hat{V}_i \mathbb{E} \left[ \hat{V}_i \left( \sum_{j=1}^R x_j V_j \right) \right] \right\|_{L^2(\Omega;H)}^2 \\ &= 1 - \inf_{x \in H^R, \|x\|_{H^R}=1} \sum_{i=1}^R \left\| \sum_{j=1}^R x_j \underbrace{\mathbb{E}[\hat{V}_i V_j]}_{=: (G)_{ij}} \right\|_H^2 \\ &= 1 - \inf_{x \in H^R, \|x\|_{H^R}=1} \|Gx\|_{H^R}^2. \end{aligned} \tag{A.8}$$

Considering the Stiefel manifold  $\text{St}(H; R) = \{(q_1, \dots, q_R) : q_i \in H, \langle q_i, q_j \rangle_H =$

$\delta_{ij}, \forall i, j = 1, \dots, R$  the following equality holds

$$\begin{aligned}
 \inf_{x \in H^R, \|x\|_{H^R}=1} \|Gx\|_{H^R}^2 &= \inf_{x \in H^R, \|x\|_{H^R}=1} \left\| \sum_{i=1}^R \sum_{j=1}^R x_j G_{ij} \right\|_H^2 \\
 &= \inf_{(q_1, \dots, q_R) \in \text{St}(H; R)} \inf_{\substack{b_1, \dots, b_R \in \mathbb{R}^R \\ \sum_{i=1}^R \|b_i\|_2^2 = 1}} \left\| \sum_{i=1}^R \sum_{j=1}^R \sum_{k=1}^R b_j^{(k)} q_k G_{ij} \right\|_H^2 \\
 &= \inf_{(q_1, \dots, q_R) \in \text{St}(H; R)} \inf_{\substack{b_1, \dots, b_R \in \mathbb{R}^R \\ \sum_{i=1}^R \|b_i\|_2^2 = 1}} \sum_{i=1}^R \sum_{j=1}^R \sum_{l=1}^R \sum_{k=1}^R b_l^{(k)} G_{il} G_{ij} b_j^{(k)} \\
 &= \inf_{\substack{B \in \mathbb{R}^{R \times R} \\ \|B\|_F = 1}} \sum_{l=1}^R \sum_{j=1}^R \sum_{k=1}^R B_{kl} (G^\top G)_{lj} B_{kj} \\
 &= \inf_{\substack{B \in \mathbb{R}^{R \times R} \\ \|B\|_F = 1}} \sum_{k=1}^R B_{k,:} G^\top G B_{k,:}^\top \\
 &\geq \inf_{\substack{B \in \mathbb{R}^{R \times R} \\ \|B\|_F = 1}} \sum_{k=1}^R \lambda_R(G^\top G) \sum_{j=1}^R B_{k,j}^2 = \lambda_R(G^\top G),
 \end{aligned} \tag{A.9}$$

where in the fourth line we defined  $B = (B_{ij})_{ij} = (b_j^{(i)})_{ij} \in \mathbb{R}^{R \times R}$ , in the second-to-last line  $B_{k,:} \in \mathbb{R}^R$  indicates the  $k$ -th row of  $B$ , and in the last line  $\lambda_R(G^\top G)$  indicates the  $R$ -th eigenvalue of  $G^\top G$ . The infimum above can be attained by taking  $B \in \mathbb{R}^{R \times R}$  in (A.9) such that  $B_{1,:}^\top$  is the unit-norm eigenvector of  $G^\top G$  associated with  $\lambda_R(G^\top G)$ , and  $B_{j,:}^\top = 0 \in \mathbb{R}^R$  for  $j > 1$ . Then,  $\inf_{x \in H^R, \|x\|_{H^R}=1} \|Gx\|_{H^R}^2 = \lambda_R(G^\top G)$  independently of  $(q_1, \dots, q_R) \in \text{St}(H; R)$ . Coming back to (A.8), we get

$$\begin{aligned}
 \|P_{\hat{V}^\perp} P_V\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)}^2 &= 1 - \inf_{x \in H^R, \|x\|_{H^R}=1} \|Gx\|_{H^R}^2 = 1 - \lambda_R(G^\top G) \\
 &= 1 - \lambda_R(GG^\top) = 1 - \inf_{b \in \mathbb{R}^R, \|b\|_2=1} \|G^\top b\|_2^2 \\
 &= \sup_{\|f\|_{L^2(\Omega; H)}=1} \|P_{V^\perp} P_{\hat{V}} f\|_{L^2(\Omega; H)}^2 = \|P_{V^\perp} P_{\hat{V}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)}^2,
 \end{aligned}$$

where in the last line we just follow the same discussion of (A.8) and (A.9) in a reverse order. Since  $P_{V^\perp}$  and  $P_{\hat{V}}$  are orthogonal projections and thus self-adjoint, we get  $\|P_{\hat{V}^\perp} P_V\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} = \|P_{\hat{V}} P_{V^\perp}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)}$ . The first equality in (A.2) follows similarly.  $\square$

We now come back to the following Lipschitz continuity of projector-valued mappings of finite rank functions.

**Proposition A.3.** *Suppose that  $X, \hat{X} \in L^2(\Omega; H)$  have the representation (A.2) with  $\sigma_j, \hat{\sigma}_j > 0$ , and  $U_j, \hat{U}_j \in H, V_j, \hat{V}_j \in L^2(\Omega)$  all orthonormal in their respective Hilbert spaces, for  $j = 1, \dots, R$ . Here,  $\sigma_j$  and  $\hat{\sigma}_j, j = 1, \dots, R$ , are ordered in*

the descending order. Then, the projections  $P_U = \sum_{j=1}^R \langle U_j, \cdot \rangle_H U_j: H \rightarrow H$  and  $P_V = \sum_{j=1}^R \mathbb{E}[V_j \cdot] V_j: L^2(\Omega; H) \rightarrow L^2(\Omega; H)$  satisfy

$$\begin{aligned} \|P_U - P_{\hat{U}}\|_{H \rightarrow H} &= \|P_U - P_{\hat{U}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} \leq \min\left\{\frac{1}{\sigma_R}, \frac{1}{\hat{\sigma}_R}\right\} \|X - \hat{X}\|_{L^2(\Omega; H)}; \\ \|P_V - P_{\hat{V}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} &\leq \min\left\{\frac{1}{\sigma_R}, \frac{1}{\hat{\sigma}_R}\right\} \|X - \hat{X}\|_{L^2(\Omega; H)}. \end{aligned} \quad (\text{A.10})$$

Moreover, we have

$$\|P_U + P_V - P_U P_V - (P_{\hat{U}} + P_{\hat{V}} - P_{\hat{U}} P_{\hat{V}})\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} \leq 2 \min\left\{\frac{1}{\sigma_R}, \frac{1}{\hat{\sigma}_R}\right\} \|X - \hat{X}\|_{L^2(\Omega; H)}. \quad (\text{A.11})$$

*Proof.* First, let us prove that  $\|P_U - P_{\hat{U}}\|_{H \rightarrow H} = \|P_U - P_{\hat{U}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)}$ . On the one hand, it holds for all  $v \in L^2(\Omega; H)$  that

$$\|(P_U - P_{\hat{U}})v\|_{L^2(\Omega; H)}^2 = \mathbb{E}[\|(P_U - P_{\hat{U}})v(\omega)\|_H^2] \leq \|P_U - P_{\hat{U}}\|_{H \rightarrow H}^2 \|v\|_{L^2(\Omega; H)}^2.$$

On the other hand, since  $v \in H$  with  $\|v\|_H = 1$  is trivially in  $L^2(\Omega; H)$  with  $\|v\|_{L^2(\Omega; H)} = 1$  we have

$$\begin{aligned} \|P_U - P_{\hat{U}}\|_{H \rightarrow H} &= \sup_{v \in H, \|v\|_H=1} \|(P_U - P_{\hat{U}})v\|_{L^2(\Omega; H)} \leq \sup_{f \in L^2(\Omega; H), \|f\|_{L^2(\Omega; H)}=1} \|(P_U - P_{\hat{U}})f\|_{L^2(\Omega; H)} \\ &= \|P_U - P_{\hat{U}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)}. \end{aligned}$$

Fix  $f \in L^2(\Omega; H)$ . Using  $\langle U_j, X \rangle_H = \sigma_j V_j$  in  $\|P_{\hat{V}^\perp} P_V f\|_{L^2(\Omega; H)} = \|(\text{id} - P_{\hat{V}}) \sum_{j=1}^R \mathbb{E}[V_j f] V_j\|_{L^2(\Omega; H)}$  yields

$$\|P_{\hat{V}^\perp} P_V f\|_{L^2(\Omega; H)} = \left\| (\text{id} - P_{\hat{V}}) \sum_{j=1}^R \frac{1}{\sigma_j} \mathbb{E}[V_j f] \langle U_j, X \rangle_H \right\|_{L^2(\Omega; H)}.$$

We have  $\sigma_j^{-1} \mathbb{E}[V_j f] (\text{id} - P_{\hat{V}}) \langle U_j, \hat{X} \rangle_H = 0$  for  $j \in 1, \dots, R$ ; indeed,

$$0 = \langle U_j, (\text{id} - P_{\hat{V}}) \hat{X} \rangle_H = \langle U_j, \hat{X} \rangle_H - \langle U_j, P_{\hat{V}} \hat{X} \rangle_H$$

holds because  $(\text{id} - P_{\hat{V}}) \hat{X} = 0$ , but [23, Theorem 8.13] implies

$$\begin{aligned} \langle U_j, P_{\hat{V}} \hat{X} \rangle_H &= \sum_{j=1}^R \langle U_j, \mathbb{E}[\hat{V}_j \hat{X}] \rangle_H \hat{V}_j \\ &= \sum_{j=1}^R \mathbb{E}[\langle U_j, \hat{V}_j \hat{X} \rangle_H] \hat{V}_j = \sum_{j=1}^R \mathbb{E}[\hat{V}_j \langle U_j, \hat{X} \rangle_H] \hat{V}_j = P_{\hat{V}}(\langle U_j, \hat{X} \rangle_H), \end{aligned}$$

and thus  $(\text{id} - P_{\hat{V}}) \langle U_j, \hat{X} \rangle_H = 0$  for  $j \in 1, \dots, R$ . Hence, the Cauchy–Schwarz inequality

implies

$$\begin{aligned}
 \|P_{\hat{V}^\perp} P_V f\|_{L^2(\Omega; H)} &= \left\| (\text{id} - P_{\hat{V}}) \sum_{j=1}^R \frac{1}{\sigma_j} \mathbb{E}[V_j f] \langle U_j, (X - \hat{X}) \rangle_H \right\|_{L^2(\Omega; H)} \\
 &\leq \| \text{id} - P_{\hat{V}} \|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} \cdot \left( \sum_{j=1}^R \frac{1}{\sigma_j} \sqrt{\| \mathbb{E}[V_j f] \|_H^2 \mathbb{E}[|\langle U_j, (X - \hat{X}) \rangle_H|^2]} \right) \\
 &\leq \frac{1}{\sigma_R} \left( \sum_{j=1}^R \| \mathbb{E}[V_j f] \|_H^2 \right)^{1/2} \left( \sum_{j=1}^R \mathbb{E}[|\langle U_j, (X - \hat{X}) \rangle_H|^2] \right)^{1/2}.
 \end{aligned}$$

Recalling the orthonormality assumptions on  $U_j$  and  $V_j$ , we use Bessel's inequality to obtain  $\sum_{j=1}^R \mathbb{E}[|\langle U_j, (X - \hat{X}) \rangle_H|^2] \leq \|X - \hat{X}\|_{L^2(\Omega; H)}^2$ . Moreover, one can also prove that  $\sum_{j=1}^R \| \mathbb{E}[f V_j] \|_H^2 \leq \|f\|_{L^2(\Omega; H)}^2$  following a similar argument to show Bessel's inequality.

It follows  $\|P_{\hat{V}^\perp} P_V f\|_{L^2(\Omega; H)} \leq \frac{\|f\|_{L^2(\Omega; H)}}{\sigma_R} \|X - \hat{X}\|_{L^2(\Omega; H)}$ . Therefore, via using Lemma A.2 we conclude that

$$\|P_V - P_{\hat{V}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} \leq \frac{1}{\sigma_R} \|X - \hat{X}\|_{L^2(\Omega; H)}.$$

Likewise,  $\|P_U - P_{\hat{U}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} \leq \frac{1}{\hat{\sigma}_R} \|X - \hat{X}\|_{L^2(\Omega; H)}$  holds. By exchanging the roles of  $U, V$  with  $\hat{U}, \hat{V}$ , respectively, using the homogeneity property of the norm and (A.3), then the inequalities (A.10) follow. Finally, the last claim (A.11) comes from

$$\begin{aligned}
 &\|P_U + P_V - P_U P_V - (P_{\hat{U}} + P_{\hat{V}} - P_{\hat{U}} P_{\hat{V}})\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} \\
 &= \|(P_U - P_{\hat{U}}) + (\text{id} - P_U)P_V - (\text{id} - P_{\hat{U}})P_{\hat{V}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} \\
 &= \|(P_U - P_{\hat{U}}) + (\text{id} - P_U)P_V - (\text{id} - P_U)P_{\hat{V}} \\
 &\quad + (\text{id} - P_U)P_{\hat{V}} - (\text{id} - P_{\hat{U}})P_{\hat{V}}\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} \\
 &\leq \|(P_U - P_{\hat{U}})(\text{id} - P_{\hat{V}})\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} + \|(\text{id} - P_U)(P_V - P_{\hat{V}})\|_{L^2(\Omega; H) \rightarrow L^2(\Omega; H)} \\
 &\leq 2 \min\left\{ \frac{1}{\sigma_R}, \frac{1}{\hat{\sigma}_R} \right\} \|X - \hat{X}\|_{L^2(\Omega; H)}.
 \end{aligned}$$

□

We conclude this section giving another useful result inherent to orthogonal projections.

**Lemma A.4.** *Given  $\gamma, N > 0$  two positive constants, let*

$$\mathbb{B}_{\gamma, N} := \{t \in [0, T] \rightarrow \mathbf{U}_t \in \mathbb{R}^{R \times d} \mid \sup_{s \in [0, T]} \|(\mathbf{U}_s \mathbf{U}_s^\top)^{-1}\|_F \leq \gamma \text{ and } \sup_{s \in [0, T]} \|\mathbf{U}_s\|_F < \sqrt{N}\},$$

where the invertibility of  $\mathbf{U}_t \mathbf{U}_t^\top$  is implicitly assumed. Given  $\mathbf{U}, \tilde{\mathbf{U}} \in \mathbb{B}_{\gamma, N}$ , define the orthogonal projectors  $P_{\mathbf{U}_t}^{\text{row}} := \mathbf{U}_t^\top (\mathbf{U}_t \mathbf{U}_t^\top)^{-1} \mathbf{U}_t$ ,  $P_{\tilde{\mathbf{U}}_t}^{\text{row}} := \tilde{\mathbf{U}}_t^\top (\tilde{\mathbf{U}}_t \tilde{\mathbf{U}}_t^\top)^{-1} \tilde{\mathbf{U}}_t$  onto the rows of  $\mathbf{U}_t$  and  $\tilde{\mathbf{U}}_t$ , respectively. Then there exists a constant  $C_1 > 0$  such that the following holds

$$\sup_{t \in [0, T]} \|P_{\mathbf{U}_t}^{\text{row}} - P_{\tilde{\mathbf{U}}_t}^{\text{row}}\|_F \leq C_1 \sup_{t \in [0, T]} \|\mathbf{U}_t - \tilde{\mathbf{U}}_t\|_F.$$

Further, given  $\tilde{\gamma}, M > 0$  two positive constants, let

$$\mathbb{B}_{\tilde{\gamma}, M} := \{t \in [0, T] \rightarrow \mathbf{Y}_t \in [L^2(\Omega)]^R \mid \sup_{s \in [0, T]} \|\mathbb{E}[\mathbf{Y}_s \mathbf{Y}_s^\top]^{-1}\|_{\text{F}} \leq \tilde{\gamma} \text{ and } \mathbb{E}[\sup_{s \in [0, T]} |\mathbf{Y}_s|^2] < M\}.$$

Given  $\mathbf{Y}, \tilde{\mathbf{Y}} \in \mathbb{B}_{\tilde{\gamma}, M}$ , let  $P_{\mathbf{Y}_t}[\cdot] := \mathbf{Y}_t^\top \mathbb{E}[\mathbf{Y}_t \mathbf{Y}_t^\top]^{-1} \mathbb{E}[\mathbf{Y}_t \cdot]$ ,  $P_{\tilde{\mathbf{Y}}_t}[\cdot] := \tilde{\mathbf{Y}}_t^\top \mathbb{E}[\tilde{\mathbf{Y}}_t \tilde{\mathbf{Y}}_t^\top]^{-1} \mathbb{E}[\tilde{\mathbf{Y}}_t \cdot]$  be the orthogonal projectors onto the subspace spanned by the components of  $\mathbf{Y}_t$  and  $\tilde{\mathbf{Y}}_t$ , respectively. Then there exists a constant  $C_2 > 0$  such that the following holds

$$\mathbb{E}[\sup_{t \in [0, T]} \|P_{\mathbf{Y}_t} - P_{\tilde{\mathbf{Y}}_t}\|_{L^2(\Omega) \rightarrow L^2(\Omega)}^2]^{\frac{1}{2}} \leq C_2 \mathbb{E}[\sup_{t \in [0, T]} |\mathbf{Y}_t - \tilde{\mathbf{Y}}_t|^2]^{\frac{1}{2}}$$

*Proof.* For  $\mathbf{Y}_t, \tilde{\mathbf{Y}}_t \in [L^2(\Omega)]^R$  we have

$$\begin{aligned} & \mathbf{Y}_t^\top \mathbb{E}[\mathbf{Y}_t \mathbf{Y}_t^\top]^{-1} \mathbb{E}[\mathbf{Y}_t g] - \tilde{\mathbf{Y}}_t^\top \mathbb{E}[\tilde{\mathbf{Y}}_t \tilde{\mathbf{Y}}_t^\top]^{-1} \mathbb{E}[\tilde{\mathbf{Y}}_t g] \\ &= (\mathbf{Y}_t^\top - \tilde{\mathbf{Y}}_t^\top) \mathbb{E}[\mathbf{Y}_t \mathbf{Y}_t^\top]^{-1} \mathbb{E}[\mathbf{Y}_t g] \\ & \quad + \tilde{\mathbf{Y}}_t^\top (\mathbb{E}[\mathbf{Y}_t \mathbf{Y}_t^\top]^{-1} - \mathbb{E}[\tilde{\mathbf{Y}}_t \tilde{\mathbf{Y}}_t^\top]^{-1}) \mathbb{E}[\mathbf{Y}_t g] \\ & \quad + \tilde{\mathbf{Y}}_t^\top \mathbb{E}[\tilde{\mathbf{Y}}_t \tilde{\mathbf{Y}}_t^\top]^{-1} (\mathbb{E}[\mathbf{Y}_t g] - \mathbb{E}[\tilde{\mathbf{Y}}_t g]). \end{aligned}$$

Let  $\mathbf{Y}, \tilde{\mathbf{Y}} \in \mathbb{B}_{\tilde{\gamma}, M}$ . From [16, Lemma 3.5], there exists a constant  $C(R, M) > 0$  such that

$$\begin{aligned} \|\mathbb{E}[\mathbf{Y}_t \mathbf{Y}_t^\top]^{-1} - \mathbb{E}[\tilde{\mathbf{Y}}_t \tilde{\mathbf{Y}}_t^\top]^{-1}\|_{\text{F}} &\leq \tilde{\gamma} C(R, M) \mathbb{E}[|\mathbf{Y}_t - \tilde{\mathbf{Y}}_t|^2]^{\frac{1}{2}} \\ &\leq \tilde{\gamma} C(R, M) \mathbb{E}[\sup_{t \in [0, T]} |\mathbf{Y}_t - \tilde{\mathbf{Y}}_t|^2]^{\frac{1}{2}}, \end{aligned}$$

and thus

$$\begin{aligned} \sup_{t \in [0, T]} |(P_{\mathbf{Y}_t} - P_{\tilde{\mathbf{Y}}_t})g| &\leq \tilde{\gamma} \sup_{t \in [0, T]} |\mathbf{Y}_t - \tilde{\mathbf{Y}}_t| \mathbb{E}[\sup_{t \in [0, T]} |\mathbf{Y}_t| |g|] \\ & \quad + \tilde{\gamma} C(R, M) \left( \sup_{t \in [0, T]} |\tilde{\mathbf{Y}}_t| \right) \left( \mathbb{E}[\sup_{t \in [0, T]} |\mathbf{Y}_t - \tilde{\mathbf{Y}}_t|^2]^{\frac{1}{2}} \right) \mathbb{E}[\sup_{t \in [0, T]} |\mathbf{Y}_t| |g|] \\ & \quad + \left( \sup_{t \in [0, T]} |\tilde{\mathbf{Y}}_t| \right) \tilde{\gamma} (\mathbb{E}[\sup_{t \in [0, T]} |\mathbf{Y}_t - \tilde{\mathbf{Y}}_t| |g|]). \end{aligned}$$

Then, using the Cauchy–Schwarz inequality to  $\mathbb{E}[\sup_{t \in [0, T]} |\mathbf{Y}_t| |g|]$  and  $\mathbb{E}[\sup_{t \in [0, T]} |\mathbf{Y}_t - \tilde{\mathbf{Y}}_t| |g|]$ , taking the square of both sides and taking the expectation yields the result. The result for  $P_{\mathbf{U}_t}^{\text{row}}$  can be analogously shown.  $\square$

## References

- [1] Michael F Adamer, Heather A Harrington, Eamonn A Gaffney, and Thomas E Woolley. “Coloured noise from stochastic inflows in reaction–diffusion systems”. In: *Bulletin of Mathematical Biology* 82 (2020), pp. 1–28.
- [2] Markus Bachmayr. “Low-Rank Tensor Methods for Partial Differential Equations”. In: *Acta Numerica* 32 (May 2023), pp. 1–121. ISSN: 0962-4929, 1474-0508. DOI: 10.1017/S0962492922000125.

- [3] Markus Bachmayr, Henrik Eisenmann, Emil Kieri, and André Uschmajew. “Existence of dynamical low-rank approximations to parabolic problems”. In: *Mathematics of Computation* 90.330 (2021), pp. 1799–1830.
- [4] Yu Cao and Jianfeng Lu. “Stochastic dynamical low-rank approximation method”. In: *Journal of Computational Physics* 372 (2018), pp. 564–586.
- [5] Gianluca Ceruti and Christian Lubich. “An unconventional robust integrator for dynamical low-rank approximation”. In: *BIT Numerical Mathematics* 62.1 (2022), pp. 23–44.
- [6] Mulin Cheng, Thomas Y Hou, and Zhiwen Zhang. “A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations I: Derivation and algorithms”. In: *Journal of Computational Physics* 242 (2013), pp. 843–868.
- [7] Mulin Cheng, Thomas Y Hou, and Zhiwen Zhang. “A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations II: Adaptivity and generalizations”. In: *Journal of Computational Physics* 242 (2013), pp. 753–776.
- [8] Minseok Choi, Themistoklis P Sapsis, and George Em Karniadakis. “On the equivalence of dynamically orthogonal and bi-orthogonal methods: Theory and numerical simulations”. In: *Journal of Computational Physics* 270 (2014), pp. 1–20.
- [9] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge university press, 2014.
- [10] Florian Feppon and Pierre FJ Lermusiaux. “Dynamically orthogonal numerical schemes for efficient stochastic advection and Lagrangian transport”. In: *Siam Review* 60.3 (2018), pp. 595–625.
- [11] Yuri E Gliklikh. *Global and stochastic analysis with applications to mathematical physics*. Springer Science & Business Media, 2010.
- [12] Elton P Hsu. *Stochastic analysis on manifolds*. 38. American Mathematical Soc., 2002.
- [13] Kiyosi Itô. “On a formula concerning stochastic differentials”. In: *Nagoya Mathematical Journal* 3 (1951), pp. 55–65.
- [14] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*. Vol. 113. Springer Science & Business Media, 2012.
- [15] Tosio Kato. *Perturbation theory for linear operators*. Classics in mathematics. Springer-Verlag, Berlin, 1995. xxii+619. ISBN: 3-540-58661-X.
- [16] Yoshihito Kazashi and Fabio Nobile. “Existence of dynamical low rank approximations for random semi-linear evolutionary equations on the maximal interval”. In: *Stochastics and Partial Differential Equations: Analysis and Computations* 9.3 (2021), pp. 603–629.
- [17] Yoshihito Kazashi, Fabio Nobile, and Eva Vidličková. “Stability properties of a projector-splitting scheme for dynamical low rank approximation of random parabolic equations”. In: *Numerische Mathematik* 149 (2021), pp. 973–1024.
- [18] Emil Kieri, Christian Lubich, and Hanna Walach. “Discretized dynamical low-rank approximation in the presence of small singular values”. In: *SIAM Journal on Numerical Analysis* 54.2 (2016), pp. 1020–1038.

- [19] Peter E Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*. Heidelberg: Springer, 1992.
- [20] Othmar Koch and Christian Lubich. “Dynamical low-rank approximation”. In: *SIAM Journal on Matrix Analysis and Applications* 29.2 (2007), pp. 434–454.
- [21] Othmar Koch and Christian Lubich. “Regularity of the multi-configuration time-dependent Hartree approximation in quantum molecular dynamics”. In: *ESAIM Math. Model. Numer. Anal.* 41.2 (Mar. 2007), pp. 315–331. ISSN: 0764-583X. DOI: 10.1051/m2an:2007020. URL: <http://www.esaim-m2an.org/10.1051/m2an:2007020>.
- [22] Hiroshi Kunita and Shinzo Watanabe. “On square integrable martingales”. In: *Nagoya Mathematical Journal* 30 (1967), pp. 209–245.
- [23] Giovanni Leoni. *A first course in Sobolev spaces*. 2nd ed. Vol. 181. Graduate studies in mathematics. American Mathematical Society, Providence, RI, 2017. ISBN: 978-1-4704-2921-8.
- [24] Christian Lubich. *From quantum to classical molecular dynamics: reduced models and numerical analysis*. Vol. 12. European Mathematical Society, 2008.
- [25] Christian Lubich and Ivan V Oseledets. “A projector-splitting integrator for dynamical low-rank approximation”. In: *BIT Numerical Mathematics* 54.1 (2014), pp. 171–188.
- [26] Eleonora Musharbash and Fabio Nobile. “Dual dynamically orthogonal approximation of incompressible Navier Stokes equations with random boundary conditions”. In: *Journal of Computational Physics* 354 (2018), pp. 135–162.
- [27] Eleonora Musharbash, Fabio Nobile, and Eva Vidličková. “Symplectic dynamical low rank approximation of wave equations with random parameters”. In: *BIT Numerical Mathematics* 60 (2020), pp. 1153–1201.
- [28] Eleonora Musharbash, Fabio Nobile, and Tao Zhou. “Error analysis of the dynamically orthogonal approximation of time dependent random PDEs”. In: *SIAM Journal on Scientific Computing* 37.2 (2015), A776–A810.
- [29] Boris Muzellec, Kanji Sato, Mathurin Massias, and Taiji Suzuki. “Dimension-free convergence rates for gradient Langevin dynamics in RKHS”. In: *Conference on Learning Theory*. PMLR. 2022, pp. 1356–1420.
- [30] Achim Nonnenmacher and Christian Lubich. “Dynamical low-rank approximation: applications and numerical experiments”. In: *Mathematics and Computers in Simulation* 79.4 (2008), pp. 1346–1357.
- [31] Philip E. Protter. *Stochastic Integration and Differential Equations*. Vol. 21. Stochastic Modelling and Applied Probability. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005. ISBN: 978-3-642-05560-7 978-3-662-10061-5. DOI: 10.1007/978-3-662-10061-5. URL: <http://link.springer.com/10.1007/978-3-662-10061-5> (visited on 06/02/2023).
- [32] Alfio Quarteroni, Gianluigi Rozza, et al. *Reduced order methods for modeling and computational reduction*. Vol. 9. New York: Springer, 2014.



- [33] Martin Redmann. “Solving high-dimensional optimal stopping problems using optimization based model order reduction”. In: *Applied Mathematical Finance* (2022), pp. 1–31.
- [34] Martin Redmann, Christian Bayer, and Pawan Goyal. “Low-dimensional approximations of high-dimensional asset price models”. In: *SIAM Journal on Financial Mathematics* 12.1 (2021), pp. 1–28.
- [35] Themistoklis P Sapsis and Pierre FJ Lermusiaux. “Dynamically orthogonal field equations for continuous stochastic dynamical systems”. In: *Physica D: Nonlinear Phenomena* 238.23-24 (2009), pp. 2347–2360.
- [36] Themistoklis P Sapsis and Andrew J Majda. “Blended reduced subspace algorithms for uncertainty quantification of quadratic systems with a stable mean state”. In: *Physica D: Nonlinear Phenomena* 258 (2013), pp. 61–76.
- [37] Themistoklis P Sapsis and Andrew J Majda. “Blending modified Gaussian closure and non-Gaussian reduced subspace methods for turbulent dynamical systems”. In: *Journal of Nonlinear Science* 23 (2013), pp. 1039–1071.
- [38] Wilhelmus HA Schilders, Henk A Van der Vorst, and Joost Rommes. *Model order reduction: theory, research aspects and applications*. Vol. 13. Springer, 2008.
- [39] René L Schilling. *Brownian Motion: A Guide to Random Processes and Stochastic Calculus*. 3rd ed. De Gruyter, 2021.
- [40] Mattheus P Ueckermann, Pierre FJ Lermusiaux, and Themistoklis P Sapsis. “Numerical schemes for dynamically orthogonal equations of stochastic fluid and ocean flows”. In: *Journal of Computational Physics* 233 (2013), pp. 272–294.
- [41] Bart Vandereycken. “Low-rank matrix completion by Riemannian optimization”. In: *SIAM Journal on Optimization* 23.2 (2013), pp. 1214–1236.
- [42] Eva Vidlicková. “Dynamical low rank approximation for uncertainty quantification of time-dependent problems”. N. 9087. EPFL, 2022.
- [43] Ke Wei, Jian-Feng Cai, Tony F Chan, and Shingyu Leung. “Guarantees of Riemannian optimization for low rank matrix recovery”. In: *SIAM Journal on Matrix Analysis and Applications* 37.3 (2016), pp. 1198–1222.