

Sampling-based model order reduction for stochastic differential equations driven by fractional Brownian motion

Nahid Jamshidi^{1,*} and Martin Redmann^{1,**}

¹ Martin Luther University Halle-Wittenberg, Institute of Mathematics, Theodor-Lieser-Str. 5, 06120 Halle (Saale),

In this paper, we study large-scale linear fractional stochastic systems representing, e.g., spatially discretized stochastic partial differential equations (SPDEs) driven by fractional Brownian motion (fBm) with Hurst parameter $H > 1/2$. Such equations are more realistic in modeling real-world phenomena in comparison to frameworks not capturing memory effects. To the best of our knowledge, dimension reduction schemes for such settings have not been studied so far.

In this work, we investigate empirical reduced order methods that are either based on snapshots (e.g. proper orthogonal decomposition) or on approximated Gramians. In each case, dominant subspaces are learned from data. Such model reduction techniques are introduced and analyzed for stochastic systems with fractional noise and later applied to spatially discretized SPDEs driven by fBm in order to reduce the computational cost arising from both the high dimension of the considered stochastic system and a large number of required Monte Carlo runs.

We validate our proposed techniques with numerical experiments for some large-scale stochastic differential equations driven by fBm.

Copyright line will be provided by the publisher

1 Introduction

The proper orthogonal decomposition (POD) is a data-driven strategy for the reduction of large-scale models that is based on the SVD of snapshot matrices and utilizes an offline-online splitting. During the offline stage, the available empirical data regarding the solution of the system is used to determine a dominant subspace of the full configuration space. After that, the high-dimensional system is projected onto that subspace with the aim of capturing the original dynamics, and in the online stage, such a reduced model is solved numerically at a lower cost of computation than the original model. The POD method has been successfully utilized in a wide variety of scientific and engineering problems, such as fluid dynamics [7, 12, 19], electric circuit analysis [16], or structural dynamics [2]; see also [4, 5], for additional references.

There have been other applications in which POD has been used to investigate stochastic systems; for example, see [11], and [20]. However, POD techniques for stochastic differential equations driven by fractional Brownian motion (fBm) have not been studied yet.

The fBm is an excellent candidate to simulate numerous phenomena thanks to its self-similar and long-range dependency qualities. Some examples of these phenomena include financial markets (see, for example, [8, 13]), and traffic networks (see, e.g., [10, 21]). It is vital to point out that if $H \neq \frac{1}{2}$ the process W^H is neither a semi-martingale nor a Markov process, and standard stochastic calculus approaches are no longer applicable. Therefore, in order to construct a stochastic calculus framework for fBm, it will be necessary to use methods that are distinct from the standard Ito calculus. In recent years, numerous developments of stochastic calculus and stochastic differential equations with respect to the fBm have taken place, particularly for $H \in (\frac{1}{2}, 1)$ (see, for example, [1, 9, 14, 15]).

Throughout this entire work, we focus on the Hurst parameter $H \in (\frac{1}{2}, 1)$, and we also employ Young's integral [22] to define an integral with respect to fBm. This concept is discussed in detail in Section 2.

In this research, we investigate large-scale linear fractional stochastic systems, resulting from the spatial discretization of SPDEs driven by fBm. We employ three approaches for this goal. The first two approaches are based on the POD, and the third method is based on an empirical Gramian. The difference between the first and second POD techniques is that the second method divides the system into two subsystems. Subsequently, POD applies to reduce each subsystem individually. The sum of these two reduced systems is then the reduced order candidate for the original system. To employ the third technique, a Gramian is introduced which characterizes dominant subspaces in terms of its eigenspaces. Such approaches have the advantage of finding a good quality reduced system uniformly in the control. We approximate the Gramian using samples of the solution since it cannot be determined via a Lyapunov equation as it can be done having Wiener or Lévy noise [3]. This is one of the main obstacles to studying model reduction for fractional stochastic systems.

The paper is structured as follows. In Section 2, we provide a quick introduction to fBm and describe some of its key characteristics before defining Young's integral. In Section 3, we briefly discuss the setting and the general reduced system structure. In Section 4, we present the POD and empirical Gramian approaches. We end the paper in Section 5 with numerical experiments illustrating our results.

* Corresponding author: nahid.jamshidi@mathematik.uni-halle.de

** martin.redmann@mathematik.uni-halle.de

2 Fractional Brownian motion and Young integration

A mean zero Gaussian process $W^H = \{W^H(t), t \geq 0\}$ is called fractional Brownian motion (fBm) with Hurst parameter $H \in (0, 1)$ if it has the covariance function

$$\mathbb{E}[W^H(t)W^H(s)] = \frac{1}{2} (s^{2H} + t^{2H} - |t - s|^{2H}). \quad (1)$$

Mandelbrot and Van Ness investigated this process after it was proposed by Kolmogorov, and they constructed a stochastic integral representation of it in terms of a standard Brownian motion. The climatologist Hurst created a statistical analysis of annual Nile river runoffs, and the resulting parameter is known as the Hurst index. The fBm has the following properties:

- Self-similarity: The probability distributions of the processes $\{a^{-H}W^H(at), t \geq 0\}$ and $\{W^H(t), t \geq 0\}$ are the same for any constant $a > 0$. This is a direct consequence of (1).
- Stationary increments: By definition, the increments of a fBm are normally distributed with zero mean. From (1), it further follows that $\mathbb{E}[(W^H(t) - W^H(s))^2] = |t - s|^{2H}$, so that the distribution only depends on $t - s$.
- Hölder continuity: [6] allows us to calculate the following modulus of continuity for fBm trajectories: There exists a non negative random variable $G_{\epsilon, T}$ for all $\epsilon > 0$ and $T > 0$ such that, almost surely,

$$|W^H(t) - W^H(s)| \leq G_{\epsilon, T} |t - s|^{H - \epsilon}$$

for all $s, t \in [0, T]$. The Hurst parameter H accounts not only for the sign of the correlation of the increments, but also for the regularity of the sample paths, i.e., for any $\epsilon > 0$, the trajectories are Hölder continuous of order $H - \epsilon$.

We will now consider the integral defined by Young [22] in 1936 in order to define a stochastic integral w.r.t. W^H , $H > 1/2$. This scenario covers integrands and integrators with a certain Hölder regularity.

Definition 2.1 Let C^α be the set of Hölder continuous functions defined on $[0, T]$, with index $0 < \alpha \leq 1$.

- Suppose that $f \in C^\alpha$ and $g \in C^\beta$, where $\alpha + \beta > 1$. Given a sequence $(t_i^n)_{i=1}^{k_n}$ of partitions of $[0, T]$ with $\lim_{n \rightarrow \infty} \max_{i=0}^{k_n-1} \{t_{n,i+1} - t_{n,i}\} = 0$. Then, the Young integral $\int_0^T f(s)dg(s)$ is then defined as

$$\int_0^T f(s)dg(s) := \lim_{n \rightarrow \infty} \sum_{i=1}^{k_n-1} f(t_i^n)[g(t_{i+1}^n) - g(t_i^n)].$$

- As pointed out above, the paths of W^H are a.s. Hölder continuous for $\alpha = H - \epsilon$. We can define $\int_0^T Y(s)dW^H(s)$ for a process Y with $(H - \epsilon)$ -Hölder continuous trajectories path-wise in the sense of Young if $H > 1/2$.

3 Setting and projection based reduced system

We consider the following stochastic system with control u satisfying $\|u\|_{L^2_T}^2 := \mathbb{E} \int_0^T \|u(t)\|_2^2 dt < \infty$:

$$dx(t) = [Ax(t) + Bu(t)]dt + \sum_{i=1}^q N_i x(t) dW_i^H(t), \quad x(0) = x_0 = X_0 z, \quad y(t) = Cx(t), \quad t \in [0, T], \quad (2)$$

where $A, N_i \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$, $X_0 \in \mathbb{R}^{n \times v}$ and $z \in \mathbb{R}^v$. W_1^H, \dots, W_q^H are independent fBm with Hurst index $H \in (1/2, 1)$. (2) is defined as an integral equation using Definition 2.1 to make sense of $\int_0^t N_i x(s) dW_i^H(s)$. We aim to find a space that is spanned by the columns of $V \in \mathbb{R}^{n \times r}$ such that $x(t) \approx V\bar{x}(t)$. Inserting this into (2) yields

$$Vd\bar{x}(t) = [AV\bar{x}(t) + Bu(t)]dt + \sum_{i=1}^q N_i V\bar{x}(t) dW_i^H(t) + e(t), \quad \bar{y}(t) = CV\bar{x}(t). \quad (3)$$

We enforce the error $e(t)$ to be orthogonal to some space spanned by columns of $W \in \mathbb{R}^{n \times r}$. Multiplying (3) with $(W^\top V)^{-1}W^\top$ from the left yields

$$d\bar{x}(t) = [A_r\bar{x}(t) + B_r u(t)]dt + \sum_{i=1}^q N_{i,r}\bar{x}(t) dW_i^H(t), \quad \bar{x}(0) = x_{0,r} = \bar{X}_0 z \quad \bar{y}(t) = C_r\bar{x}(t), \quad t \in [0, T], \quad (4)$$

where $\bar{X}_0 = (W^\top V)^{-1}W^\top X_0$ and

$$A_r = (W^\top V)^{-1}W^\top AV, \quad B_r = (W^\top V)^{-1}W^\top B, \quad N_{i,r} = (W^\top V)^{-1}W^\top N_i V, \quad C_r = CV.$$

Often $W^\top V = I$ is chosen. If $W = V$, we have a Galerkin approximation. In this paper, we focus on empirical Galerkin based methods, where $W = V$ has orthonormal columns. We start with the POD that is based on snapshots and subsequently discuss techniques based on empirical Gramians representing a uniform reduced model approach in u .

4 Empirical reduced order methods

4.1 Proper orthogonal decomposition

The first strategy As a first strategy, we apply the POD technique to problem (2), meaning that we create a snapshots matrix and pick V from the SVD of this matrix in order to find (4) with $W = V$. To be more precise, we sample the solution for a fixed u and x_0 , so we get

$$Z_j = [x(t_1, w_j), x(t_2, w_j), \dots, x(t_N, w_j)], \quad \text{for } \omega_j \in \Omega, \quad j = 1, \dots, N_s$$

where $N, N_s > 0$ are the number of time points and samples. We want this method to be computationally inexpensive. Therefore, N_s should be small. We consider $Z := [X_0, Z_1, Z_2, \dots, Z_{N_s}]$ and calculate the SVD of Z :

$$Z = \begin{pmatrix} V & * \end{pmatrix} \begin{pmatrix} \Sigma \\ * \end{pmatrix} \begin{pmatrix} U^\top \\ * \end{pmatrix}$$

Therefore, $V_Z = \begin{pmatrix} V & * \end{pmatrix}$ is a weighted basis for the snapshots. If Σ contains all large singular values, then V contains the dominant basis vectors and we obtain a good reduced system (ROM) from that V if the data is representative.

The second strategy We split the main system (2) into subsystems as follows:

$$dx_1(t) = [Ax_1(t) + Bu(t)]dt + \sum_{i=1}^q N_i x_1(t) dW_i^H(t), \quad x_1(0) = 0, \quad y_1(t) = Cx_1(t), \quad (5)$$

$$dx_2(t) = Ax_2(t)dt + \sum_{i=1}^q N_i x_2(t) dW_i^H(t), \quad x_2(0) = x_0 = X_0 z, \quad y_2(t) = Cx_2(t) \quad (6)$$

and observe that $y = y_1 + y_2$. In contrast to the first strategy, the POD method is applied to each subsystem individually to receive a ROM. The sum of the subsystem reduced order outputs then is an approximation for y .

4.2 Empirical Gramian method

In this subsection, we introduce a Gramian P that allows us to identify the dominant subspaces of (2). In particular, the projection matrix V shall represent the eigenspaces of P corresponding to the large eigenvalues of the Gramian. This approach has the advantage of providing good quality ROMs uniformly in the control u . However, since fBm with $H > 1/2$ does not have independent increments, no Lyapunov equations associated with P can be derived. Therefore, we consider an empirical version of this Gramian meaning that we discretize the integral representation of P by a Monte-Carlo method.

The fundamental solution to (6) is defined as the matrix valued stochastic process Φ solving

$$\Phi(t, s) = I + \int_s^t A\Phi(\tau, s)d\tau + \sum_{i=1}^q \int_s^t N_i \Phi(\tau, s) dW_i^H(\tau) \quad (7)$$

for $t \geq s \geq 0$. We further set $\Phi(t) := \Phi(t, 0)$. Then, we can show that the state of (2) is given by

$$x(t) = \Phi(t)x_0 + \int_0^t \Phi(t, s)Bu(s)ds = \Phi(t)X_0 z + \int_0^t \Phi(t, s)Bu(s)ds. \quad (8)$$

We define a (time-limited) Gramian $P := P_1 + P_2$ for (2), where

$$P_1 := \mathbb{E} \left[\int_0^T \Phi(s)X_0 X_0^\top \Phi^\top(s)ds \right], \quad P_2 := \mathbb{E} \left[\int_0^T \Phi(T, s)BB^\top \Phi^\top(T, s)ds \right], \quad (9)$$

where P_1 and P_2 can be interpreted as Gramians for (5) and (6). Since P can not be computed directly, we approximate it by

$$\begin{aligned} P &= \int_0^T \mathbb{E} [\Phi(s)X_0 X_0^\top \Phi^\top(s)] ds + \int_0^T \mathbb{E} [\Phi(T, s)BB^\top \Phi^\top(T, s)] ds \\ &\approx \frac{T}{N} \sum_{i=1}^N (\mathbb{E} [\Phi(s_i)X_0 X_0^\top \Phi^\top(s_i)] + \mathbb{E} [\Phi(T, s_i)BB^\top \Phi^\top(T, s_i)]) \\ &\approx \frac{T}{N \cdot N_s} \sum_{i=1}^N \sum_{j=1}^{N_s} (x_2(s_i, 0, X_0, w_j)x_2(s_i, 0, X_0, w_j)^\top + x_2(T, s_i, B, w_j)x_2(T, s_i, B, w_j)^\top) =: \tilde{P}, \end{aligned} \quad (10)$$

where $x_2(t, s, X_0, \cdot)$ denotes the solution to (6) with initial time s and initial state X_0 . \tilde{P} is called empirical Gramian. As mentioned above, the ROM (4) with $W = V$ based on \tilde{P} is obtained by conducting an eigenvalue decomposition of the empirical Gramian and subsequently choosing V to be the matrix of eigenvectors associated with the large eigenvalues of \tilde{P} .

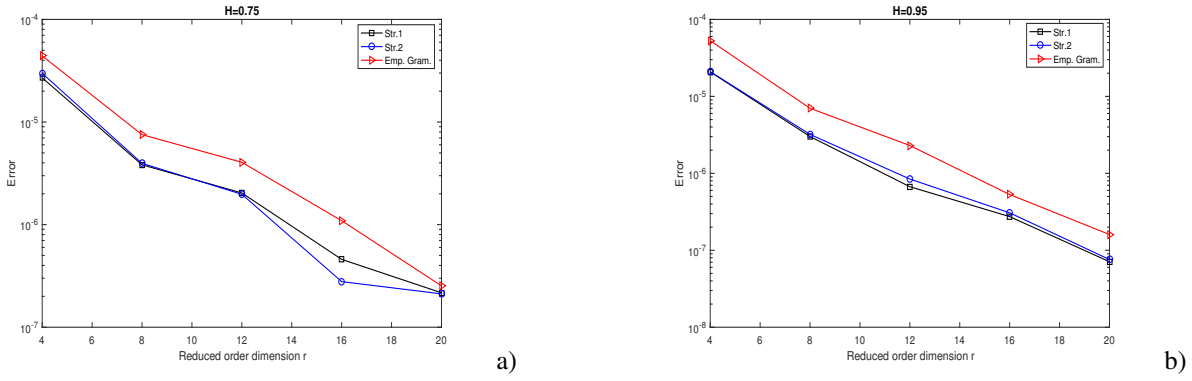


Fig. 1: The error for the three approaches with two different Hurst parameters $H = 0.75$ (a) and $H = 0.95$ (b).

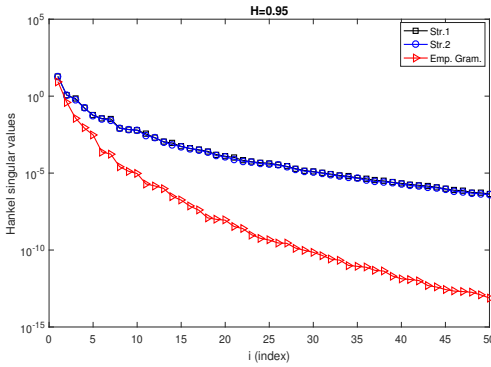


Fig. 2: Decay of first 50 HSVs of the system (2) based on POD Str. 1 and Str. 2 as well as Emp. Gram.

Table 1: $\sup_{t \in [0,1]} \mathbb{E} \|y(t) - \bar{y}(t)\|_2$ for $r \in \{4, 8, 12, 16, 20\}$ and $H = 0.95$.

red.dim. r	Str.1	Str.2	Emp. Gram.
4	$2.08e-05$	$2.09e-05$	$5.28e-05$
8	$3.01e-06$	$3.21e-06$	$7.01e-06$
12	$6.70e-07$	$8.43e-07$	$2.30e-06$
16	$2.74e-07$	$3.07e-07$	$5.30e-07$
20	$7.10e-08$	$7.58e-08$	$1.60e-07$

5 Numerical examples

Example 1 We begin with a modified version of an example studied in [3] and consider the following controlled stochastic partial differential equation:

$$\begin{aligned} \frac{\partial X(t, \zeta)}{\partial t} &= a \Delta X(t, \zeta) + 1_{[\frac{\pi}{4}, \frac{3\pi}{4}]^2}(\zeta) u(t) + \gamma e^{-|\zeta_1 - \frac{\pi}{2}| - \zeta_2} X(t, \zeta) \frac{\partial W^H(t)}{\partial t}, \quad t \in [0, 1], \quad \zeta \in [0, \pi]^2, \\ X(t, \zeta) &= 0, \quad t \in [0, 1], \quad \zeta \in \partial[0, \pi]^2, \quad \text{and} \quad X(0, \zeta) = b \sin(\zeta), \end{aligned} \quad (11)$$

where $a, b > 0$, $\gamma \in \mathbb{R}$ and W^H is an one-dimensional fBm and $m = 1$. We set $\mathcal{H} = L^2([0, \pi])$, $\mathcal{B} = 1_{[\frac{\pi}{4}, \frac{3\pi}{4}]^2}(\cdot)$ and $\mathcal{N}x = e^{-|\cdot - \frac{\pi}{2}| - \cdot} x$ for $x \in \mathcal{H}$. Due to the fact that the Dirichlet Laplacian generates a C_0 -semigroup, the \mathcal{H} -valued solution X of (11) is meant to be in the mild sense. The eigenfunctions $(h_k)_{k \in \mathbb{N}}$ of this Laplacian demonstrate an orthonormal basis of \mathcal{H} which is ordered based on the corresponding eigenvalues $(-\lambda_k)_{k \in \mathbb{N}}$. The average temperature in the uncontrolled area is the quantity of interest, i.e.,

$$Y(t) = \frac{4}{3\pi^2} \int_{[0, \pi]^2 \setminus [\frac{\pi}{4}, \frac{3\pi}{4}]^2} X(t, \zeta) d\zeta.$$

We define the associated output operator as $\mathcal{C}x = \frac{4}{3\pi^2} \int_{[0, \pi]^2 \setminus [\frac{\pi}{4}, \frac{3\pi}{4}]^2} x(\zeta) d\zeta$ for $x \in \mathcal{H}$.

A numerical solution to this SPDE can be found by first considering a spatial discretization as a possible first step. We exploit a spectral Galerkin technique in this context to build an approximation X_n to X by an orthogonal projection of (11) onto the subspace $\mathcal{H}_n = \text{span}\{h_1, \dots, h_n\}$. X_n converges to the SPDE solution with $n \rightarrow \infty$ and is determined by its Fourier coefficients $x(t) = (\langle X_n(t), h_1 \rangle_{\mathcal{H}}, \dots, \langle X_n(t), h_n \rangle_{\mathcal{H}})$, that satisfy (2) with coefficients

- $A = a \text{diag}(-\lambda_1, -\lambda_2, \dots, -\lambda_n)$, $B = (\langle \mathcal{B}, h_k \rangle_{\mathcal{H}})_{k=1 \dots n}$, $C = (\mathcal{C}h_k)_{k=1 \dots n}$,
- $N_1 = (\langle \mathcal{N}(h_i), h_k \rangle_{\mathcal{H}})_{k,i=1 \dots n}$, $x_0 = b(\langle \sin(\cdot), h_k \rangle_{\mathcal{H}})_{k=1 \dots n}$,

where we fix $a = 0.2$ and $b = 0.1$. We refer to [3] for more details on the discretization. We now fix the dimension of spatial discretization to be $n = 225$. For the discretization in time, the Runge-Kutta scheme in [18] is used, where the number

of time steps is $N = 100$. The control $u(t) = e^{-bt}$ and the initial condition $x_0 = b \langle (\sin(\cdot), h_k)_{\mathcal{H}} \rangle_{k=1 \dots n}$ were used to calculate snapshots for the POD methods. The quality of the resulting ROMs is calculated for $u(t) = \sqrt{\frac{2}{\pi}} \cos(t)$ and $x_0 = 0$, respectively. The reduction error is measured by the quantity $\sup_{t \in [0,1]} \mathbb{E} \|y(t) - \bar{y}(t)\|_2$ which is approximated based on $N_s = 10^4$ samples. We apply POD with its two different strategies, see Section 4.1, and the empirical Gramian approach, see Section 4.2, to obtain the ROM. We present the error for these three approaches for two different Hurst parameters $H = 0.75$ and $H = 0.95$ in Figure 1. According to this figure, although the performance of the first and second POD strategies differs somewhat in a few of the resulting errors, these two ways produce almost identical performance, which is superior to the result of the empirical Gramian method. In addition, the quantity of errors caused by the three aforementioned approaches with $H = 0.95$ for the system (11) can be observed in Table 1.

Figure 2 depicts the first 50 singular values of the snapshot matrices and the eigenvalues of the empirical Gramian \tilde{P} . For simplicity, these values are called Hankel singular values (HSVs) of system (2) for the three aforementioned methods. The second strategy involves plotting $\sigma_{i,1} + \sigma_{i,2}$, where $\sigma_{i,1}$ and $\sigma_{i,2}$ represent the POD singular values of subsystems (5) and (6), respectively. It is understood that small truncated HSV correspond to a small reduction error in the MOR scheme. Due to the rapid decay of these values, we can conclude that even for small reduced dimensions r , small errors can be achieved.

Example 2 We consider the following controlled stochastic partial differential equation which is a modification of the example studied in [17]. For $t \in [0, 1]$ and $\zeta \in [0, \pi]$

$$\frac{\partial^2 X(t, \zeta)}{\partial t^2} + a \frac{\partial X(t, \zeta)}{\partial t} = \frac{\partial^2}{\partial \zeta^2} X(t, \zeta) + e^{-|\zeta - \frac{\pi}{2}|} u(t) + 2e^{-|\zeta - \frac{\pi}{2}|} X(t, \zeta) \frac{\partial W^H(t)}{\partial t},$$

$$X(t, 0) = 0 = X(t, \pi), \quad t \in [0, 1], \quad X(0, \zeta) \equiv 0, \quad \dot{X}(t, \zeta)|_{t=0} = b \sin(\zeta)$$

is investigated and the output is an approximation of the string's midpoint, which is

$$Y(t) = \frac{1}{2\epsilon} \int_{\frac{\pi}{2} - \epsilon}^{\frac{\pi}{2} + \epsilon} X(t, \zeta) d\zeta,$$

where $a, b > 0$ and $\epsilon > 0$. We know from [17] that the Galerkin system looks as follows:

$$x(t) = x_0 + \int_0^t Ax(s) + Bu(s) ds + \int_0^t N_1 x(s) dW^H(s) \quad y(t) = Cx(t), \quad t \in [0, T], \quad (12)$$

where the coefficients are represented by

- $A = \text{diag}(E_1, \dots, E_{\frac{n}{2}})$ with $E_l = \begin{pmatrix} 0 & l \\ -l & -a \end{pmatrix}$ ($l = 1, \dots, \frac{n}{2}$),
- $B = (b_k)_{k=1, \dots, n}$ with $b_{2l-1} = 0$, $b_{2l} = \sqrt{\frac{2}{\pi}} \langle e^{-|\cdot - \frac{\pi}{2}|}, \sin(l \cdot) \rangle_{\mathcal{H}}$,
- $x_0 = (x_{0,k})_{k=1, \dots, n}$ with $x_{0,2l-1} = 0$, $x_{0,2l} = b \sqrt{\frac{2}{\pi}} \langle \sin(\cdot), \sin(l \cdot) \rangle_{\mathcal{H}}$,
- $N_1 = (n_{kj})_{k,j=1, \dots, n}$ with

$$n_{(2l-1),j} = 0, \quad n_{(2l),j} = \begin{cases} 0, & \text{if } j = 2\nu \\ \frac{4}{\pi\nu} \langle \sin(l \cdot), e^{-|\cdot - \frac{\pi}{2}|} \sin(\nu \cdot) \rangle_{\mathcal{H}}, & \text{if } j = 2\nu - 1, \end{cases}$$

for $j = 1 \dots, n$ and $\nu = 1, \dots, \frac{n}{2}$,

- the output matrix is given by $C^T = (c_k)_{k=1, \dots, n}$ with $c_{2l} = 0$ and $c_{2l-1} = \frac{1}{\sqrt{2\pi}l^2} [\cos(l(\frac{\pi}{2})) - \cos(l(\frac{\pi}{2} + \epsilon))]$.

In this case, we assume $b = 0.1$, $a = 2$ and $H = 0.75$. Further, the sizes of spatial and time discretization are $n = 100$ and $N = 100$, respectively. In this example, we consider the same scenario as we did in the first example; which means we calculate snapshots and errors using different control functions and different initial values. We use $N_s = 100$, $u(t) = \sqrt{\frac{2}{1-e^{-2\pi}}} e^{-t}$ and the initial state defined above in the computation of the snapshots. We calculated the problem's error using the following values: $N_s = 10^4$, $u(t) = \sqrt{\frac{2}{\pi}} \cos(t)$ and $x_0 = 0$. As can be seen in Figure 3 and Table 2, the performance of the three techniques is nearly identical. With a more thorough perspective, however, the results of the first and second POD strategies are totally consistent, and these two ways are superior to the empirical Gramian method by a minor difference.

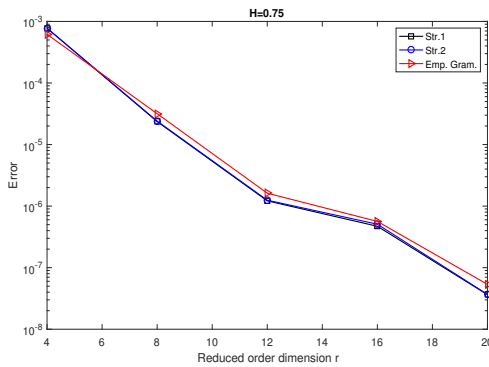


Fig. 3: The error for the three approaches with Hurst parameters $H = 0.75$.

Table 2: $\sup_{t \in [0,1]} \mathbb{E} \|y(t) - \bar{y}(t)\|_2$ for $r \in \{4, 8, 12, 16, 20\}$ and $H = 0.75$.

red.dim. r	Str.1	Str.2	Emp. Gram.
4	$7.77e-04$	$7.83e-04$	$6.18e-04$
8	$2.36e-05$	$2.40e-05$	$3.17e-05$
12	$1.22e-06$	$1.25e-06$	$1.62e-06$
16	$4.70e-07$	$5.08e-07$	$5.65e-07$
20	$3.64e-08$	$3.68e-08$	$5.35e-08$

References

- [1] E. Alos and D. Nualart, Stochastic integration with respect to the fractional Brownian motion. *Stoch. Stoch. Rep.* **75**, 24 (2003).
- [2] M. Amabili, A. Sarkar, and M. Paidoussis, Reduced-order models for nonlinear vibrations of cylindrical shells via the proper orthogonal decomposition method, *J. Fluids. Struct.* (2)**18**, 24 (2003).
- [3] P. Benner and M. Redmann, Model Reduction for Stochastic Systems, *Stoch PDE: Anal Comp*, 3(3):291–338, 2015.
- [4] P. Bhatt, and B. Haasdonk, Symplectic model order reduction with nonorthonormal bases, *Math. Comput. Appl.* 19 (2019).
- [5] S. Chaturantabut and D. C. Sorensen, Nonlinear model reduction via discrete empirical interpolation, *SIAM. J. Sci. Comput.* (5)**32**, 28 (2010).
- [6] A. M. Garsia, E. Rodemich, and H. Rumsey, A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Univ. Math. J.* **20**, 14 (1971).
- [7] P. Holmes, J. Lumley, G. Berkooz, and C. Rowley. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry* (Cambridge university press, Cambridge,2012), p. 386.
- [8] Y. Hu and B. Oksendal, Fractional white noise calculus and applications to finance, *Infin. Dimens. Quantum Probab. Relat.* **6**, 32 (2003).
- [9] M. Kamrani and N. Jamshidi, Implicit Euler approximation of stochastic evolution equations with fractional Brownian motion, *Commun. Nonlinear. Sci.* **44**, 10 (2017).
- [10] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, On the self-similar nature of ethernet traffic (Extended version). *IEEE/ACM Trans.* **2**, 15 (1994).
- [11] T. Iliescu, H. Liu, and X. Xie. Regularized reduced order models for a stochastic Burgers equation, *Int. J. Numer. Anal. Model.* (4-5)**15**, 14 (2018).
- [12] K. Ito and S. Ravindran, A reduced-order method for simulation and control of fluid flows. *J. Comput. Phys.* (2)**143**, 23 (1998).
- [13] B. B. Mandelbrot, The variation of certain speculative prices. *J. Bus.* **36**, 26 (1963).
- [14] Y. Mishura, *Stochastic Calculus for Fractional Brownian Motion and Related Processes* (Springer, Berlin, 2008), p. 411.
- [15] A. Neuenkirch, Optimal pointwise approximation of stochastic differential equations driven by fractional Brownian motion, *Stoch. Process. Their. Appl.* (12)**118**, 40 (2008).
- [16] P. Parrilo, S. Lall, F. Paganini, G. Verghese, B. Lesieutre, and J. Marsden, Model reduction for analysis of cascading failures in power systems, *Proc. Am. Control Conf.* **6**, 5 (1999).
- [17] M. Redmann, P. Benner, Approximation and Model Order Reduction for Second Order Systems with Lévy-Noise, *AIMS Proceedings*, (2015).
- [18] M. Redmann, S. Riedel, Runge-Kutta methods for rough differential equations, *arXiv:2003.12626* (2020).
- [19] C. W. Rowley, T. Colonius, and R. M. Murray, Model reduction for compressible flows using POD and Galerkin projection, *Phys. D: Nonlinear Phenom.* (1)**189**, 15 (2004).
- [20] T. M. Tyrnowski. Data-driven structure-preserving model reduction for stochastic Hamiltonian systems. *arXiv:2201.13391* (2022).
- [21] W. Willinger, M. S. Taqqu, W. E. Leland, and D. V. Wilson, Self-similarity in high-speed packet traffic: analysis and modelling of Ethernet traffic measurements. *Stat. Sci.* **10**, 19 (1995).
- [22] L. C. Young, An inequality of Hölder type, connected with Stieltjes integration, *Acta. Math.* **67**, 32 (1936).