Max Planck Institute Magdeburg
Preprints

Martin Redmann Peter Benner

Singular Perturbation Approximation for Linear Systems with Lévy Noise

MPIMD/15-22 December 17, 2015

Impressum:

Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg

Publisher:
Max Planck Institute for
Dynamics of Complex Technical Systems

Address:
Max Planck Institute for
Dynamics of Complex Technical Systems
Sandtorstr. 1
39106 Magdeburg

www.mpi-magdeburg.mpg.de/preprints

Abstract

To solve a stochastic linear evolution equation numerically, finite dimensional approximations are commonly used. For a good approximation, one might end up with a sequence of ordinary stochastic linear equations of high order. To reduce the high dimension for practical computations, we consider the singular perturbation approximation as a model order reduction technique in this paper. This approach is well-known from deterministic control theory and here we generalize it for controlled linear systems with Lévy noise. Additionally, we discuss properties of the reduced order model, provide an error bound, and give some examples to demonstrate the quality of this model order reduction technique.

1 Introduction

Model order reduction (MOR) is of major importance in the field of deterministic control theory. It is used to save computational time by replacing large scale systems by systems of low order in which the main information of the original system should be captured. Such kind of high dimensional problems occur for example after the special discretization of a PDE which can be used to model chemical, physical or biological phenomena. A particular MOR scheme is balanced truncation (BT) assuming asymptotic stability of the original system. The idea is to balance the system such that one creates a system where the dominant reachable and observable states are the same. Afterwards, the difficult to observe and difficult to reach states are truncated. This was considered first in Moore [13]; Antoulas [1] or Obinata, Anderson [14] for a thorough treatment of the topic.
Since many phenomena in natural sciences contain uncertainties, it is natural to extend PDE models by adding a noise term. This leads to stochastic PDEs (SPDEs) which are studied, e.g., in Da Prato, Zabczyk [5] and in Prévôt and Röckner [16] for the Wiener case. Peszat, Zabczyk consider more general equations with Lévy noise in [15], where the solutions may have jumps. Again, SPDEs can be reduced to large scale ordinary SDEs by a semi-discretization. A possibility to do that is the Galerkin method which is for example investigated in Grecksch, Kloeden [8], Hausenblas [10], Jentzen, Kloeden [11] and Redmann, Benner [17]. For that reason, generalizing model order reduction techniques to stochastic systems can easily be motivated. Inspired by this application, balanced truncation is considered for SDEs with Wiener noise in Benner, Damm [2] and for systems with Lévy noise by Benner, Redmann in [4]. Benner and Redmann additionally pointed out the benefit of BT in the field of SPDEs in detail by applying it to particular examples, see [4] and [17].
An alternative method to obtain a reduced order model (ROM) is the singular perturbation approximation (SPA), see Liu, Anderson [12] and Fernando, Nicholson [7] for deterministic linear systems. Rather than setting all truncated states to zero as in BT, they are assumed constant which allows to solve for them and thus include this information in the differential equation for the remaining states. This has the advantage of a zero steady-state error, a property often important in applications. The SPA also exists for bilinear systems. For that framework, we refer to Hartmann et al. [9].
In this paper, we generalize the work of Liu and Anderson to linear systems with Lévy noise. In Section 2, we motivate the SPA for stochastic systems and derive the ROM which coincides with the deterministic case ROM if $N = 0$. Next, in Section 3, we analyze the properties of the ROM. First, we consider the stability of the reduced system. We show that it is mean square stable and discuss why the ideas from Benner et al. [3] cannot be adopted in order to prove the preservation of mean square asymptotic stability. Additionally, we state the remaining part to complete the proof of mean square asymptotic stability for the ROM. Besides the stability analysis of the ROM, we investigate the reachability and observability in the reduced model by the SPA. Therefore, we repeat the concepts used in Benner, Damm [2] and Benner, Redmann [4] and show, by an example, that one can lose the complete reachability and observability in the ROM even if one starts with a completely reachable and observable original model which is in contrast to the deterministic case. In Section 4, we assume to have a ROM that preserves the mean square asymptotic stability

which is vital for the existence of the error bound we provide in that section. This error bound we obtain by modifying the coefficients of the ROM in order to have the same structure as in the original system. The modified matrices coincide with the ones that are used in the bilinear case by Hartmann et al. [9]. Furthermore, from that error bound, we can point out the cases in which we have a good approximation by the SPA. Finally, in Section 5, we compare BT and the SPA by reducing a large scale system we get from a special discretization of a second order SPDE with Poisson noise. There, we see that SPA can be better if one considers the underlying equations on a larger time interval. We present a second example, which we generate randomly, to illustrate further advantages of the SPA.

2 SPA

Let M be a scalar and square integrable Lévy process with mean zero defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.[1] In addition we assume M to be $(\mathcal{F}_t)_{t\geq 0}$-adapted and the increments $M(t+h) - M(t)$ to be independent of $\mathcal{F}_t$ for $t, h \geq 0$. We consider the following equations:

$$\begin{aligned} dX(t) &= [AX(t) + Bu(t)]dt + NX(t-)dM(t), \quad X(0) = x_0 \in \mathbb{R}^n, \\ Y(t) &= CX(t), \quad t \geq 0, \end{aligned} \tag{1}$$

where A, $N \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{p\times n}$, $B \in \mathbb{R}^{n\times m}$ and $X(t-) := \lim_{s\uparrow t} X(s)$. With L^2_T we denote the space of all $(\mathcal{F}_t)_{t\geq 0}$-adapted (cadlag) processes v with values in $\mathbb{R}^m$, which are square integrable with respect to $\mathbb{P}\otimes dt$. The norm in L^2_T is given by

$$\|v\|^2_{L^2_T} := \mathbb{E}\int_0^T v^T(t)v(t)dt = \mathbb{E}\int_0^T \|v(t)\|_2^2\, dt,$$

where we define the processes v_1 and v_2 to be equal in L^2_T if they coincide almost surely with respect to $\mathbb{P}\otimes dt$. Further, we assume the control $u \in L^2_T$ for every $T > 0$. Below, the solution of (1) at time $t \geq 0$ with initial condition $x_0 \in \mathbb{R}^n$ and given control u is always denoted by $X(t, x_0, u)$. We assume

$$\mathbb{E}\|X(t, x_0, 0)\|_2^2 \to 0 \tag{2}$$

for $t \to \infty$ and $x_0 \in \mathbb{R}^n$. This concept of stability is also used in [2] and is necessary to define (infinite) Gramians, which are the solutions of the generalized Lyapunov equations (3) and (4) below.

Theorem 2.1. *The following are equivalent:*

(i) The homogeneous equation ($u \equiv 0$) of (1) is asymptotically mean square stable.

(ii) There exists a matrix $P > 0$, such that

$$A^TP + PA + N^TPN\mathbb{E}\left[M^2(1)\right] < 0.$$

(iii) The eigenvalues of $\left(I_n \otimes A + A \otimes I_n + N \otimes N \cdot \mathbb{E}\left[M^2(1)\right]\right)$ have negative real parts.

Especially, condition (iii) implies the stability of A, that is $\sigma(A) \subset \mathbb{C}_-$.

Proof. With Proposition 3.2 in [4] the proof is similar to the Wiener case. Therefore, we refer to Theorem 3.6.1 in [6] where these results are proven for equations with Wiener noise. □

[1] We assume that $(\mathcal{F}_t)_{t\geq 0}$ is right continuous and that $\mathcal{F}_0$ contains all $\mathbb{P}$ null sets.

We assume that the system (1) is balanced, meaning that solutions of the following generalized Lyapunov equations are diagonal and equal

$$A^T\Sigma + \Sigma A + N^T\Sigma N\mathbb{E}\left[M^2(1)\right] = -C^TC, \tag{3}$$

$$A\Sigma + \Sigma A^T + N\Sigma N^T\mathbb{E}\left[M^2(1)\right] = -BB^T, \tag{4}$$

where $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \ldots \geq \sigma_n > 0$. We introduce the following partitions

$$\Sigma = \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix}, \; A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \; N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}, \; C = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \text{ and } B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix},$$

where Σ_1, A_{11}, $N_{11} \in \mathbb{R}^{r\times r}$, $C_1 \in \mathbb{R}^{p\times r}$ and $B_1 \in \mathbb{R}^{r\times m}$. Using the partition $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, the idea of balanced truncation is to select the first r rows in equation (1) and to neglect X_2 which means that we set $X_2 = 0$. This yields a reduced order model with coefficients $(A_{11}, N_{11}, C_1, B_1)$. A detailed motivation regarding BT in the stochastic case one can find in [2] and [4]. From [3] we know that balanced truncation preserves asymptotic stability also in the stochastic case if $\sigma_r \neq \sigma_{r+1}$:

$$\sigma\left(I_r \otimes A_{11} + A_{11} \otimes I_r + N_{11} \otimes N_{11} \cdot \mathbb{E}\left[M^2(1)\right]\right) \subset \mathbb{C}_-. \tag{5}$$

The same is true for the truncated part meaning

$$\sigma\left(I_{n-r} \otimes A_{22} + A_{22} \otimes I_{n-r} + N_{22} \otimes N_{22} \cdot \mathbb{E}\left[M^2(1)\right]\right) \subset \mathbb{C}_-. \tag{6}$$

From the properties (5) and (6) we can also conclude that A_{11} and A_{22} are invertible.

The method we introduce below is called singular perturbation approximation (SPA) with a more general idea of setting the symbolic derivative $\frac{dX_2}{dt}$ equal to zero instead. We obtain a system

$$\begin{aligned} \begin{pmatrix} dX_1(t) \\ 0 \end{pmatrix} &= \left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(t)\right) dt + \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \begin{pmatrix} X_1(t-) \\ X_2(t-) \end{pmatrix} dM(t), \\ Y(t) &= \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}, \quad t \geq 0, \end{aligned} \tag{7}$$

where we assume $x_0 = 0$ below. From the second line in (7), we obtain

$$0 = \int_0^t A_{21}X_1(s) + A_{22}X_2(s) + B_2u(s)ds + \int_0^t N_{21}X_1(s-) + N_{22}X_2(s-)dM(s), \tag{8}$$

such that an Ito integral equals an ordinary integral which is a strange situation, since the ordinary integral is differentiable and the Ito integral is not in general. We define the process $S(t) = \int_0^t a(s)ds + \int_0^t b(s)dM(s)$, where $a(s) := A_{21}X_1(s) + A_{22}X_2(s) + B_2u(s)$ and $b(s) := N_{21}X_1(s-) + N_{22}X_2(s-)$ and determine the mean of the stochastic differential of $S^T(t)S(t)$, $t \geq 0$. For that reason, we introduce an Ito product formula stated for example in Corollary 2.4 in [4]:

$$S^T(t)S(t) = \int_0^t dS^T(s)S(s) + \int_0^t S^T(s)dS(s) + \sum_{i=1}^{n-r}[S_i, S_i]_t,$$

with $[S_i, S_i]_t$ being the quadratic variation part of the i-th component of S. Inserting the differential of S and using the property

$$\mathbb{E}\left[\sum_{i=1}^{n-r}[S_i, S_i]_t\right] = \int_0^t \mathbb{E}\left[b^T(s)b(s)\right] ds\, \mathbb{E}\left[M^2(1)\right]$$

from Section 2.4 in [4] yields

$$\mathbb{E}\left[S^T(t)S(t)\right] = \mathbb{E}\left[\int_0^t a^T(s)S(s)ds\right] + \mathbb{E}\left[\int_0^t S^T(s)a(s)ds\right] + \int_0^t \mathbb{E}\left[b^T(s)b(s)\right] ds\, \mathbb{E}\left[M^2(1)\right],$$

Setting $S \equiv 0$ provides

$$0 = \int_0^t \mathbb{E}\left[b^T(s)b(s)\right] ds\, \mathbb{E}\left[M^2(1)\right] = \mathbb{E}\left\|\int_0^t b(s)dM(s)\right\|_2^2,$$

which implies $\int_0^t b(s)dM(s) = 0$ $\mathbb{P}$-a.s. If we apply this to equation (8), we get

$$X_2(t) = -(A_{22}^{-1}A_{21}X_1(t) + A_{22}^{-1}B_2u(t)). \tag{9}$$

By inserting this in the first line in equation (7), we have

$$X_1(t) = \int_0^t \bar{A}X_1(s) + \bar{B}u(s)ds + \int_0^t \bar{N}X_1(s-) + \bar{B}_0u(s-)dM(s) \tag{10}$$

and

$$\bar{Y}(t) = \bar{C}X_1(t) + \bar{D}u(t),$$

where

$\bar{A} = A_{11} - A_{12}A_{22}^{-1}A_{21},\ \bar{B} = B_1 - A_{12}A_{22}^{-1}B_2,\ \bar{N} = N_{11} - N_{12}A_{22}^{-1}A_{21},\ \bar{C} = C_1 - C_2A_{22}^{-1}A_{21},$
$\bar{B}_0 = -N_{12}A_{22}^{-1}B_2$ and $\bar{D} = -C_2A_{22}^{-1}B_2$.

Remark. *(i) The SPA yields a reduced order model (10) which has a different structure than the original model (1), meaning that we obtained a system in which the output equation is controlled and the control in the state equation is disturbed by Lévy noise. If we use this ROM, we have to restrict ourselves to controls with existing left limits $u(t-)$, $t \geq 0$, in order to have equation (10) well defined. Since we prefer a ROM having the same shape like the original model we will often emphasize the case $(\bar{B}, \bar{B}_0, \bar{D}) = (B_1, 0, 0)$ which we get by setting $B_2 = 0$ in equation (9).*

(ii) If we set $(\bar{B}, \bar{B}_0, \bar{D}) = (B_1, 0, 0)$, we precisely obtain the matrices that are recommended for the SPA in the bilinear case in [9].

3 Properties of the ROM

In this section, we discuss properties of the ROM which we obtain by the SPA. In the first subsection, we state the first steps how to prove the asymptotic mean square stability of the ROM. Unfortunately, this proof is not complete but our conjecture is that this property is preserved. In the second subsection, we point out that starting with a completely observable and reachable original system, one can lose these properties in the ROM.

3.1 Preservation of (asymptotic) mean square stability

We multiply A^{-T} from the left and A^{-1} from the right hand side in equation (3) and get

$$\Sigma\tilde{A} + \tilde{A}^T\Sigma + \tilde{N}^T\Sigma\tilde{N}\mathbb{E}\left[M^2(1)\right] = -\tilde{C}^T\tilde{C}, \tag{11}$$

where $\tilde{A} = A^{-1}$, $\tilde{N} = NA^{-1}$ and $\tilde{C} = CA^{-1}$. It can be shown easily that using these transformed coefficients $\tilde{A}$ and $\tilde{N}$ instead of A and N does not effect the asymptotic mean square stability. By equation (4), the corresponding dual equation is

$$(A\Sigma A^T)\tilde{A}^T + \tilde{A}(A\Sigma A^T) + \tilde{N}(A\Sigma A^T)\tilde{N}^T\mathbb{E}\left[M^2(1)\right] = -BB^T. \tag{12}$$

The reason to consider the matrices $\tilde{A}$ and $\tilde{N}$ is the following equivalence between its left upper blocks and the reduced order model coefficients:

$$\sigma\left(I_r \otimes \tilde{A}_{11} + \tilde{A}_{11} \otimes I_r + \tilde{N}_{11} \otimes \tilde{N}_{11} \cdot c\right) \subset \mathbb{C}_- \Leftrightarrow \sigma\left(I_r \otimes \bar{A} + \bar{A} \otimes I_r + \bar{N} \otimes \bar{N} \cdot c\right) \subset \mathbb{C}_-$$

with $c = \mathbb{E}\left[M^2(1)\right]$. Since one can show that

$$\tilde{A} = \begin{bmatrix} \bar{A}^{-1} & -A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \\ -A_{22}^{-1}A_{21}\bar{A}^{-1} & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix},$$

we have $\tilde{A}_{11} = \bar{A}^{-1}$, $\tilde{N}_{11} = \bar{N}\bar{A}^{-1}$. So, proving asymptotic mean square stability in the ROM is now transformed into the following problem:
Starting with a system with coefficients $\tilde{A}$ and $\tilde{N}$, show that this property is preserved if one truncates the system, i.e. one chooses the reduced order coefficients $\tilde{A}_{11}$ and $\tilde{N}_{11}$.
The main difficulty is the fact that this system is not balanced since the solution of equation (12) is neither diagonal nor it coincides with the one from equation (11). For that reason, the ideas that are used for the stability analysis of BT in [3] (see Sections 4.3 – 4.5) cannot be applied. In the deterministic case, where $N = 0$, the dual equation (12) is obtained by pre- and post-multiplying equation (4) with A^{-1} and A^{-T} which in that case yields a balanced system, see [12]. Unfortunately, this does not work in the more general framework $N \neq 0$ because we would get $A^{-1}N$ instead of the desired matrix $\tilde{N} = NA^{-1}$. We could state the desired result then under the assumption that A and N commute, which could at least partially prove the conjecture.

Since the solution of equation (11) is in diagonal form, we can adopt at least a few arguments from [3] which we state in the proof of the lemma below.

Lemma 3.1. *The reduced order models with the coefficients* $(\tilde{A}_{11}, \tilde{N}_{11})$ *or* $(\bar{A}, \bar{N})$ *are mean square stable, i.e.*

$$\sigma\left(I_r \otimes \tilde{A}_{11} + \tilde{A}_{11} \otimes I_r + \tilde{N}_{11} \otimes \tilde{N}_{11} \cdot c\right) \subset \overline{\mathbb{C}_-} \tag{13}$$

and

$$\sigma\left(I_r \otimes \bar{A} + \bar{A} \otimes I_r + \bar{N} \otimes \bar{N} \cdot c\right) \subset \overline{\mathbb{C}_-},$$

where $c = \mathbb{E}\left[M^2(1)\right]$.

Proof. We use a suitable partition of $\tilde{A}$, $\tilde{N}$, $\tilde{C}$, Σ and obtain the following equation for the left upper block of (11):

$$\Sigma_1\tilde{A}_{11} + \tilde{A}_{11}^T\Sigma_1 + \tilde{N}_{11}^T\Sigma_1\tilde{N}_{11} \cdot c = -\tilde{C}_1^T\tilde{C}_1 - \tilde{N}_{21}^T\Sigma_2\tilde{N}_{21} \cdot c, \tag{14}$$

with $\tilde{N}_{21} = (N_{21} - N_{22}A_{22}^{-1}A_{21})\bar{A}^{-1}$ and $\tilde{C}_1 = \bar{C}\bar{A}^{-1}$. Consequently, by Corollary 3.2 in [3], we obtain property (13). With the same argument, it also holds

$$\sigma\left(I_r \otimes \bar{A} + \bar{A} \otimes I_r + \bar{N} \otimes \bar{N} \cdot \mathbb{E}\left[M^2(1)\right]\right) \subset \overline{\mathbb{C}_-},$$

since by pre- and post-multiplying (14) with $\bar{A}^T$ and $\bar{A}$, we get

$$\bar{A}^T\Sigma_1 + \Sigma_1\bar{A} + \bar{N}^T\Sigma_1\bar{N} \cdot c = -\bar{C}^T\bar{C} - (N_{21} - N_{22}A_{22}^{-1}A_{21})^T\Sigma_2(N_{21} - N_{22}A_{22}^{-1}A_{21}) \cdot c.$$

□

Using Theorem 3.1 in [3], we obtain

$$\alpha(K) := \max\{\Re(\lambda) : \lambda \in \sigma(K)\} \in \sigma(K)$$

with $K = I_r \otimes \tilde{A}_{11} + \tilde{A}_{11} \otimes I_r + \tilde{N}_{11} \otimes \tilde{N}_{11} \cdot c$. By (13) it remains to show that $0 \notin \sigma(K)$ to get the desired asymptotic mean square stability. This we summarize as follows:

Conjecture 3.2. *The reduced order model with coefficients $(\bar{A}, \bar{N})$ is asymptotically mean square stable, i.e. $0 \notin \sigma(K)$.*

The result in Conjecture 3.2 is theoretically important for the existence of the error bound we state in Section 4. Practically, it is easy to check if zero is an eigenvalue of K or not since the reduced order dimension r is usually small.

3.2 Observability and reachability in the ROM

We introduce the fundamental solution of the state equation (1) as an $\mathbb{R}^{n\times n}$-valued process Φ satisfying

$$d\Phi(t) = A\Phi(t)dt + N\Phi(t-)dM(t), \quad \Phi(0) = I_n \quad t \geq 0.$$

Now, we can introduce the observability Gramian $Q = \int_0^\infty \mathbb{E}\left[\Phi^T(s)C^TC\Phi(s)\right] ds$ and the reachability Gramian $P = \int_0^\infty \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] ds$ which exist by assumption (2). Q solves equation (3) and P fulfills (4). That is proven in Section 3 in [4]. Here, we are in a balanced situation which means that

$$P = Q = \Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n).$$

We know that system (1) is completely observable if and only if the Gramian Q is positive definite, see Section 3.2 in [4]. Since the reachability concept for system (1), used in Section 3.1 in [4], neglects the information that is contained in N, it is not surprising that P can only provide partial information about the reachability of a state $x \in \mathbb{R}^n$. To be more precise, if x is reachable, then $x \in \operatorname{im} P$ but the other direction is not true. So, it is necessary to introduce the deterministic Gramian $P_D = \int_0^\infty \mathrm{e}^{At} BB^T \mathrm{e}^{A^Tt} dt$. Following [4] again, system (1) is completely reachable if and only if $P_D > 0$. This is analogous to the deterministic case, where the results are stated in [1].
Since the ROM (10) has a different structure than the original model one might think that the Gramian of the ROM has to be defined differently in order to characterize observability and reachability of the system. We will see soon that the additional matrices $\bar{B}_0$ and $\bar{D}$ have no impact in that context. In order to discuss this property we repeat the concepts of observability and reachability of the ROM:

$$dX_1(t) = [\bar{A}X_1(t) + \bar{B}u(t)]dt + [\bar{N}X_1(t-) + \bar{B}_0u(t-)]dM(t), \quad X_1(0) = \bar{x}_0, \tag{15}$$
$$\bar{Y}(t) = \bar{C}X_1(t) + \bar{D}u(t). \tag{16}$$

Since the observation concept is considered in the uncontrolled case ($u \equiv 0$), the matrix $\bar{D}$ does not enter in the following definition.

Definition 3.3. *An initial state $\bar{x}_0$ is called observable if the corresponding observation energy is positive:*

$$\left\|\bar{C}X_1(\cdot, \bar{x}_0, 0)\right\|_{L^2}^2 := \mathbb{E}\int_0^\infty \left\|\bar{C}X_1(t, \bar{x}_0, 0)\right\|_2^2 dt > 0.$$

Since we have $\bar{C}X_1(t, \bar{x}_0, 0) = \bar{C}\bar{\Phi}(t)\bar{x}_0$, $t \geq 0$, it follows

$$\left\|\bar{C}X_1(\cdot, \bar{x}_0, 0)\right\|_{L^2}^2 = \bar{x}_0^T Q_R x_0$$

with $Q_R := \mathbb{E}\int_0^\infty \bar{\Phi}^T(t)\bar{C}^T\bar{C}\bar{\Phi}(t)dt$. Here, $\bar{\Phi}$ denotes the fundamental solution of the ROM. Hence, the ROM is completely reachable if and only if $Q_R > 0$. Below, we distinguish between the solution of (15) for general $\bar{B}_0$ which we denote by $X_1(t, \bar{x}_0, u)$ and $X_1^0(t, \bar{x}_0, u)$, $t \geq 0$, denoting the solution of (15) in case $B_0 = 0$. Now, we define reachable average states.

Definition 3.4. *A state $\bar{x}$ is called reachable on average (from zero) if there is a time $T > 0$ and a control function $u \in L^2_T$, such that we have*

$$\mathbb{E}\left[X_1(T, 0, u)\right] = \bar{x}.$$

Applying the expectation on both sides of equation (15) and using the property that the Ito integral has mean zero yields that the functions $\mathbb{E}\left[X_1(t, \bar{x}_0, u)\right]$ and $\mathbb{E}\left[X_1^0(t, \bar{x}_0, u)\right]$, $t \geq 0$, are both solutions of the ODE

$$\dot{\mathcal{X}}_1(t) = \bar{A}\mathcal{X}_1(t) + \bar{B}\mathbb{E}[u(t)], \quad \mathcal{X}_1(0) = \bar{x}_0, \quad t \geq 0.$$

Hence, both expected values coincide, such that the matrix $\bar{B}_0$ can be completely neglected in the reachability concept. Setting $B_0 = 0$ provides a system having the same form like the original model (1). Consequently, we know that the ROM (15) is completely reachable if and only if $P_{D,R} := \int_0^\infty \mathrm{e}^{\bar{A}t}\,\bar{B}\bar{B}^T\,\mathrm{e}^{\bar{A}^T t}\,dt > 0$. The next example shows that starting with a completely observable and completely reachable system does not mean that the ROM has these properties as well.

Example 3.5. *We define a system (1) with $\mathbb{E}\left[M^2(1)\right] = 1$ and coefficients*

$$(A, B, C, N) = \left(\begin{pmatrix} -\frac{17}{2} & 8 & 8 \\ -8 & -20 & -20 \\ -8 & -20 & -\frac{41}{2} \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 4 & 4 \\ -4 & 2 & 2 \\ -4 & 2 & 2 \end{pmatrix}\right),$$

which is asymptotically mean square stable. In addition, we have a balanced system since for the solution of the equations (3) and (4) it holds $P = Q = \operatorname{diag}(2, 1, 1)$. Consequently, it is also completely observable. The complete reachability we obtain by $P_D > 0$. The corresponding one dimensional ROM has the coefficients

$$(\bar{A}, \bar{B}, \bar{B}_0, \bar{C}, \bar{D}, \bar{N}) = (-\frac{117}{10}, 0, 0, 0, 2, -\frac{3}{5}).$$

Since there is no control in the state equation of the ROM and the output of the uncontrolled ROM is identically zero, the reduced order system is neither completely reachable nor completely observable. Of course, this also holds for the modified ROM, where one sets $(\bar{B}, \bar{B}_0, \bar{D}) := (B_1, 0, 0) = (0, 0, 0)$.

The fact that reachability and observability are not necessarily preserved by the SPA is not surprising since analogous observations are made for BT in [4].

4 Error bound

In this section, we provide an error bound for the case $(\bar{B}, \bar{B}_0, \bar{D}) = (B_1, 0, 0)$ and $x_0 = 0$. In the error bound below, the matrix $P_R := \mathbb{E}\int_0^\infty \bar{\Phi}(t)B_1B_1^T\bar{\Phi}^T(t)dt$ enters. For its existence we assume that the mean square asymptotic stability is preserved in the ROM. This means that

$$0 \notin \sigma\left(I_r \otimes \bar{A} + \bar{A} \otimes I_r + \bar{N} \otimes \bar{N} \cdot \mathbb{E}\left[M^2(1)\right]\right), \tag{17}$$

which we know from Section 3.1. Condition (17) is usually easy to check since the reduced order dimension r is small.

Following the arguments in Section 4.2 in [4], the error of the SPA is bounded as follows:

$$\sup_{t\in[0,T]} \mathbb{E}\left\|Y(t)-\bar{Y}(t)\right\|_2 \leq \left(\operatorname{tr}\left(C\Sigma C^T\right)+\operatorname{tr}\left(\bar{C}P_R\bar{C}^T\right)-2\ \operatorname{tr}\left(CP_G\bar{C}^T\right)\right)^{\frac{1}{2}} \|u\|_{L_T^2}^2, \tag{18}$$

where

$$\begin{aligned} AP_G + P_G\bar{A}^T + NP_G\bar{N}^T\mathbb{E}\left[M^2(1)\right] &= -BB_1^T, \\ \bar{A}P_R + P_R\bar{A}^T + \bar{N}P_R\bar{N}^T\mathbb{E}\left[M^2(1)\right] &= -B_1B_1^T. \end{aligned}$$

Below, we specify this bound to emphasize the cases in which the SPA performs well.

Theorem 4.1. *If the ROM is asymptotically mean square stable, then*

$$\begin{aligned} &\operatorname{tr}\left(C\Sigma C^T\right)+\operatorname{tr}\left(\bar{C}P_R\bar{C}^T\right)-2\ \operatorname{tr}\left(CP_G\bar{C}^T\right) \\ &= \operatorname{tr}(2\Sigma_2((N_{22}P_{G,2}+N_{21}P_{G,1})(N_{21}-N_{22}A_{22}^{-1}A_{21})^Tc-(A_{22}P_{G,2}+A_{21}P_{G,1})(A_{22}^{-1}A_{21})^T)) \\ &\quad + \operatorname{tr}(\Sigma_2(B_2B_2^T-(N_{21}-N_{22}A_{22}^{-1}A_{21})P_R(N_{21}-N_{22}A_{22}^{-1}A_{21})^Tc)), \end{aligned}$$

where $P_{G,1}$ are the first r and $P_{G,2}$ the last $n-r$ rows of P_G, $c = \mathbb{E}\left[M^2(1)\right]$ and $\Sigma_2 = \operatorname{diag}(\sigma_{r+1},\dots,\sigma_n)$.

Proof. The right lower block of (3) satisfies

$$A_{22}^T\Sigma_2+\Sigma_2A_{22}+N_{22}^T\Sigma_2N_{22}c+N_{12}^T\Sigma_1N_{12}c=-C_2^TC_2. \tag{19}$$

If we multiply (3) with A^{-T} from the left hand side and select the left and right upper block of this equation, we obtain

$$\begin{aligned} \Sigma_1+\bar{A}^{-T}(\Sigma_1A_{11}-A_{21}^TA_{22}^{-T}\Sigma_2A_{21}+\bar{N}^T\Sigma_1N_{11}c+(N_{21}-N_{22}A_{22}^{-1}A_{21})^T\Sigma_2N_{21}c) &= -\bar{A}^{-T}\bar{C}^TC_1, \\ \bar{A}^{-T}(\Sigma_1A_{12}-A_{21}^TA_{22}^{-T}\Sigma_2A_{22}+\bar{N}^T\Sigma_1N_{12}c+(N_{21}-N_{22}A_{22}^{-1}A_{21})^T\Sigma_2N_{22}c) &= -\bar{A}^{-T}\bar{C}^TC_2. \end{aligned}$$

and thus

$$\bar{A}^T\Sigma_1+\Sigma_1A_{11}-A_{21}^TA_{22}^{-T}\Sigma_2A_{21}+\bar{N}^T\Sigma_1N_{11}c+(N_{21}-N_{22}A_{22}^{-1}A_{21})^T\Sigma_2N_{21}c=-\bar{C}^TC_1, \tag{20}$$

$$\Sigma_1A_{12}-A_{21}^TA_{22}^{-T}\Sigma_2A_{22}+\bar{N}^T\Sigma_1N_{12}c+(N_{21}-N_{22}A_{22}^{-1}A_{21})^T\Sigma_2N_{22}c=-\bar{C}^TC_2 \tag{21}$$

Furthermore, using (4) one can conclude

$$A_{11}\Sigma_1+\Sigma_1A_{11}^T+N_{11}\Sigma_1N_{11}^Tc+N_{12}\Sigma_2N_{12}^Tc=-B_1B_1^T \tag{22}$$

and

$$A_{22}\Sigma_2+\Sigma_2A_{22}^T+N_{22}\Sigma_2N_{22}^Tc+N_{21}\Sigma_1N_{21}^Tc=-B_2B_2^T. \tag{23}$$

From

$$\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix}P_{G,1}\\ P_{G,2}\end{bmatrix}+\begin{bmatrix}P_{G,1}\\ P_{G,2}\end{bmatrix}\bar{A}^T+\begin{bmatrix}N_{11} & N_{12}\\ N_{21} & N_{22}\end{bmatrix}\begin{bmatrix}P_{G,1}\\ P_{G,2}\end{bmatrix}\bar{N}^Tc=-\begin{bmatrix}B_1B_1^T\\ B_2B_1^T\end{bmatrix}$$

we also know that

$$A_{11}P_{G,1}+A_{12}P_{G,2}+P_{G,1}\bar{A}^T+N_{11}P_{G,1}\bar{N}^Tc+N_{12}P_{G,2}\bar{N}^Tc=-B_1B_1^T, \tag{24}$$

$$A_{21}P_{G,1}+A_{22}P_{G,2}+P_{G,2}\bar{A}^T+N_{22}P_{G,2}\bar{N}^Tc+N_{21}P_{G,1}\bar{N}^Tc=-B_2B_1^T. \tag{25}$$

We define $\mathcal{E} := \left(\operatorname{tr}\left(C\Sigma C^T\right) + \operatorname{tr}\left(\bar{C}P_R\bar{C}^T\right) - 2\ \operatorname{tr}\left(CP_G\bar{C}^T\right)\right)^{\frac{1}{2}}$ and obtain

$$\begin{aligned}\mathcal{E}^2 &= \operatorname{tr}\left(\begin{bmatrix}C_1 & C_2\end{bmatrix}\begin{bmatrix}\Sigma_1 & \\ & \Sigma_2\end{bmatrix}\begin{bmatrix}C_1^T \\ C_2^T\end{bmatrix}\right) + \operatorname{tr}\left(\bar{C}P_R\bar{C}^T\right) - 2\ \operatorname{tr}\left(\begin{bmatrix}C_1 & C_2\end{bmatrix}\begin{bmatrix}P_{G,1} \\ P_{G,2}\end{bmatrix}\bar{C}^T\right)\\ &= \operatorname{tr}(C_2\Sigma_2C_2^T + C_1\Sigma_1C_1^T + \bar{C}P_R\bar{C}^T - 2C_1P_{G,1}\bar{C}^T - 2C_2P_{G,2}\bar{C}^T).\end{aligned}$$

Using equation (21) yields

$$\begin{aligned}&\operatorname{tr}(-C_2P_{G,2}\bar{C}^T) = \operatorname{tr}(-\bar{C}^TC_2P_{G,2})\\ &= \operatorname{tr}(\Sigma_1A_{12}P_{G,2} - A_{21}^TA_{22}^{-T}\Sigma_2A_{22}P_{G,2} + \bar{N}^T\Sigma_1N_{12}P_{G,2}c + (N_{21} - N_{22}A_{22}^{-1}A_{21})^T\Sigma_2N_{22}P_{G,2}c)\\ &= \operatorname{tr}(A_{12}P_{G,2}\Sigma_1 - A_{21}^TA_{22}^{-T}\Sigma_2A_{22}P_{G,2} + N_{12}P_{G,2}\bar{N}^T\Sigma_1c + (N_{21} - N_{22}A_{22}^{-1}A_{21})^T\Sigma_2N_{22}P_{G,2}c).\end{aligned}$$

By equation (24) we obtain

$$\begin{aligned}\operatorname{tr}(-C_2P_{G,2}\bar{C}^T) =& \operatorname{tr}(-A_{21}^TA_{22}^{-T}\Sigma_2A_{22}P_{G,2} + (N_{21} - N_{22}A_{22}^{-1}A_{21})^T\Sigma_2N_{22}P_{G,2}c)\\ &- \operatorname{tr}((B_1B_1^T + P_{G,1}\bar{A}^T + A_{11}P_{G,1} + N_{11}P_{G,1}\bar{N}^Tc)\Sigma_1).\end{aligned}$$

Using equation (20), we have

$$\begin{aligned}&\operatorname{tr}(P_{G,1}\bar{A}^T + A_{11}P_{G,1} + N_{11}P_{G,1}\bar{N}^Tc)\Sigma_1) = \operatorname{tr}(\bar{A}^T\Sigma_1 + \Sigma_1A_{11} + \bar{N}^T\Sigma_1N_{11}c)P_{G,1})\\ &= -\operatorname{tr}(\bar{C}^TC_1P_{G,1} + (N_{21} - N_{22}A_{22}^{-1}A_{21})^T\Sigma_2N_{21}P_{G,1}c - (A_{22}^{-1}A_{21})^T\Sigma_2A_{21}P_{G,1})\end{aligned}$$

and hence,

$$\begin{aligned}\mathcal{E}^2 =& \operatorname{tr}(C_2\Sigma_2C_2^T + C_1\Sigma_1C_1^T + \bar{C}P_R\bar{C}^T) - 2\operatorname{tr}(B_1B_1^T\Sigma_1)\\ &+ 2\operatorname{tr}(-(A_{22}^{-1}A_{21})^T\Sigma_2A_{22}P_{G,2} + (N_{21} - N_{22}A_{22}^{-1}A_{21})^T\Sigma_2N_{22}P_{G,2}c)\\ &+ 2\operatorname{tr}((N_{21} - N_{22}A_{22}^{-1}A_{21})^T\Sigma_2N_{21}P_{G,1}c - (A_{22}^{-1}A_{21})^T\Sigma_2A_{21}P_{G,1}).\end{aligned}$$

Thus,

$$\begin{aligned}\mathcal{E}^2 =& \operatorname{tr}(\Sigma_2(C_2^TC_2 - 2(A_{22}P_{G,2} + A_{21}P_{G,1})(A_{22}^{-1}A_{21})^T)\\ &+ \operatorname{tr}(2\Sigma_2(N_{22}P_{G,2} + N_{21}P_{G,1})(N_{21} - N_{22}A_{22}^{-1}A_{21})^Tc))\\ &+ \operatorname{tr}(C_1\Sigma_1C_1^T + \bar{C}P_R\bar{C}^T - 2B_1B_1^T\Sigma_1).\end{aligned}\tag{26}$$

By definition, the Gramians P_R and Q_R satisfy

$$\bar{A}^TQ_R + Q_R\bar{A} + \bar{N}^TQ_R\bar{N}c = -\bar{C}^T\bar{C}\tag{27}$$

and

$$\bar{A}P_R + P_R\bar{A}^T + \bar{N}P_R\bar{N}^Tc = -B_1B_1^T,\tag{28}$$

such that one can conclude $\operatorname{tr}(\bar{C}P_R\bar{C}^T) = \operatorname{tr}(B_1^TQ_RB_1)$ from inserting (27) into $\operatorname{tr}(\bar{C}P_R\bar{C}^T)$. Consequently,

$$\operatorname{tr}(C_1\Sigma_1C_1^T + \bar{C}P_R\bar{C}^T - 2B_1B_1^T\Sigma_1) = \operatorname{tr}(C_1\Sigma_1C_1^T - B_1B_1^T\Sigma_1) + \operatorname{tr}(B_1^T(Q_R - \Sigma_1)B_1).$$

Inserting equation (22) provides

$$\begin{aligned}\operatorname{tr}(-B_1B_1^T\Sigma_1) &= \operatorname{tr}(A_{11}\Sigma_1\Sigma_1 + \Sigma_1A_{11}^T\Sigma_1 + N_{11}\Sigma_1N_{11}^Tc\Sigma_1 + N_{12}\Sigma_2N_{12}^Tc\Sigma_1)\\ &= \operatorname{tr}(\Sigma_1\Sigma_1A_{11} + \Sigma_1A_{11}^T\Sigma_1 + \Sigma_1N_{11}^T\Sigma_1N_{11}c + N_{12}\Sigma_2N_{12}^T\Sigma_1c)\\ &= -\operatorname{tr}(\Sigma_1C_1^TC_1) - \operatorname{tr}(\Sigma_1N_{21}^T\Sigma_2N_{21}c) + \operatorname{tr}(N_{12}\Sigma_2N_{12}^T\Sigma_1c).\end{aligned}$$

So, it holds

$$\mathrm{tr}(C_1\Sigma_1 C_1^T - B_1 B_1^T \Sigma_1) = \mathrm{tr}(N_{12}\Sigma_2 N_{12}^T \Sigma_1 c) - \mathrm{tr}(\Sigma_1 N_{21}^T \Sigma_2 N_{21} c).$$

From (19) it follows

$$\begin{aligned}\mathrm{tr}(\Sigma_2 N_{12}^T \Sigma_1 N_{12} c) &= \mathrm{tr}(-\Sigma_2(A_{22}^T\Sigma_2 + \Sigma_2 A_{22} + N_{22}^T\Sigma_2 N_{22}c + C_2^T C_2)) \\ &= \mathrm{tr}(-\Sigma_2(\Sigma_2 A_{22}^T + A_{22}\Sigma_2 + N_{22}\Sigma_2 N_{22}^T c + C_2^T C_2)).\end{aligned}$$

Using (23) yields

$$\mathrm{tr}(\Sigma_2 N_{12}^T \Sigma_1 N_{12} c) = \mathrm{tr}(\Sigma_2(N_{21}\Sigma_1 N_{21}^T c + B_2 B_2^T - C_2^T C_2)),$$

such that

$$\mathrm{tr}(C_1\Sigma_1 C_1^T - B_1 B_1^T \Sigma_1) = \mathrm{tr}(\Sigma_2(B_2 B_2^T - C_2^T C_2)).$$

Below, we analyze the term $\mathrm{tr}(B_1^T(Q_R - \Sigma_1)B_1)$. First, notice that the following holds:

$$\bar{A}^T\Sigma_1 + \Sigma_1\bar{A} + \bar{N}^T\Sigma_1\bar{N}c = -\bar{C}^T\bar{C} - (N_{21} - N_{22}A_{22}^{-1}A_{21})^T\Sigma_2(N_{21} - N_{22}A_{22}^{-1}A_{21})c.$$

With (27) we thus know that

$$\begin{aligned}&\bar{A}^T(Q_R - \Sigma_1) + (Q_R - \Sigma_1)\bar{A} + \bar{N}^T(Q_R - \Sigma_1)\bar{N}c \\ &= (N_{21} - N_{22}A_{22}^{-1}A_{21})^T\Sigma_2(N_{21} - N_{22}A_{22}^{-1}A_{21})c.\end{aligned} \quad (29)$$

Applying the equations (28) and (29) yields

$$\begin{aligned}\mathrm{tr}(B_1^T(Q_R - \Sigma_1)B_1) &= -\,\mathrm{tr}((\bar{A}P_R + P_R\bar{A}^T + \bar{N}P_R\bar{N}^T c)(Q_R - \Sigma_1)) \\ &= -\,\mathrm{tr}(P_R((Q_R - \Sigma_1)\bar{A} + \bar{A}^T(Q_R - \Sigma_1) + \bar{N}^T(Q_R - \Sigma_1)\bar{N}c)) \\ &= -\,\mathrm{tr}(P_R(N_{21} - N_{22}A_{22}^{-1}A_{21})^T\Sigma_2(N_{21} - N_{22}A_{22}^{-1}A_{21})c).\end{aligned}$$

We apply these results to (26) and obtain

$$\begin{aligned}\mathcal{E}^2 =& \,\mathrm{tr}(2\Sigma_2((N_{22}P_{G,2} + N_{21}P_{G,1})(N_{21} - N_{22}A_{22}^{-1}A_{21})^T c - (A_{22}P_{G,2} + A_{21}P_{G,1})(A_{22}^{-1}A_{21})^T)) \\ &+ \mathrm{tr}(\Sigma_2(B_2B_2^T - (N_{21} - N_{22}A_{22}^{-1}A_{21})P_R(N_{21} - N_{22}A_{22}^{-1}A_{21})^T c)).\end{aligned}$$

□

The error bound representation in Theorem 4.1 depends on Σ_2 which contains the $n-r$ smallest Hankel singular values $\sigma_{r+1}, \dots, \sigma_n$ of the original system. In case these values are small, the reduced order model obtained by the SPA is of good quality.

5 Comparison between balanced truncation and singular perturbation approximation

In this section, we compare BT which is discussed in [4] and the SPA which we consider in this paper. The aim is to point out the cases, when the SPA is better to motivate the practical relevance of this method. We start with an example which we obtain by discretizing an SPDE in the spatial component and afterwards we state a random example to illustrate further effects. Both examples are not in the balanced form already but balancing these systems can be done easily by the procedure stated in Section 4 in [4].
The numerical experiments are run on a desktop computer with a dual-core Intel Pentium processor

E5400 and 3GB RAM. All algorithms are implemented and executed in MATLAB 7.14.0.739 (R2012a) running on Ubuntu 10.04.1 LTS.

5.1 SPDE example

To compare BT and the SPA we use an example created in [17]. There, a second order SPDE with Lévy noise is considered and approximated by a large scale system of ordinary SDEs.

Example 5.1. *The lateral displacement of an electricity cable impacted by wind can be modeled by*

$$\frac{\partial^2}{\partial t^2}X(t,\zeta)+2\frac{\partial}{\partial t}X(t,\zeta)=\frac{\partial^2}{\partial \zeta^2}X(t,\zeta)+\mathrm{e}^{-(\zeta-\frac{\pi}{2})^2}\,u(t)+2\,\mathrm{e}^{-(\zeta-\frac{\pi}{2})^2}\,X(t-,\zeta)\frac{\partial}{\partial t}M(t)$$

for $t\in[0,T]$ *and* $\zeta\in[0,\pi]$. *Here,* $M(t)=-(N(t)-t)$ *with* $(N(t))_{t\geq 0}$ *being a Poisson process with parameter* 1. *The boundary and initial conditions are*

$$X(0,t)=0=X(\pi,t)\ \textit{and}\ X(0,\zeta),\ \frac{\partial}{\partial t}X(t,\zeta)\bigg|_{t=0}\equiv 0.$$

The output is an approximation for the position of the middle of the string

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

where $\epsilon>0$.

We introduce the following approximating SDE with state space dimension n, initial condition $\mathcal{X}(0)=0$ and output Y_n:

$$d\mathcal{X}(t)=[A\mathcal{X}(t)+Bu(t)]\,dt+N\mathcal{X}(s-)dM(s),\quad Y_n(t)=C\mathcal{X}(t),\quad t\geq 0, \tag{30}$$

where

- $A=\operatorname{diag}\left(E_1,\ldots,E_{\frac{n}{2}}\right)$ with $E_\ell=\left(\begin{smallmatrix}0 & \ell\\ -\ell & -\alpha\end{smallmatrix}\right)$,
- $B=(b_k)_{k=1,\ldots,n}$ with

$$b_{2\ell-1}=0,\quad b_{2\ell}=\sqrt{\frac{2}{\pi}}\left\langle \mathrm{e}^{-(\cdot-\frac{\pi}{2})^2},\sin(\ell\cdot)\right\rangle_H,$$

- $N=(n_{kj})_{k,j=1,\ldots,n}$ with

$$n_{(2\ell-1)j}=0,\quad n_{(2\ell)j}=\begin{cases}0, & \text{if } j=2v,\\ \frac{4}{\pi v}\left\langle\sin(\ell\cdot),\mathrm{e}^{-(\cdot-\frac{\pi}{2})^2}\sin(v\cdot)\right\rangle_H, & \text{if } j=2v-1,\end{cases}$$

  for $j=1,\ldots,n$ and $v=1,\ldots,\frac{n}{2}$,
- the output matrix C is given by $C^T=(c_k)_{k=1,\ldots,n}$ with

$$c_{2\ell}=0 \text{ and } c_{2\ell-1}=\frac{1}{\sqrt{2\pi}\ell^2}\left[\cos\left(\ell\left(\frac{\pi}{2}-\epsilon\right)\right)-\cos\left(\ell\left(\frac{\pi}{2}+\epsilon\right)\right)\right],$$

  where we assume n to be even, $\ell=1,\ldots,\frac{n}{2}$ and $H=L^2([0,\pi])$.

Following the arguments in [17] this approximation is meaningful, since

$$\mathbb{E}\left|Y_n(t)-Y(t)\right|^2\to 0$$

for $n \to \infty$ and $t \geq 0$. Now, we fix the dimension of (30) to $n = 1000$. The uncontrolled state equation is asymptotically mean square stable (see [17]) which means that

$$\mathbb{E} \left\| \mathcal{X}(t, x_0, 0) \right\|_2^2 \to 0$$

for $t \to \infty$ and any initial condition such that we can apply balanced truncation and the singular perturbation approximation, respectively below. We consider the deviation between Y_n and the outputs of the ROMs via BT and via the SPA in the norm on the left hand side of (18). We insert particular normalized control functions $u_1(t) = \sqrt{\frac{2}{\pi}} 1_{[0, \frac{\pi}{2}]}(t)$ and $u_2(t) = \frac{\sqrt{8}}{\pi} 1_{[0, \frac{\pi}{2}]}(t) w(t)$ ($t \in [0, \pi]$), where w is a Wiener process. The exact errors and the error bound $\mathcal{E}_1$ of BT we take from [17] and we additionally determine these values for the SPA, where $\mathcal{E}_2$ denotes the corresponding error bound stated in Theorem 4.1. Furthermore, we set $(\bar{B}, \bar{B}_0, \bar{D}) = (B_1, 0, 0)$.

Dim. ROM	BT Exact Error ($u = u_1$)	BT Exact Error ($u = u_2$)	Bound $\mathcal{E}_1$
40	$1.4484 \cdot 10^{-6}$	$1.1182 \cdot 10^{-6}$	$4.0103 \cdot 10^{-5}$
20	$7.2173 \cdot 10^{-6}$	$8.5996 \cdot 10^{-6}$	$1.2695 \cdot 10^{-4}$
10	$5.1396 \cdot 10^{-5}$	$3.8038 \cdot 10^{-5}$	$3.6395 \cdot 10^{-4}$
5	$5.2740 \cdot 10^{-4}$	$4.3632 \cdot 10^{-4}$	$2.3446 \cdot 10^{-3}$
3	0.0113	$8.6287 \cdot 10^{-3}$	0.0380

Dim. ROM	SPA Exact Error ($u = u_1$)	SPA Exact Error ($u = u_2$)	Bound $\mathcal{E}_2$
40	$2.0858 \cdot 10^{-6}$	$1.8654 \cdot 10^{-6}$	$4.1799 \cdot 10^{-5}$
20	$8.3989 \cdot 10^{-6}$	$1.0239 \cdot 10^{-5}$	$1.2808 \cdot 10^{-4}$
10	$5.6005 \cdot 10^{-5}$	$3.9154 \cdot 10^{-5}$	$3.4039 \cdot 10^{-4}$
5	$6.4096 \cdot 10^{-4}$	$6.5180 \cdot 10^{-4}$	$2.3876 \cdot 10^{-3}$
3	0.0183	0.0148	0.0629

From the numerical results above we see that BT is slightly better than the SPA on a time interval $[0, \pi]$. We are also interested in the long run behavior of the system (30). Therefore, we increase the length of the time interval and consider (30) on $[0, 8.5\pi]$ next and repeat the procedure. This is done due to the expected zero steady-state error that is known for the deterministic case. Again, we use normalized controls $\tilde{u}_1(t) = \sqrt{\frac{2}{8.5\pi}} 1_{[0, \frac{8.5\pi}{2}]}(t)$ and $\tilde{u}_2(t) = \frac{\sqrt{8}}{8.5\pi} 1_{[0, \frac{8.5\pi}{2}]}(t) w(t)$ ($t \in [0, 8.5\pi]$) and obtain better results for the SPA for growing dimensions of the ROM.

Dim. ROM	Error SPA ($u = \tilde{u}_1$)	Error BT ($u = \tilde{u}_1$)	Error SPA ($u = \tilde{u}_2$)	Error BT ($u = \tilde{u}_2$)
20	$6.0324 \cdot 10^{-6}$	$1.1826 \cdot 10^{-5}$	$5.0916 \cdot 10^{-6}$	$8.0841 \cdot 10^{-6}$
10	$2.5374 \cdot 10^{-5}$	$3.6819 \cdot 10^{-5}$	$2.7988 \cdot 10^{-5}$	$4.1106 \cdot 10^{-5}$
6	$8.1671 \cdot 10^{-5}$	$1.0461 \cdot 10^{-4}$	$7.5523 \cdot 10^{-5}$	$9.3595 \cdot 10^{-5}$
5	$7.2951 \cdot 10^{-4}$	$6.7156 \cdot 10^{-4}$	$9.5126 \cdot 10^{-4}$	$8.1803 \cdot 10^{-4}$
3	0.0195	0.0106	0.0154	0.0100

Below, we would like to compare different outputs visually. Since the reduced order models of BT and the SPA are quite accurate it is not possible to distinguish between the trajectories. For that reason, we create a random example in the next section.

5.2 Random example

Here, we consider an example of the form (30) with Wiener noise which we generate as follows: We set the state space dimension of the original model to $n = 500$, the reduced order system dimension to $r = 2$ and

$$A = JDJ^{-1}$$

with

$$D = -\operatorname{diag}(10\,\text{abs}(\text{randn}(n,1))) - 2I_n \text{ and } J = \text{randn}(n),$$

where we use "randn('state',1)" for D and "randn('state',2)" for J. The matrices B, C, N are also random and generated by

$$B = \text{randn}(n,n),\ C = \text{randn}(1,n) \text{ and } N = \text{rand}(n)/100,$$

where we use "rand('state',1)" for N, "rand('state',3)" for B and "rand('state',4)" for C. One can check numerically that there is a positive definite solution X to

$$A^T X + XA + N^T XN = -I.$$

By Theorem 2.1 this mean that the system is asymptotically mean square stable. We insert the controls u_i $(i = 1, \dots, n)$ on $[0, 12]$

$$u_i(t) = \begin{cases} k_i & \text{if } t \in [0,2] \cup [5,7] \\ 0 & \text{else,} \end{cases}$$

where the k_i are randomly generated constants. In Figure 1 we visualize a trajectory of the output

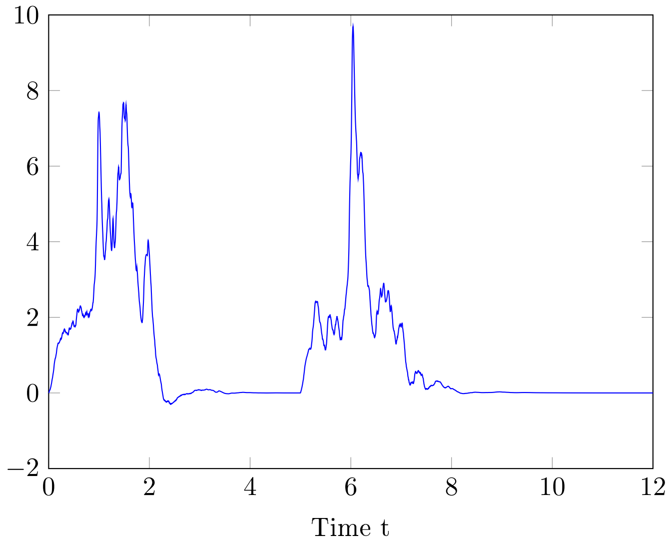

Figure 1: Output $Y(\omega, \cdot)$ of the original system

of the original model and in Figure 2 we compare the pointwise error of BT with the pointwise error of the SPA for a particular trajectory. If the graph in Figure 2 is above the red line, then the SPA is better. From the two plots we observe that the SPA is a better approximation if the output curve is flat. In this case, it seems to be a good assumption to suppose certain state components to be constant (symbolic derivative $\frac{dX_2}{dt} = 0$, see (7)), whereas BT provides a smaller error, where the slope of the output is big.

6 Conclusions

We generalized the singular perturbation approximation for stochastic systems with noise processes having jumps as an alternative to balanced truncation. In particular, we focused on a linear controlled state equation driven by a Lévy process which is asymptotically mean square stable and

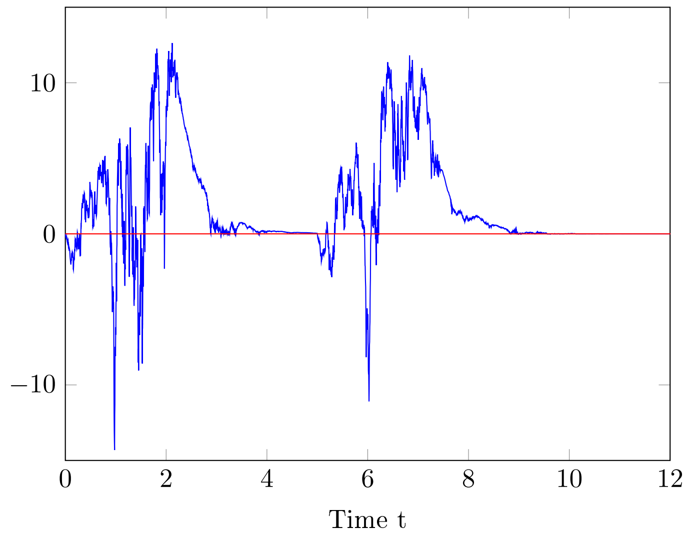

Figure 2: $|Y(\omega,\cdot) - Y_{BT}(\omega,\cdot)| - |Y(\omega,\cdot) - Y_{SPA}(\omega,\cdot)|$

equipped with an output equation. We showed that the reduced order model is mean square stable, but the question of preserving the asymptotic mean square stability is still open. Additionally, we demonstrated the possibility to lose complete observability and reachability, we provided an error bound for the singular perturbation approximation of Lévy driven systems and pointed out the cases in which the approach is good. Finally, we compared balanced truncation and the singular perturbation approximation for stochastic systems. We applied it in the context of the numerical solution of linear controlled SPDEs with Lévy noise and to a random example. There, we emphasized the advantages of using the singular perturbation approximation.

References

[1] Athanasios C. Antoulas. *Approximation of large-scale dynamical systems.* Advances in Design and Control 6. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2005.

[2] Peter Benner and Tobias Damm. Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems. *SIAM J. Control Optim.*, 49(2):686–711, 2011.

[3] Peter Benner, Tobias Damm, Martin Redmann, and Yolanda Rocio Rodriguez Cruz. Positive Operators and Stable Truncation. *Linear Algebra Appl.*, 2014.

[4] Peter Benner and Martin Redmann. Model Reduction for Stochastic Systems. *Stoch PDE: Anal Comp*, 3(3):291–338, 2015.

[5] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions.* Encyclopedia of Mathematics and Its Applications. 44. Cambridge etc.: Cambridge University Press, 1992.

[6] Tobias Damm. *Rational Matrix Equations in Stochastic Control.* Lecture Notes in Control and Information Sciences 297. Berlin: Springer, 2004.

[7] K. V. Fernando and H Nicholson. Singular perturbational model reduction of balanced systems. *IEEE Trans. Automat. Control*, 27:466–468, 1982.

[8] Wilfried Grecksch and Peter E. Kloeden. Time-discretised Galerkin approximations of parabolic stochastic PDEs. *Bull. Aust. Math. Soc.*, 54(1):79–85, 1996.

[9] Carsten Hartmann, Boris Schäfer-Bung, and Anastasia Thöns-Zueva. Balanced averaging of bilinear systems with applications to stochastic control. *SIAM J. Control Optim.*, 51(3):2356–2378, 2013.

[10] Erika Hausenblas. Approximation for Semilinear Stochastic Evolution Equations. *Potential Anal.*, 18(2):141–186, 2003.

[11] Arnulf Jentzen and Peter E. Kloeden. Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space-time noise. *Proc. R. Soc. A 2009*, 465:649–667, 2009.

[12] Yi Liu and Brian D.O. Anderson. Singular perturbation approximation of balanced systems. *Int. J. Control*, 50(4):1379–1405, 1989.

[13] Bruce C. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Trans. Autom. Control*, 26:17–32, 1981.

[14] Goro Obinata and Brian D.O. Anderson. *Model Reduction for Control System Design.* Communications and Control Engineering Series. London: Springer, 2001.

[15] Szymon Peszat and Jerzy Zabczyk. *Stochastic Partial Differential Equations with Lévy Noise. An evolution equation approach.* Encyclopedia of Mathematics and Its Applications 113. Cambridge: Cambridge University Press, 2007.

[16] Claudia Prévôt and Michael Röckner. *A Concise Course on Stochastic Partial Differential Equations.* Lecture Notes in Mathematics 1905. Berlin: Springer, 2007.

[17] Martin Redmann and Peter Benner. Approximation and Model Order Reduction for Second Order Systems with Levy-Noise. *AIMS Proceedings*, accepted: 12 August 15.