Learning non-intrusive ROMs from linear SDEs with additive noise

M. A. Freitag * **J. M. Nicolaus** ** **M. Redmann** ***

* *University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam*
** *University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam(e-mail: jan.martin.nicolaus@uni-potsdam.de)*
*** *Martin-Luther-University Halle-Wittenberg, Theodor-Lieser-Str. 5, 06120 Halle (Saale)*

Keywords: black box modelling, data-driven modelling, non-intrusive model order reduction, stochastic linear systems

1. INTRODUCTION

Model Order Reduction (MOR) for stochastic linear systems is concerned with approximating the high-dimensional Full Order Model (FOM)

$$\mathrm{d}x(t) = [\mathrm{A}x(t) + \mathrm{B}u(t)]\,\mathrm{d}t + \mathrm{M}\mathrm{d}W(t), \quad x(0) = x_0, \quad (1)$$

by a surrogate model. Such arise, for example, from spatial discretisations of PDEs, with noisy boundary conditions. Here, the so-called drift of (1)is a function of the state $x(t) \in \mathbb{R}^n$ and the *control* $u(t) \in \mathbb{R}^m$ given by the respective multiplication with $\mathrm{A} \in \mathbb{R}^{n\times n}$ and $\mathrm{B} \in \mathbb{R}^{n\times m}$. The term $\mathrm{M}\mathrm{d}W(t)$ is called the diffusion term and describes the influence of the noise-generating process $W(t) \in \mathbb{R}^d$ on the state variable. In this case the process is a d-dimensional standard *Brownian motion.* The matrix $\mathrm{M} \in \mathbb{R}^{n\times d}$ is called the diffusion coefficient. Due to the structure of the SDE (1) the FOM state variable $x(t)$ is Gaussian for each fixed time t if x_0 is Gaussian or constant. Hence, the distribution of x at the time $t \in [0, T]$ is completely determined by the expectation $\mathrm{E}(t) \in \mathbb{R}^n$ and covariance $\mathrm{C}(t) \in \mathbb{R}^{n\times n}$.

2. PROJECTION-BASED MOR

Projection-based MOR constructs such Reduced Order Models (ROMs) by approximating the FOM state variable in an r-dimensional subspace $\mathcal{V}_r$, that is, one assumes $x(t) \approx \mathrm{V}_r\mathrm{V}_r^T x(t) = \mathrm{V}_r x_r(t)$, where orthogonal columns of $\mathrm{V}_r = [v_1, \ldots, v_r] \in \mathbb{R}^{n\times r}$ span the subspace $\mathcal{V}_r$. By the requiring a Galerkin condition on the residual of the FOM dynamics, one obtains

$$\mathrm{d}x_r(t) = [\mathrm{A}_r x_r(t) + \mathrm{B}_r u(t)]\,\mathrm{d}t + \mathrm{M}_r\mathrm{d}W(t), \quad (2)$$

with the *reduced coefficients*

$$\mathrm{A}_r := \mathrm{V}_r^T \mathrm{A}\mathrm{V}_r \in \mathbb{R}^{r\times r},\ \mathrm{B}_r := \mathrm{V}_r^T\mathrm{B} \in \mathbb{R}^{r\times m},$$
$$\mathrm{M}_r := \mathrm{V}_r^T\mathrm{M} \in \mathbb{R}^{r\times d},\ x_r(0) := \mathrm{V}_r^T x_0 \in \mathbb{R}^r.$$

If $r \ll n$, then (2) is much cheaper to compute than the FOM (1). Since projections retain the linear structure of the original FOM equations, the ROM state $x_r(t)$ is Gaussian as well for each fixed $t \geq 0$. One can show that the expectation of the projected state $E_r(t) := \mathrm{E}[x_r(t)]$ satisfies the ODE

$$\dot{E}_r(t) = \mathrm{A}_r E_r(t) + \mathrm{B}_r u(t),\ E_r(0) = \mathrm{E}[x_{r,0}]$$

and approximates the FOM expectation after lifting with V_r, since $\mathrm{V}_r E_r(t) = \mathrm{V}_r\mathrm{E}[x_r(t)] = \mathrm{E}[\mathrm{V}_r\mathrm{V}_r^T x(t)] \approx \mathrm{E}[x(t)]$. Analogous results for the covariance matrix $C_r(t) = \mathrm{cov}[x_r(t)]$ hold, see Freitag et al. (2024).

A popular data-driven method to construct $\mathcal{V}_r$ is *Proper Orthogonal Decomposition* (POD) method. In this method, the dominant subspace of observed *snapshots* $\mathrm{X}^s = [x(t_1), \ldots, x(t_s)]$ is chosen as $\mathcal{V}_r$. This is achieved by taking the r leading left-singular vectors of X^s as the columns of $\mathrm{V}_r = [v_1, \ldots, v_r]$. To construct a ROM in such a way, it is necessary to have access to the FOM matrices A, B, and M. Such methods are called *intrusive* and can be infeasible in the case of, for instance, black-box or legacy code.

3. NON-INTRUSIVE MOR

To address this issue, so-called *non-intrusive* methods have been developed. These methods do not require the availability of the FOM system coefficients, but instead rely on the availability of large amounts of data or the ability to query the FOM. One well-known method is the Operator Inference (OpInf) approach by Peherstorfer and Willcox (2016), which recently has been extended to the SDE setting by Freitag et al. (2024) We briefly illustrate this extension. In the standard OpInf approach for SDEs, one first collects L samples of s trajectory observations $x(t_1), \ldots, x(t_s)$ of the FOM state, which are then used to compute approximations of the reduced expectation

$$\mathrm{E}_{r,i}^L = \mathrm{V}_r^T\mathrm{E}_i^L,\ \mathrm{E}_i^L \approx E(t_i) := \mathrm{E}[x_{t_i}],\ i \in \{1, \ldots, s\}.$$

of the ROM state variable x_r at the observation times. An approximation of the time derivative $\dot{E}_r(t)$ of the reduced expectation can be obtained by a finite difference approximation $\mathrm{E}_{r,i}^{L,h}$ using $\mathrm{E}_{r,i}^L := \mathrm{V}_r^T\mathrm{E}_i^L$, where h is given by the difference between the (equidistant) observation times t_i. Thus, to obtain approximations to A_r and B_r, one can solve the least-squares problem

$$[\mathrm{A}_r^*\ \mathrm{B}_r^*] = \underset{\tilde{\mathrm{O}}\in\mathbb{R}^{r\times(r+m)}}{\mathrm{argmin}} \|\tilde{\mathrm{O}}\mathrm{D}^L - \mathrm{R}^{L,h}\|_F, \quad (3)$$

* The research has been partially funded by the Deutsche Forschungsgemeinschaft (DFG) - Project-ID 318763901 - SFB1294 as well as by the DFG individual grant “Low-order approximations for large-scale problems arising in the context of high-dimensional PDEs and spatially discretized SPDEs”– project number 499366908.

DOI: 10.34726/9010

where

$$\mathrm{D}^L = \begin{bmatrix} \mathrm{E}^L_{r,1} & \dots & \mathrm{E}^L_{r,s} \\ u(t_1) & \dots & u(t_s) \end{bmatrix} \text{ and } \mathrm{R}^{L,h} = \left[\mathrm{E}^{L,h}_{r,1}, \dots, \mathrm{E}^{L,h}_{r,s}\right].$$

Since D^L and $\mathrm{R}^{L,h}$ are constructed from approximations, one can view (3) as a perturbed version of an unperturbed least-squares problem, where direct observations would be available. Freitag et al. (2024) show that if the data matrix $\mathrm{D}^L \in \mathbb{R}^{r+m\times s}$ is of full rank, the unique solution is an almost-surely convergent estimator of the unperturbed least squares solution in the limit of $L \to \infty$ and $h \to 0$. Using an analogous approach, Freitag et al. (2024) obtain an estimation of the product $\mathrm{M}_r\mathrm{M}_r^T \in \mathbb{R}^{d\times d}$ for the diffusion coefficient by utilising the inferred A_r^* instead of A_r.

One drawback of the standard OpInf formulation, as Peherstorfer (2020) points out, is that the projected FOM trajectories can differ from the trajectories of the intrusive ROM (2). This so-called *closure error* arises due to the inability of ROMs of the form (2) to model the non-Markovian dynamics of the projected FOM state variable with respect to the subspace $\mathcal{V}_r$. Thus, the ROM obtained from standard OpInf can fail to approximate the reduced dynamics in $\mathcal{V}_r$. To address this issue, Peherstorfer (2020) proposes a modified sampling scheme called *re-projection*. We illustrate this method in the SDE setting of this paper. The core idea is to estimate the re-projection sampling, performed directly on the expectation, by computing the empirical mean of re-projected samples. To perform the re-projection scheme, access to the stepping function $f(x, u)$

$$\tilde{x}_{i+1} = f(\tilde{x}_i, u_i) = \tilde{\mathrm{A}}\tilde{x}_i + \tilde{\mathrm{B}}u_i + \tilde{\mathrm{M}}z_i, z_i \sim \mathcal{N}(0_d, \mathrm{I}_d),$$

of the time-discretised FOM is required. Here, 0_d is the d-dimensional zero vector and I_d the identity matrix of size $d \times d$. The sampling algorithm then computes trajectories $\{\hat{x}_i,\ i = 1, \dots, s\} \subset \mathbb{R}^r$ by projecting each query result onto $\mathcal{V}_r$, that is, one computes $\hat{x}_{i+1} = \mathrm{V}_r^T f(\mathrm{V}_r\hat{x}_i, u_i)$ and constructs the matrices

$$\hat{\mathrm{X}}^s = \begin{bmatrix} \hat{x}^E_1 & \dots & \hat{x}^E_{s-1} \\ u(t_1) & \dots & u(t_{s-1}) \end{bmatrix} \in \mathbb{R}^{(r+m)\times(s-1)} \text{ and}$$

$$\hat{\mathrm{Y}}^s = \left[\hat{x}^E_2 \ \dots \ \hat{x}^E_s\right] \in \mathbb{R}^{r\times(s-1)}$$

from the empirical estimation $\hat{x}^E_i$ of the expectation of $\hat{x}_i$. An approximation to the time-discrete reduced operators is obtained by solving the least-squares problem

$$\left[\hat{\mathrm{A}}_r^* \ \hat{\mathrm{B}}_r^*\right] = \underset{\hat{\mathrm{O}}\in\mathbb{R}^{r\times(r+m)}}{\operatorname{argmin}} \|\hat{\mathrm{O}}\hat{\mathrm{X}}^s - \hat{\mathrm{Y}}^s\|_F. \tag{4}$$

As in the standard OpInf method, the condition number of the data-matrix X^s can be improved by sampling linearly independent pairs of initial conditions and control. Lastly, the availability of f enables us to easily obtain an estimation of the projected time-discrete diffusion operator $\tilde{\mathrm{M}}_r$. While one could proceed as in Freitag et al. (2024) by using the covariance matrices of the re-projected time-steps, it is much simpler to sample the projected time-stepping function f with a zero initial condition and control, since

$$\mathrm{V}_r^T f(0_r, 0) = \mathrm{V}_r^T\tilde{\mathrm{M}}z,\ z \sim \mathcal{N}(0_d, \mathrm{I}_d).$$

The covariance matrix $\tilde{\mathrm{C}}_f$ of such samples is an estimation of $\mathrm{V}_r^T\tilde{\mathrm{M}}\tilde{\mathrm{M}}^T\mathrm{V}_r$. An approximation of $\tilde{\mathrm{M}}_r$ is then obtained by, e.g., an eigenvalue decomposition of $\tilde{\mathrm{C}}_f$. Note, that this approach approximates the reduced system coefficients of the time-discretised FOM, instead of the reduced system coefficients of the time-continuous FOM.

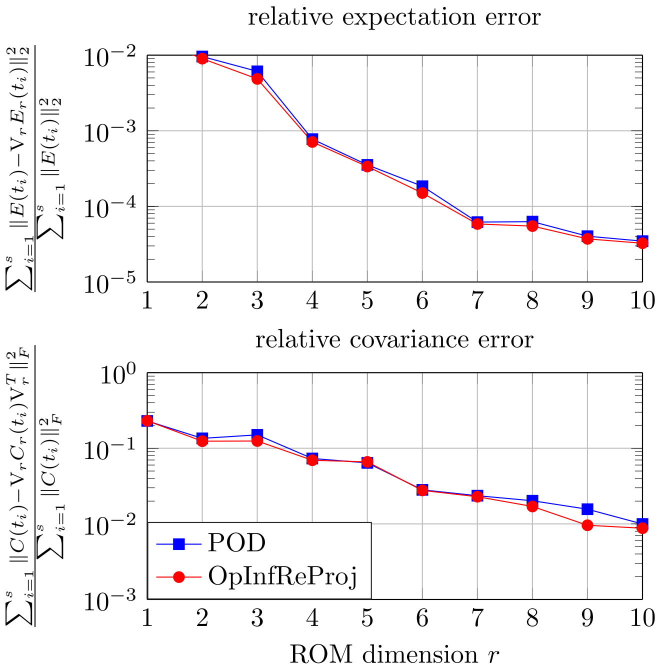

Fig. 1. Relative errors of expectation and covariance. FOM dimension $n = 1357$. Number of samples and time steps for inference and testing $L = 10^4$ and $s = 10^3$. Time-step size $h = 10^{-3}$.

4. NUMERICAL EXAMPLE

The FOM is obtained from the *Steel Profile* Benchmark from the Oberwolfach Benchmark Collection (2005). The control function models the temperature controls, which can be applied on $m = 7$ sections of the boundary of the profile. One can model a noisy control u, which is perturbed by white noise, by choosing the diffusion coefficient to be $\mathrm{M} = \frac{\mathrm{B}}{\|\mathrm{B}\|}$. The step function f is given by a semi-implicit Euler-Maruyama time-discretisation of the corresponding SDE. Figure 1 reports the relative summed errors in the expectation and covariance between the FOM and the POD and the FOM and Operator Inference with re-projection ROMs. The code used to perform the experiment displayed in Figure 1 is available at `https://github.com/JMNicolaus/SDE_OpInfRP`

REFERENCES

Freitag, M.A., Nicolaus, J.M., and Redmann, M. (2024). Learning stochastic reduced models from data: A non-intrusive approach. URL `https://arxiv.org/abs/2407.05724`.

Oberwolfach Benchmark Collection (2005). Steel profile. hosted at MORwiki – Model Order Reduction Wiki. URL `http://modelreduction.org/index.php/Steel_Profile`.

Peherstorfer, B. (2020). Sampling low-dimensional markovian dynamics for preasymptotically recovering reduced models from data with operator inference. *SIAM Journal on Scientific Computing*, 42(5), A3489–A3515.

Peherstorfer, B. and Willcox, K. (2016). Data-driven operator inference for nonintrusive projection-based model reduction. *Computer Methods in Applied Mechanics and Engineering*, 306, 196–215.

DOI: 10.34726/9010