MAX-PLANCK-GESELLSCHAFT

Martin Redmann Peter Benner

Model Reduction for Stochastic Systems

MAX–PLANCK–INSTITUT FÜR DYNAMIK KOMPLEXER TECHNISCHER SYSTEME MAGDEBURG

Max Planck Institute Magdeburg
Preprints

MPIMD/14-03 February 13, 2014

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. If one uses the well known Galerkin scheme one can end up with a sequence of ordinary stochastic linear equations of high order. To reduce the high dimension for practical computations we consider balanced truncation being a model order reduction technique known from deterministic control theory. So, we generalize balanced truncation for controlled linear systems with Levy noise, discuss properties of the reduced order model, provide an error bound and give some examples.

1 Introduction

Model order reduction is of major importance for example in the field of control theory. A commonly used method is balanced truncation, which was first introduced by Moore [18] for linear deterministic systems and where a good overview containing all results of this scheme is stated in Antoulas [1]. Balanced truncation also works for deterministic bilinear equations (see Benner, Damm [4] and Zhang and others [12]). Benner and Damm additionally pointed out the relation between balanced truncation for deterministic bilinear control systems and linear stochastic systems with Wiener noise. So, in both cases the reachability and observability Gramians are solutions of generalized Lyapunov equations. We resume with working on balanced truncation for stochastic system and want to generalize the results known for the Wiener case. Thereby, the main idea is to allow the states to have jumps. Furthermore, we want to ensure that the Gramians we define are still solutions of generalized Lyapunov equations. So, a convenient noise process is given by a square integrable Levy process.

In Section 2 we provide the necessary background on semimartingales, square integrable Levy processes and stochastic calculus in order to render this paper as self-contained as possible. Detailed information regarding general Levy processes one can find in Bertoin [6] and Sato [23] and we refer to Applebaum [2] and Kuo [16] for an extended version of stochastic integration theory. In Section 3 we focus on a linear controlled state equation driven by uncorrelated Levy processes, which is asymptotically mean square stable and equipped with an output equation. We introduce the fundamental solution Φ of the state equation and point out the differences compared to fundamental solutions of deterministic systems. Using Φ we introduce reachability and observability Gramians the same way like Benner, Damm [4] and proof that the reachable and observable (average) states and the corresponding energy are characterized by these Gramians. In Section 4 we describe the procedure of balanced truncation for the linear system with Levy noise, which is similar to the procedure known from the deterministic case (see Antoulas [1] and Obinata, Anderson [19]). We discuss properties of the resulting reduced order model (ROM). We will show that it is mean square stable, not balanced, the Hankel singular values (HV) of the ROM are not a subset of the HVs of the original system and one can lose complete observability and reachability. Also we discuss the open question of the preservation of mean square asymptotic stability. Finally, we provide an error bound for balanced truncation of the Levy driven system assuming mean square asymptotic stability of the ROM. This error bound has the same structure as the $\mathcal{H}^2$ error bound of linear deterministic systems. In Section 5 we deal with a linear controlled stochastic evolution equation with Levy noise (compare Da Prato and Zabczyk [8], Peszat and Zabczyk [20], Prévôt and Röckner [21]). To solve such a problem numerically, finite dimensional approximations are commonly used. The scheme we state here is the well known Galerkin method (see Grecksch, Kloeden [10]) leading to a sequence of ordinary stochastic differential equation of that kind we considered in Section 4. For a particular case we apply balanced truncation on that Galerkin solution and compute the error bounds and the exact errors of the approximation.

2 Basics from Stochastics

Let all stochastic processes appearing in this section be defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$[1]. The set of all cadlag[2] square integrable $\mathbb{R}$-valued martingales with respect to $(\mathcal{F}_t)_{t\geq 0}$ we denote by $\mathcal{M}^2(\mathbb{R})$.

2.1 Semimartingales and Ito's formula

Below, we introduce the class of semimartingales.

Definition 2.1. *(i) An $(\mathcal{F}_t)_{t\geq 0}$-adapted cadlag process X with values in $\mathbb{R}$ is called* semimartingale *if it has the representation $X = X_0 + M + A$. Thereby, X_0 is an $\mathcal{F}_0$-measurable random variable, $M \in \mathcal{M}^2(\mathbb{R})$ and A is a cadlag process of bounded variation.*[3]

(ii) An $\mathbb{R}^d$-valued process $\vec{X}$ is called semimartingale *if all components are real-valued semimartingales.*

The following is based on Proposition 17.2 in [17].

Proposition 2.2. *Let $M, N \in \mathcal{M}^2(\mathbb{R})$, then there exists a unique predictable*[4] *process $\langle M, N\rangle$ of bounded variation such that $MN - \langle M, N\rangle$ is a martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$.*

Next, we consider a decomposition of square integrable martingales (see Theorem 4.18 in [14]).

Theorem 2.3. *A process $M \in \mathcal{M}^2(\mathbb{R})$ has the following representation:*

$$M(t) = M_0 + M^c(t) + M^d(t), \quad t \geq 0,$$

where $M^c(0) = M^d(0) = 0$, M_0 is $\mathcal{F}_0$-measurable random variable, M^c is a continuous process in $\mathcal{M}^2(\mathbb{R})$ and $M^d \in \mathcal{M}^2(\mathbb{R})$.

We need the quadratic covariation $[Z_1, Z_2]$ of two real-valued semimartingales Z_1 and Z_2, which can be introduced by

$$[Z_1, Z_2]_t := Z_1(t)Z_2(t) - Z_1(0)Z_2(0) - \int_0^t Z_1(s-)dZ_2(s) - \int_0^t Z_2(s-)dZ_1(s) \tag{1}$$

for $t \geq 0$. By the linearity of the integrals in (1) we obtain the property

$$[Z_1, Z_2]_t = \frac{1}{2}\left([Z_1 + Z_2, Z_1 + Z_2]_t - [Z_1, Z_1]_t - [Z_2, Z_2]_t\right), \quad t \geq 0.$$

From Theorem 4.52 in [14] we know that $[Z_1, Z_2]$ is also given by

$$[Z_1, Z_2]_t = \langle M_1^c, M_2^c\rangle_t + \sum_{0\leq s\leq t} \Delta Z_1(s)\Delta Z_2(s) \tag{2}$$

for $t \geq 0$, where M_1^c and M_2^c are the continuous martingale parts of Z_1 and Z_2. Furthermore, we set $\Delta Z(s) := Z(s) - Z(s-)$ with $Z(s-) := \lim_{t\uparrow s} Z(t)$ for a real-valued semimartingale Z. If we rearrange equation (1), we obtain the Ito product formula

$$Z_1(t)Z_2(t) = Z_1(0)Z_2(0) + \int_0^t Z_1(s-)dZ_2(s) + \int_0^t Z_2(s-)dZ_1(s) + [Z_1, Z_2]_t \tag{3}$$

for $t \geq 0$, which we use for the following corollaries:

[1] $(\mathcal{F}_t)_{t\geq 0}$ shall be right continuous and complete.
[2] Cadlag means that $\mathbb{P}$-almost all paths are right continuous and the left limits exist.
[3] This means that $\mathbb{P}$-almost all paths are of bounded variation.
[4] The process $\langle M, N\rangle$ is measurable with respect to $\mathcal{P}$, which we characterize below Definition 2.10.

Corollary 2.4. *Let Y and Z be two $\mathbb{R}^d$-valued semimartingales, then*

$$Y^T(t)Z(t) = Y^T(0)Z(0) + \int_0^t Z^T(s-)dY(s) + \int_0^t Y^T(s-)dZ(s) + \sum_{i=1}^d [Y_i, Z_i]_t$$

for all $t \geq 0$.

Proof. We have

$$\begin{aligned} Y^T(t)Z(t) &= \sum_{i=1}^d Y_i(t)Z_i(t) = \sum_{i=1}^d \left(Y_i(0)Z_i(0) + \int_0^t Z_i(s-)dY_i(s) + \int_0^t Y_i(s-)dZ_i(s) + [Y_i, Z_i]_t \right) \\ &= Y^T(0)Z(0) + \int_0^t Z^T(s-)dY(s) + \int_0^t Y^T(s-)dZ(s) + \sum_{i=1}^d [Y_i, Z_i]_t \end{aligned}$$

by applying the product formula in (3). □

Corollary 2.5. *Let Y be an $\mathbb{R}^d$-valued and Z be an $\mathbb{R}^n$-valued semimartingale, then*

$$Y(t)Z^T(t) = Y(0)Z^T(0) + \int_0^t dY(s)Z^T(s-) + \int_0^t Y(s-)dZ^T(s) + ([Y_i, Z_j]_t)_{\substack{i=1,\dots,d \\ j=1,\dots,n}}$$

for all $t \geq 0$.

Proof. We consider the stochastic differential of the ijth component of the matrix-valued process $Y(t)Z^T(t)$, $t \geq 0$, and obtain the following via the product formula in (3):

$$\begin{aligned} e_i^T Y(t)Z^T(t)e_j &= e_i^T Y(0)Z^T(0)e_j + \int_0^t Z^T(s-)e_j d(e_i^T Y(s)) + \int_0^t e_i^T Y(s-)d(Z^T(s)e_j) \\ &\quad + [e_i^T Y, Z^T e_j]_t \\ &= e_i^T Y(0)Z^T(0)e_j + e_i^T \int_0^t d(Y(s))Z^T(s-)e_j + e_i^T \int_0^t Y(s-)d(Z^T(s))e_j + [Y_i, Z_j]_t \end{aligned}$$

for all $t \geq 0$, $i \in \{1, \dots, d\}$ and $j \in \{1, \dots, n\}$, where e_i is the ith unit vector in $\mathbb{R}^d$ or in $\mathbb{R}^n$, respectively. Hence, in compact form we have

$$Y(t)Z^T(t) = Y(0)Z^T(0) + \int_0^t dY(s)Z^T(s-) + \int_0^t Y(s-)dZ^T(s) + ([Y_i, Z_j]_t)_{\substack{i=1,\dots,d \\ j=1,\dots,n}}$$

for all $t \geq 0$. □

2.2 Levy processes

Definition 2.6. *Let $L = (L(t))_{t \geq 0}$ be a cadlag stochastic process with values in $\mathbb{R}$ having independent and homogenous increments. If furthermore $L(0) = 0$ $\mathbb{P}$-almost surely holds and L is continuous in probability, then L is called (real-valued) Levy process.*

Below, we focus on Levy processes L being square integrable. The following theorem is proven analogously to the Theorem 4.44 in [20].

Theorem 2.7. *We set $\tilde{m} = \mathbb{E}[L(1)]$. For square integrable Levy processes L and $t, s \geq 0$ it holds*

$$\begin{aligned} &\mathbb{E}[L(t)] = t\mathbb{E}[L(1)] \text{ and} \\ &\mathrm{Cov}(L(s), L(t)) = \mathbb{E}[(L(t) - \tilde{m}t)(L(s) - \tilde{m}s)] = \min\{t, s\}\ \mathrm{Var}(L(1)). \end{aligned}$$

Proposition 2.8. *Let L be a square integrable Levy process adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$, such that the increments $L(t+h) - L(t)$ are independent of $\mathcal{F}_t$ $(t, h \geq 0)$, then L is a martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$ if and only if L has mean zero.*

Proof. First, we assume that L has mean zero, then the conditional expactation $\mathbb{E}\{L(t)|\mathcal{F}_s\}$ fulfills

$$\begin{aligned}\mathbb{E}\{L(t)|\mathcal{F}_s\} &= \mathbb{E}\{L(t) - L(s) + L(s)|\mathcal{F}_s\} = \mathbb{E}\{L(t) - L(s)|\mathcal{F}_s\} + L(s) \\ &= \mathbb{E}[L(t) - L(s)] + L(s) = L(s)\end{aligned}$$

for $0 \leq s < t$. If we know that L is a martingale it easily follows that it has a constant mean function, since

$$\mathbb{E}[L(t)] = \mathbb{E}[\mathbb{E}\{L(t)|\mathcal{F}_s\}] = \mathbb{E}[L(s)]$$

for $0 \leq s < t$. But by Theorem 2.7 we know that the mean function is linear. Thus, $\mathbb{E}[L(t)] = 0$ for all $t \geq 0$. $\square$

We set $M(t) := L(t) - t\mathbb{E}[L(1)]$, $t \geq 0$, where L is square integrable. By Proposition 2.8 M is a square integrable martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$ and a Levy process as well. So, we have the following representation for square intergrable Levy processes L:

$$L(t) = M(t) + \mathbb{E}[L(1)]t, \quad t \geq 0.$$

The compensator $\langle M, M \rangle$ of M is deterministic and continuous and given by

$$\langle M, M \rangle_t = \mathbb{E}\left[M^2(t)\right] = \mathbb{E}\left[M^2(1)\right] t,$$

because $M^2(t) - \mathbb{E}\left[M^2(1)\right] t$, $t \geq 0$, is a martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$.

2.3 Stochastic integration

We assume that $M \in \mathcal{M}^2(\mathbb{R})$. The definition of an integral with respect to M is similar to that with respect to a Wiener process W. This makes things comfortable. A definition for an integral based on W can for example be found in Applebaum [2], Arnold [3] and Kloeden, Platen [15]. Furthermore, Applebaum [2] gives a definition of an integral with respect to the so called "martingale-valued measures", which is a generalization of the integral introduced here. The definition of the integral with respect to M we take from Chapter 6.5 in the book of Kuo [16].

Fist of all, we characterize the class of simple processes.

Definition 2.9. *A process $\Psi = (\Psi(t))_{t\in[0,T]}$ is called simple if it has the following representation:*

$$\Psi(s) = \sum_{i=0}^{m} \chi_{(t_i, t_{i+1}]}(s)\Psi_i, \quad s \in [0, T], \tag{4}$$

for $0 = t_0 < t_1 < \ldots < t_{m+1} = T$. Here, the random variables Ψ_i are $\mathcal{F}_{t_i}$-measurable and bounded, $i \in \{0, 1, \ldots, m\}$.

For simple processes Ψ we define

$$I_T^M(\Psi) := \int_0^T \Psi(s)dM(s) := \sum_{i=0}^{m} \Psi_i\left(M(t_{i+1}) - M(t_i)\right)$$

and for $0 \leq t_0 \leq t \leq T$ we set

$$I_{t_0,t}^M(\Psi) := \int_{t_0}^t \Psi(s)dM(s) := \int_0^T \chi_{[t_0,t]}(s)\Psi(s)dM(s).$$

Definition 2.10. *Let $(F(t))_{t\in[0,T]}$ be adapted to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ with left continuous trajectories. We define $\mathcal{P}_T$ as the smallest sub σ-algebra of $\mathcal{B}([0,T])\otimes\mathcal{F}$ with respect to which all mappings $F:[0,T]\times\Omega\to\mathbb{R}$ are measurable. $\mathcal{P}_T$ we call predictable σ-algebra.*

Remark. *$\mathcal{P}_T$ is generated as follows:*

$$\mathcal{P}_T = \sigma\left(\{(s,t]\times A\colon 0\le s\le t\le T, A\in\mathcal{F}_s\}\cup\{\{0\}\times A\colon A\in\mathcal{F}_0\}\right). \tag{5}$$

In Definition 2.10 we can replace the time interval $[0,T]$ by $\mathbb{R}_+$. Then the predictable σ-algebra is denoted by $\mathcal{P}$. $\mathcal{P}_T$- or $\mathcal{P}$-measurable processes we call predictable.

We want to extend the set of all integrable processes with respect to M. Therefor, we introduce $\mathcal{L}_T^2$ as the space of all predictable mappings Ψ on $[0,T]\times\Omega$ with $\|\Psi\|_T<\infty$, where

$$\|\Psi\|_T^2 := \mathbb{E}\int_0^T |\Psi(s)|^2\, d\langle M,M\rangle_s \tag{6}$$

and $\langle M,M\rangle$ is the compensator of M introduced in Proposition 2.2.

By Chapter 6.5 in Kuo [16] we can choose a sequence $(\Psi_n)_{n\in\mathbb{N}}\subset\mathcal{L}_T^2$ of simple processes, such that

$$\|\Psi_n-\Psi\|_T\to 0$$

for $\Psi\in\mathcal{L}_T^2$ and $n\to\infty$. Hence, we obtain that $\left(I_T^M(\Psi_n)\right)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega,\mathcal{F},\mathbb{P})$. Therefore, we can define

$$\int_0^T \Psi(s)dM(s) := L^2-\lim_{n\to\infty}\int_0^T \Psi_n(s)dM(s)$$

and for $0\le t_0\le t\le T$ we set

$$\int_{t_0}^t \Psi(s)dM(s) := L^2-\lim_{n\to\infty}\int_{t_0}^t \Psi_n(s)dM(s).$$

Here, "$L^2-\lim_{n\to\infty}$" denotes the limit in $L^2(\Omega,\mathcal{F},\mathbb{P})$.

By Theorem 6.5.8 in Kuo [16] the integral with respect to M has the following properties:

Theorem 2.11. *If $\Psi\in\mathcal{L}_T^2$ for $T>0$, then*

(i) the integral with respect to M has mean zero:

$$\mathbb{E}\left[\int_0^T \Psi(s)dM(s)\right]=0,$$

(ii) the second moment of $I_T^M(\Psi)$ is given by

$$\mathbb{E}\left|\int_0^T \Psi(s)dM(s)\right|^2 = \mathbb{E}\int_0^T |\Psi(s)|^2\, d\langle M,M\rangle_s$$

(iii) and the process

$$\left(\int_0^t \Psi(s)dM(s)\right)_{t\in[0,T]}$$

is a martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.

2.4 Levy type integrals

Below, we want to determine the mean of the quadratic covariation of the following processes:

$$\tilde{Z}_1(t) = \tilde{Z}_1(0) + \int_0^t A_1(s)ds + \sum_{i=1}^q \int_0^t B_1^i(s)dM^i(s), \quad t \geq 0,$$
$$\tilde{Z}_2(t) = \tilde{Z}_2(0) + \int_0^t A_2(s)ds + \sum_{i=1}^q \int_0^t B_2^i(s)dM^i(s), \quad t \geq 0,$$

where the processes M^i $(i = 1, \ldots, q)$ are uncorrelated scalar square integrable Levy processes with mean zero. In addition, the processes B_1^i, B_2^i are integrable with respect to M^i $(i = 1, \ldots, q)$, which by Section 2.3 means that they are predictable with

$$\mathbb{E} \int_0^t \left|B^i(s)\right|^2 ds < \infty, \quad t \geq 0,$$

considering (6) with $\langle M, M\rangle_t = \mathbb{E}\left[M^2(1)\right] t$. Furthermore, A_1, A_2 are $\mathbb{P}$-almost surely Lebesgue integrable and $(\mathcal{F}_t)_{t\geq 0}$-adapted.

We set $b_1(t) := \sum_{i=1}^q \int_0^t B_1^i(s)dM^i(s)$ and $b_2(t) := \sum_{i=1}^q \int_0^t B_2^i(s)dM^i(s)$ and obtain

$$\left[\tilde{Z}_1, \tilde{Z}_2\right]_t = [b_1, b_2]_t$$

for $t \geq 0$ considering equation (2) because $\tilde{Z}_i$ has the same jumps and the same martingale part as b_i $(i = 1, 2)$. We know that

$$[b_1, b_2]_t = \frac{1}{2}\left([b_1 + b_2, b_1 + b_2]_t - [b_1, b_1]_t - [b_2, b_2]_t\right) \tag{7}$$

for $t \geq 0$. Using the definition in (1) yields

$$[b_1, b_1]_t = (b_1(t))^2 - 2\int_0^t b_1(s-)db_1(s) = (b_1(t))^2 - 2\sum_{i=1}^q \int_0^t b_1(s-)B_1^i(s)dM^i(s).$$

This provides

$$\mathbb{E}\,[b_1, b_1]_t = \mathbb{E}\left[(b_1(t))^2\right].$$

Since M^i and M^j are uncorrelated processes for $i \neq j$, we have

$$\mathbb{E}\left[(b_1(t))^2\right] = \sum_{i=1}^q \mathbb{E}\left[\left(\int_0^t B_1^i(s)dM^i(s)\right)^2\right] = \sum_{i=1}^q \int_0^t \mathbb{E}\left[\left(B_1^i(s)\right)^2\right] ds \cdot c_i$$

applying Theorem 2.11 (ii), where $c_i := \mathbb{E}\left[\left(M^i(1)\right)^2\right]$. Hence,

$$\mathbb{E}\,[b_1, b_1]_t = \sum_{i=1}^q \int_0^t \mathbb{E}\left[\left(B_1^i(s)\right)^2\right] ds \cdot c_i.$$

Analogously, we can show that

$$\mathbb{E}\left[b_2, b_2\right]_t = \sum_{i=1}^{q} \int_0^t \mathbb{E}\left[\left(B_2^i(s)\right)^2\right] ds \cdot c_i \text{ and}$$
$$\mathbb{E}\left[b_1 + b_2, b_1 + b_2\right]_t = \sum_{i=1}^{q} \int_0^t \mathbb{E}\left[\left(B_1^i + B_2^i(s)\right)^2\right] ds \cdot c_i$$

holds for $t \geq 0$. Considering equation (7), we obtain

$$\mathbb{E}\left[\tilde{Z}_1, \tilde{Z}_2\right]_t = \mathbb{E}\left[b_1, b_2\right]_t = \sum_{i=1}^{q} \int_0^t \mathbb{E}\left[B_1^i B_2^i(s)\right] ds \cdot c_i. \tag{8}$$

At the end of this section we want to refer to Section 4.4.3 in Applebaum [2]. There one can find some remarks regarding the quadratic covariation of the Levy type integrals defined in that book.

3 Linear Control with Levy Noise

Before describing balanced truncation for the stochastic case, we define observability and reachability. Therefor, we introduce observability and reachability Gramians for our Levy driven system like Benner, Damm [4] do (Section 2.2). We additionally show that the sets of observable and reachable states are characterized by these Gramians. This is analogous to deterministic systems, where observability and reachability concepts are described in Sections 4.2.1 and 4.2.2 in Antoulas [1].

3.1 Reachability concept

Let $M_1, \dots, M_q$ be real-valued uncorrelated and square integrable Levy processes with mean zero defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.[5] In addition we assume M_k $(k = 1, \dots, q)$ to be $(\mathcal{F}_t)_{t\geq 0}$-adapted and the increments $M_k(t+h) - M_k(t)$ to be independent of $\mathcal{F}_t$ for $t, h \geq 0$. We consider the following equations:

$$dX(t) = [AX(t) + Bu(t)]dt + \sum_{k=1}^{q} \Psi^k X(t-) dM_k(t), \quad t \geq 0, \tag{9}$$
$$X(0) = x_0 \in \mathbb{R}^n,$$

where A, $\Psi^k \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$. With L^2_T we denote the space of all adapted stochastic 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 we call energy norm. It 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 $\mathbb{P} \otimes dt$. For the case $T = \infty$ we denote the space by L_2. Further, we assume the control $u \in L^2_T$ for every $T > 0$. We start with the definition of a solution of (9).

Definition 3.1. *An $\mathbb{R}^n$-valued and $(\mathcal{F}_t)_{t\geq 0}$-adapted cadlag process $(X(t))_{t\geq 0}$ is called solution of*

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

(9) if

$$X(t) = x_0 + \int_0^t [AX(s) + Bu(s)]ds + \sum_{k=1}^q \int_0^t \Psi^k X(s-)dM_k(s) \tag{10}$$

$\mathbb{P}$-almost surely holds for all $t \geq 0$.

Below, the solution of (9) at time $t \geq 0$ with initial condition $x_0 \in \mathbb{R}^n$ and given control u is always denoted by $X(t, x_0, u)$. For the homogeneous solution of (9) we briefly write $Y_{x_0} := X(t, x_0, 0)$. Furthermore, by $\|\cdot\|_2$ we denote the Euclidean norm. We assume the homogeneous solution to be asymptotically mean square stable, which means that

$$\mathbb{E}\,\|Y_{y_0}(t)\|_2^2 \to 0$$

for $t \to \infty$ and $y_0 \in \mathbb{R}^n$. This concept of stability is also used in Benner, Damm [4] and is necessary for defining (infinite) Gramians, which are introduced later.

Proposition 3.2. *Let Y_{y_0} be the homogeneous solution of (9) with any initial value $y_0 \in \mathbb{R}^n$, then $\mathbb{E}\left[Y_{y_0}(t)Y_{y_0}^T(t)\right]$ is the solution of the matrix integral equation*

$$\mathbb{Y}(t) = y_0 y_0^T + \int_0^t \mathbb{Y}(s)ds\ A^T + A\ \int_0^t \mathbb{Y}(s)ds + \sum_{k=1}^q \Psi^k \int_0^t \mathbb{Y}(s)ds\ \left(\Psi^k\right)^T\ \mathbb{E}\left[M_k(1)^2\right] \tag{11}$$

for $t \geq 0$.

Proof. We determine the stochastic differential of the matrix-valued process $Y_{y_0}Y_{y_0}^T$ via using the Ito formula in Corollary 2.5. This yields

$$Y_{y_0}(t)Y_{y_0}^T(t) = y_0 y_0^T + \int_0^t Y_{y_0}(s-)dY_{y_0}^T(s) + \int_0^t dY_{y_0}(s)Y_{y_0}^T(s-) + \left([e_i^T Y_{y_0}, Y_{y_0}^T e_j]_t\right)_{i,j=1,\dots,n},$$

where e_i is the ith unit vector. Hence,

$$\int_0^t Y_{y_0}(s-)dY_{y_0}^T(s) = \int_0^t Y_{y_0}(s-)Y_{y_0}^T(s)A^T ds + \sum_{k=1}^q \int_0^t Y_{y_0}(s-)Y_{y_0}^T(s-)(\Psi^k)^T dM_k(s) \text{ and}$$
$$\int_0^t dY_{y_0}(s)Y_{y_0}^T(s-) = \int_0^t AY_{y_0}(s)Y_{y_0}^T(s-)ds + \sum_{k=1}^q \int_0^t \Psi^k Y_{y_0}(s-)Y_{y_0}^T(s-)dM_k(s)$$

by inserting the stochastic differential of Y_{y_0}. Thus, by taking the expactation, we obtain

$$\mathbb{E}\left[Y_{y_0}(t)Y_{y_0}^T(t)\right] = y_0 y_0^T + \int_0^t \mathbb{E}\left[Y_{y_0}(s-)Y_{y_0}^T(s)\right] A^T ds + \int_0^t A\mathbb{E}\left[Y_{y_0}(s)Y_{y_0}^T(s-)\right] ds$$
$$+ \left(\mathbb{E}[e_i^T Y_{y_0}, Y_{y_0}^T e_j]_t\right)_{i,j=1,\dots,n}$$

applying Theorem 2.11 (i). Considering equation (8), we have

$$\mathbb{E}[e_i^T Y_{y_0}, Y_{y_0}^T e_j]_t = e_i^T \sum_{k=1}^q \int_0^t \mathbb{E}\left[\Psi^k Y_{y_0}(s)Y_{y_0}^T(s)\left(\Psi^k\right)^T\right] ds \cdot c_k e_j,$$

where $c_k := \mathbb{E}\left[M_k(1)^2\right]$. In addition, we use the property that a cadlag process has at most countably many jumps on a finite time interval (see Theorem 2.7.1 in Applebaum [2]), such that we can replace the left limit by the function value itself. Thus,

$$\mathbb{E}\left[Y_{y_0}(t)Y_{y_0}^T(t)\right] = y_0 y_0^T + \int_0^t \mathbb{E}\left[Y_{y_0}(s)Y_{y_0}^T(s)\right] ds\; A^T + A\int_0^t \mathbb{E}\left[Y_{y_0}(s)Y_{y_0}^T(s)\right] ds \qquad (12)$$
$$+\sum_{k=1}^q \Psi^k \int_0^t \mathbb{E}\left[Y_{y_0}(s)Y_{y_0}^T(s)\right] ds\; \left(\Psi^k\right)^T \cdot c_k.$$

□

We introduce an additional concept of stability of the homogeneous system corresponding to equation (9). We call Y_{y_0} exponentially mean square stable if there exist $c, \beta > 0$ such that

$$\mathbb{E}\left\|Y_{y_0}(t)\right\|_2^2 \leq \left\|y_0\right\|_2^2 c\, \mathrm{e}^{-\beta t}$$

for $t \geq 0$. This stability turns out to be equivalent to asymptotic mean square stability, which is stated in the next theorem.

Theorem 3.3. *The following are equivalent:*

(i) The homogeneous equation of (9) is asymptotically mean square stable.

(ii) The homogeneous equation of (9) is exponentially mean square stable.

(iii) The eigenvalues of $\left(I_n \otimes A + A \otimes I_n + \sum_{k=1}^q \Psi^k \otimes \Psi^k \cdot \mathbb{E}\left[M_k(1)^2\right]\right)$ *have negative real parts.*

Proof. Due to the similarity of the proofs we refer to Theorem 1.5.3 in Damm [9], where these results are proven for the Wiener case. □

As in the deterministic case, there exists a fundamental solution, which we define by

$$\Phi(t) := \left[Y_{e_1}(t), Y_{e_2}(t), \ldots, Y_{e_n}(t)\right]$$

for $t \geq 0$, where e_i is the ith unit vector ($i = 1, \ldots, n$). Thus, Φ fulfills the following integral equation:

$$\Phi(t) = I_n + \int_0^t A\Phi(s)ds + \sum_{k=1}^q \int_0^t \Psi^k \Phi(s-)dM_k(s).$$

The columns of Φ represent a minimal generating set such that we have $Y_{y_0}(t) = \Phi(t)y_0$. With $B = [b_1, b_2, \ldots, b_m]$ one can see that

$$\Phi(t)B = \left[\Phi(t)b_1, \Phi(t)b_2, \ldots, \Phi(t)b_m\right] = \left[Y_{b_1}(t), Y_{b_2}(t) \ldots, Y_{b_m}(t)\right].$$

Hence, we have

$$\Phi(t)BB^T\Phi^T(t) = Y_{b_1}(t)Y_{b_1}^T(t) + Y_{b_2}(t)Y_{b_2}^T(t) + \ldots + Y_{b_m}(t)Y_{b_m}^T(t),$$

such that

$$\mathbb{E}\left[\Phi(t)BB^T\Phi^T(t)\right] = BB^T + \int_0^t \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] ds\; A^T + A\int_0^t \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] ds \quad (13)$$
$$+\sum_{k=1}^q \Psi^k \int_0^t \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] ds\; (\Psi^k)^T\, \mathbb{E}\left[M_k(1)^2\right]$$

holds for every $t \geq 0$. Due to the assumption that the homogeneous solution Y_{y_0} is asymptotically mean square stable for an arbitrary initial value y_0, yielding $\mathbb{E}\left[Y_{y_0}^T(t)Y_{y_0}(t)\right] \to 0$ for $t \to \infty$, we obtain

$$0 = BB^T + \int_0^\infty \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] ds\ A^T + A\int_0^\infty \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] ds$$
$$+ \sum_{k=1}^q \Psi^k \int_0^\infty \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] ds\ (\Psi^k)^T\ \mathbb{E}\left[M_k(1)^2\right]$$

by taking the limit $t \to \infty$ in equation (13). Therefore, we can conclude that $P := \int_0^\infty \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] ds$, which exists by the asymptotic mean square stability assumption, is the solution of a generalized Lyapunov equation

$$AP + PA^T + \sum_{k=1}^q \Psi^k P\left(\Psi^k\right)^T\ \mathbb{E}\left[M_k(1)^2\right] = -BB^T.$$

P is the reachability Gramian of system (9), where this definition of the Gramian is also used in Benner, Damm [4] for stochastic systems driven by Wiener noise. Note that in this case $\mathbb{E}\left[M_k(1)^2\right] = 1$.

Remark. *The solution of the matrix equation*

$$0 = BB^T + AP + PA^T + \sum_{k=1}^q \Psi^k P(\Psi^k)^T \cdot \mathbb{E}\left[M_k(1)^2\right] \tag{14}$$

is unique if and only if the solution of

$$-\operatorname{vec}(BB^T) = \left(I_n \otimes A + A \otimes I_n + \sum_{k=1}^q \Psi^k \otimes \Psi^k \cdot \mathbb{E}\left[M_k(1)^2\right]\right)\operatorname{vec}(P)$$

is unique. By the assumption of mean square asymptotic stability the eigenvalues of $I\otimes A + A\otimes I + \sum_{k=1}^q \Psi^k \otimes \Psi^k \cdot \mathbb{E}\left[M_k(1)^2\right]$ *are non zero, hence the matrix equation (14) is uniquely solvable.*

More general, we consider stochastic processes $(\Phi(t,\tau))_{t\geq\tau}$ with starting time $\tau \geq 0$ and inital condition $\Phi(\tau,\tau) = I_n$ satisfying

$$\Phi(t,\tau) = I_n + \int_\tau^t A\Phi(s,\tau)ds + \sum_{k=1}^q \int_\tau^t \Psi^k\Phi(s-,\tau)dM_k(s) \tag{15}$$

for $t \geq \tau \geq 0$. Of course, we have $\Phi(t,0) = \Phi(t)$. Analogous to equation (13), we can show that

$$\begin{aligned}\mathbb{E}\left[\Phi(t,\tau)BB^T\Phi^T(t,\tau)\right] = BB^T &+ \int_\tau^t \mathbb{E}\left[\Phi(s,\tau)BB^T\Phi^T(s,\tau)\right] ds\ A^T \\ &+ A\int_\tau^t \mathbb{E}\left[\Phi(s,\tau)BB^T\Phi^T(s,\tau)\right] ds \\ &+ \sum_{k=1}^q \Psi^k \int_\tau^t \mathbb{E}\left[\Phi(s,\tau)BB^T\Phi^T(s,\tau)\right] ds\ (\Psi^k)^T\ \mathbb{E}\left[M_k(1)^2\right].\end{aligned} \tag{16}$$

This yields that $\mathbb{E}\left[\Phi(t,\tau)BB^T\Phi^T(t,\tau)\right]$ is the solution of the differential equation

$$\dot{\mathbb{Y}}(t) = A\mathbb{Y}(t) + \mathbb{Y}(t)A^T + \sum_{k=1}^q \Psi^k\mathbb{Y}(t)(\Psi^k)^T\ \mathbb{E}\left[M_k(1)^2\right] \tag{17}$$

for $t \geq \tau$ with initial condition $\mathbb{Y}(\tau) = BB^T$.

Remark. *For $t \geq \tau \geq 0$ we have $\Phi(t,\tau) = \Phi(t)\Phi^{-1}(\tau)$, since $\Phi(t)\Phi^{-1}(\tau)$ fulfills equation (15). Compared to the deterministic case ($\Psi^k = 0$) we do not have the semigroup property for the fundamental solution. So, it is not true that $\Phi(t,\tau) = \Phi(t-\tau)$ $\mathbb{P}$-almost surely holds, because the trajectories of the noise processes on $[0, t-\tau]$ and $[\tau, t]$ are different in general. We can however conclude that $\mathbb{E}\left[\Phi(t,\tau)BB^T\Phi^T(t,\tau)\right] = \mathbb{E}\left[\Phi(t-\tau)BB^T\Phi^T(t-\tau)\right]$, since both terms solve equation (17) as can be seen employing (13).*

Now, we give the solution representation of the system (9) via using the stochastic variation of constants method. For the Wiener case that result is stated in Theorem 1.4.1 in Damm [9].

Proposition 3.4. $(\Phi(t)z(t))_{t\geq 0}$ *is a solution of equation (9), where z is given by*

$$dz(t) = \Phi^{-1}(t)Bu(t)dt, \quad z(0) = x_0.$$

Proof. We want to determine the stochastic differential of $\Phi(t)z(t)$, $t \geq 0$, where its ith component is given by $e_i^T\Phi(t)z(t)$. Applying the Ito product formula from Corollary 2.4 yields

$$e_i^T\Phi(t)z(t) = e_i^Tx_0 + \int_0^t e_i^T\Phi(s-)d(z(s)) + \int_0^t z^T(s)d(\Phi^T(s)e_i).$$

Above, the quadratic covariation terms are zero, since z is a continuous semimartingale with a martingale part of zero (see equation (2)). Applying that $s \mapsto \Phi(\omega, s)$ and $s \mapsto \Phi(\omega, s-)$ coincide ds-almost everywhere for $\mathbb{P}$-almost all fixed $\omega \in \Omega$, we have

$$\begin{aligned}
e_i^T\Phi(t)z(t) &= e_i^Tx_0 + \int_0^t e_i^T\Phi(s)\Phi^{-1}(s)Bu(s)ds + \int_0^t z^T(s)\Phi^T(s)A^Te_ids \\
&\quad + \sum_{k=1}^q \int_0^t z^T(s)\Phi^T(s-)(\Psi^k)^Te_idM_k(s) \\
&= e_i^Tx_0 + e_i^T\int_0^t Bu(s)ds + e_i^T\int_0^t A\Phi(s)z(s)ds + e_i^T\sum_{k=1}^q\int_0^t \Psi^k\Phi(s-)z(s)dM_k(s).
\end{aligned}$$

This yields

$$\Phi(t)z(t) = x_0 + \int_0^t A\Phi(s)z(s)ds + \sum_{k=1}^q\int_0^t \Psi^k\Phi(s-)z(s)dM_k(s) + \int_0^t Bu(s)ds.$$

□

Below, we set $P_t := \int_0^t \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right]ds$ and call P_t finite reachability Gramian at time $t \geq 0$. By $X(T, 0, u)$ we denote the solution of the inhomogeneous system (9) at time T with initial condition 0 for a given input u. With Proposition 3.4 we already know that

$$X(T,0,u) = \int_0^T \Phi(T)\Phi^{-1}(t)Bu(t)dt = \int_0^T \Phi(T,t)Bu(t)dt.$$

Now, we have the goal to steer the average state of the system (9) from zero to any given $x \in \mathbb{R}^n$ via the control u with minimal energy. First of all we need the following definition, which is motivated by the remarks above Theorem 2.3 in [4].

Definition 3.5. *An average state $x \in \mathbb{R}^n$ is called reachable (from zero) if there is a time $T > 0$ and a control function $u \in L^2_T$, such that we have*

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

We say that the stochastic system is completely reachable if every average vector $x \in \mathbb{R}^n$ is reachable. Next, we characterize the set of all reachable average states. First of all, we need the following proposition, where we define $P := \int_0^\infty \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] ds$.

Proposition 3.6. *The finite reachability Gramians P_t, $t > 0$, have the same image as the infinite reachability Gramian P, i.e.,*

$$\operatorname{im} P_t = \operatorname{im} P$$

for all $t > 0$.

Proof. Since P and P_t are positive semidefinite and symmetric by definition it is sufficient to show that the kernels are equal. First, we assume $v \in \ker P$. Thus,

$$0 \leq v^T P_t v \leq v^T P v = 0$$

since $t \mapsto v^T P_t v$ is increasing such that $v \in \ker P_t$ follows. On the other hand, if $v \in \ker P_t$ we have

$$0 = v^T P_t v = \int_0^t v^T \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] v ds.$$

Hence, we can conclude that $v^T\mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] v = 0$ for almost all $s \in [0, t]$. Addionally, we know that $t \mapsto \mathbb{E}\left[\Phi(t)BB^T\Phi^T(t)\right]$ is the solution of the linear matrix differential equation

$$\dot{\mathbb{Y}}(t) = A\mathbb{Y}(t) + \mathbb{Y}(t)A^T + \sum_{k=1}^{q} \Psi^k \mathbb{Y}(t)(\Psi^k)^T\, \mathbb{E}\left[M_k(1)^2\right]$$

with initial condition $\mathbb{Y}(0) = BB^T$ for $t \geq 0$. The vectorized form $\operatorname{vec}(\mathbb{Y})$ satisfies

$$\operatorname{vec}(\dot{\mathbb{Y}}(t)) = \left(I_n \otimes A + A \otimes I_n + \sum_{k=1}^{q} \Psi^k \otimes \Psi^k \cdot \mathbb{E}\left[M_k(1)^2\right]\right) \operatorname{vec}(\mathbb{Y}(t)), \quad \operatorname{vec}(\mathbb{Y}(0)) = \operatorname{vec}(BB^T).$$

Thus, the entries of $\mathbb{E}\left[\Phi(t)BB^T\Phi^T(t)\right]$ are analytic functions. This implies that the function $f(t) := v^T\mathbb{E}\left[\Phi(t)BB^T\Phi^T(t)\right] v$ is analytic, such that $f \equiv 0$ on $[0, \infty)$. Thus,

$$0 = \int_0^\infty v^T\mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] v ds = v^T P v.$$

□

Proposition 3.7. *An average state $x \in \mathbb{R}^n$ is reachable (from zero) if and only if $x \in \operatorname{im} P$, where $P := \int_0^\infty \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] ds$.*

Proof. Provided $x \in \operatorname{im} P$ we will show that this average state can be reached with the following input function:

$$[0, T] \ni t \mapsto u(t, \omega) = B^T\Phi^T(T, t, \omega)P_T^{\#}x \tag{18}$$

for all $\omega \in \Omega$ where $P_T^{\#}$ denotes the Moore-Penrose pseudoinverse of P_T. Thus,

$$\mathbb{E}\left[X(T, 0, u)\right] = \mathbb{E}\left[\int_0^T \Phi(T, t)BB^T\Phi^T(T, t)P_T^{\#}xdt\right]$$

by inserting the function u. Considering the remarks above Proposition 3.4 we have

$$\mathbb{E}\left[\Phi(t-\tau)BB^T\Phi^T(t-\tau)\right] = \mathbb{E}\left[\Phi(t,\tau)BB^T\Phi^T(t,\tau)\right].$$

Using this fact we obtain

$$\mathbb{E}\left[X(T,0,u)\right] = \mathbb{E}\left[\int_0^T \Phi(T-t)BB^T\Phi^T(T-t)P_T^{\#}xdt\right].$$

We substitute $s = T - t$ and since $x \in \operatorname{im} P_T$ by Proposition 3.6 we get

$$\mathbb{E}\left[X(T,0,u)\right] = \mathbb{E}\left[\int_0^T \Phi(s)BB^T\Phi^T(s)ds\right] P_T^{\#}x = P_T P_T^{\#}x = x.$$

The energy of the input function $u(t) = B^T\Phi^T(T,t)P_T^{\#}x$ is

$$\begin{aligned}\|u\|_{L_T^2}^2 &= \mathbb{E}\left[\int_0^T x^T P_T^{\#}\Phi(T-t)BB^T\Phi^T(T-t)P_T^{\#}xdt\right] \\ &= x^T P_T^{\#}\mathbb{E}\left[\int_0^T \Phi(T-t)BB^T\Phi^T(T-t)dt\right] P_T^{\#}x = x^T P_T^{\#}P_T P_T^{\#}x = x^T P_T^{\#}x.\end{aligned}$$

On the other hand, if $x \in \mathbb{R}^n$ is reachable, then there exists an input function u and a time $t > 0$ such that

$$x = \mathbb{E}\left[X(t,0,u)\right] = \mathbb{E}\left[\int_0^t \Phi(t,s)Bu(s)ds\right]$$

by definition. We assume that $v \in \ker P$. Hence,

$$\left|\langle x,v\rangle_2\right| = \left|\mathbb{E}\left[\int_0^t \langle\Phi(t,s)Bu(s),v\rangle_2\, ds\right]\right| = \left|\mathbb{E}\left[\int_0^t \left\langle u(s), B^T\Phi^T(t,s)v\right\rangle_2 ds\right]\right|.$$

Employing the Cauchy-Schwarz inequality, we obtain

$$\left|\langle x,v\rangle_2\right| \leq \mathbb{E}\left[\int_0^t \|u(s)\|_2 \left\|B^T\Phi^T(t,s)v\right\|_2 ds\right].$$

By the Hölder inequality, we have

$$\begin{aligned}\left|\langle x,v\rangle_2\right| &\leq \|u\|_{L_t^2}\left(\mathbb{E}\left[\int_0^t \left\|B^T\Phi^T(t,s)v\right\|_2^2 ds\right]\right)^{\frac{1}{2}} \\ &= \|u\|_{L_t^2}\left(v^T\mathbb{E}\left[\int_0^t \Phi(t,s)BB^T\Phi^T(t,s)ds\right]v\right)^{\frac{1}{2}} = \|u\|_{L_t^2}\left(v^T P_t v\right)^{\frac{1}{2}}.\end{aligned}$$

Since $t \mapsto v^T P_t v$ is increasing, we obtain

$$\left|\langle x,v\rangle_2\right| \leq \|u\|_{L_t^2}\left(v^T P v\right)^{\frac{1}{2}} = 0.$$

Thus, $\langle x,v\rangle_2 = 0$, such that we can conclude that $x \in \operatorname{im} P$ due to $\operatorname{im} P = (\ker P)^{\perp}$. $\square$

Proposition 3.8. *The input function given by (18) is the one with the minimal energy to reach $x \in \mathbb{R}^n$ at any time $T > 0$. This minimal energy is given by $x^T P_T^{\#}x$.*

Proof. Let $\tilde{u}(t)$, $t \in [0,T]$, be an additional function we can reach with the average state x at time T, then

$$x = \mathbb{E}\left[\int_0^T \Phi(T,t)B\left(u(t) + (\tilde{u}(t) - u(t))\right)dt\right]$$

holds per definition of $\tilde{u}$. Thus, the residual vanishes

$$\mathbb{E}\left[\int_0^T \Phi(T,t)B\left(\tilde{u}(t) - u(t)\right)dt\right] = 0$$

such that

$$\mathbb{E}\left[\int_0^T u(t)^T\left(\tilde{u}(t) - u(t)\right)dt\right] = 0$$

follows. Hence, we have

$$\|\tilde{u}\|_{L_T^2}^2 = \|u + (\tilde{u} - u)\|_{L_T^2}^2 = \|u\|_{L_T^2}^2 + \|\tilde{u} - u\|_{L_T^2}^2 \geq \|u\|_{L_T^2}^2 .$$

From the proof of Proposition 3.7 we know that the energy of u is given by $x^T P_T^{\#} x$. □

So, by Proposition 3.8, the minimal energy that is needed to steer the system to x is given by $\inf_{T>0} x^T P_T^{\#} x$. By definition of P_T we know that it is increasing in time such that the pseudoinverse $P_T^{\#}$ is decreasing. Hence, it is clear that the minimal energy is given by $x^T P^{\#} x$, where $P^{\#}$ is the pseudoinverse of $P := \int_0^\infty \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right]ds$. The same result is obtained by Benner and Damm [4] in Theorem 2.3 for stochastic differential equations driven by Wiener processes. This is also true for the deterministic case (see Section 4.3.1 in Antoulas [1]).

3.2 Observability concept

Below, we introduce the concept of observability for the output equation

$$\mathcal{Y}(t) = CX(t) \tag{19}$$

corresponding to the stochastic linear system (9), where $C \in \mathbb{R}^{p\times n}$. Therefore, we need the following Proposition.

Proposition 3.9. *Let $\hat{Q}$ be a symmetric positive semidefinite matrix and $Y_a := X(\cdot, a, 0)$, $Y_b := X(\cdot, b, 0)$ the homogeneous solutions to (9) with initial conditions $a, b \in \mathbb{R}^n$, then*

$$\begin{aligned}\mathbb{E}\left[Y_a(t)^T\hat{Q}Y_b(t)\right] = a^T\hat{Q}b + \mathbb{E}\left[\int_0^t Y_a^T(s)\hat{Q}AY_b(s)ds\right] + \mathbb{E}\left[\int_0^t Y_a^T(s)A^T\hat{Q}Y_b(s)ds\right]\\ + \mathbb{E}\left[\int_0^t Y_a(s)^T\sum_{k=1}^q(\Psi^k)^T\hat{Q}\Psi^k\mathbb{E}\left[M_k(1)^2\right]Y_b(s)ds\right].\end{aligned} \tag{20}$$

Proof. By applying the Ito product formula from Corollary 2.4, we have

$$Y_a^T(t)\hat{Q}Y_b(t) = a^T\hat{Q}b + \int_0^t Y_a^T(s-)d(\hat{Q}Y_b(s)) + \int_0^t Y_b^T(s-)\hat{Q}d(Y_a(s)) + \sum_{i=1}^n[e_i^TY_a(t), e_i^T\hat{Q}Y_b(t)]_t,$$

where e_i is the ith unit vector ($i = 1, \ldots, n$). We get

$$\int_0^t Y_a^T(s-)d(\hat{Q}Y_b(s)) = \int_0^t Y_a^T(s-)\hat{Q}AY_b(s)ds + \sum_{k=1}^q \int_0^t Y_a^T(s-)\hat{Q}\Psi^k Y_b(s-)dM_k(s)$$

and

$$\int_0^t Y_b^T(s-)\hat{Q}d(Y_a(s)) = \int_0^t Y_b(s-)^T\hat{Q}AY_a(s)ds + \sum_{k=1}^q \int_0^t Y_b(s-)^T\hat{Q}\Psi^k Y_a(s-)dM_k(s).$$

By equation (8) the mean of the quadratic covariations is given by

$$\mathbb{E}[e_i^T Y_a(t), e_i^T \hat{Q} Y_b(t)]_t = \sum_{k=1}^q \mathbb{E}\int_0^t e_i^T \Psi^k Y_a(s) e_i^T \hat{Q}\Psi^k Y_b(s)ds\, \mathbb{E}\left[M_k(1)^2\right].$$

With Theorem 2.11 (i) we obtain

$$\begin{aligned}\mathbb{E}\left[Y_a(t)^T\hat{Q}Y_b(t)\right] &= a^T\hat{Q}b + \mathbb{E}\left[\int_0^t Y_a^T(s)\hat{Q}AY_b(s)ds\right] + \mathbb{E}\left[\int_0^t Y_a^T(s)A^T\hat{Q}Y_b(s)ds\right]\\ &+ \sum_{k=1}^q \mathbb{E}\left[\int_0^t Y_a(s)^T(\Psi^k)^T\hat{Q}\Psi^k Y_b(s)ds\right]\mathbb{E}\left[M_k(1)^2\right]\end{aligned}$$

using that the trajectories of Y_a and Y_b only have jumps on Lebesgue zero sets. $\square$

If we set $a = e_i$ and $b = e_j$ in Proposition 3.9, we obtain

$$\begin{aligned}\mathbb{E}\left[e_i^T\Phi(t)^T\hat{Q}\Phi(t)e_j\right] &= e_i^T\hat{Q}e_j + \mathbb{E}\left[\int_0^t e_i^T\Phi^T(s)\hat{Q}A\Phi(s)e_jds\right] + \mathbb{E}\left[\int_0^t e_i^T\Phi^T(s)A^T\hat{Q}\Phi(s)e_jds\right]\\ &+ \mathbb{E}\left[\int_0^t e_i^T\Phi(s)^T\sum_{k=1}^q(\Psi^k)^T\hat{Q}\Psi^k\mathbb{E}\left[M_k(1)^2\right]\Phi(s)e_jds\right].\end{aligned}$$

This yields

$$\begin{aligned}\mathbb{E}\left[\Phi(t)^T\hat{Q}\Phi(t)\right] &= \hat{Q} + \mathbb{E}\left[\int_0^t \Phi^T(s)\hat{Q}A\Phi(s)ds\right] + \mathbb{E}\left[\int_0^t \Phi^T(s)A^T\hat{Q}\Phi(s)ds\right]\\ &+ \mathbb{E}\left[\int_0^t \Phi(s)^T\sum_{k=1}^q(\Psi^k)^T\hat{Q}\Psi^k\mathbb{E}\left[M_k(1)^2\right]\Phi(s)ds\right].\end{aligned}$$

Let Q be the solution of the generalized Lyapunov equation

$$A^TQ + QA + \sum_{k=1}^q(\Psi^k)^TQ\Psi^k\,\mathbb{E}\left[M_k(1)^2\right] = -C^TC. \tag{21}$$

Then,

$$\mathbb{E}\left[\Phi(t)^TQ\Phi(t)\right] = Q - \mathbb{E}\left[\int_0^t \Phi^T(s)C^TC\Phi(s)ds\right]$$

and by taking the limit $t \to \infty$, we have

$$Q = \mathbb{E}\left[\int_0^\infty \Phi^T(s)C^TC\Phi(s)ds\right] \tag{22}$$

due to the asymptotic mean square stability of the homogenous equation, which provides the existence of the integral in equation (22) as well.

Remark. *The matrix equation (21) is uniquely solvable, since*

$$L := \left(A^T \otimes I_n + I_n \otimes A^T + \sum_{k=1}^{q} (\Psi^k)^T \otimes (\Psi^k)^T \cdot \mathbb{E}\left[M_k(1)^2\right] \right)$$

has non zero eigenvalues and hence the solution of $L \cdot \operatorname{vec}(Q) = -\operatorname{vec}(C^T C)$ *is unique.*

Next, we assume that the system (9) is uncontrolled, that means $u \equiv 0$. By using our knowledge concerning the homogeneous system, $X(t, x_0, 0)$ is given by $\Phi(t)x_0$, where here $x_0 \in \mathbb{R}^n$ denotes the initial value of the system. So, we obtain $\mathcal{Y}(t) = C\Phi(t)x_0$.
We observe $\mathcal{Y}$ on a time interval $[0, \infty)$. The problem is to find x_0 from the observations we have. The energy produced by the initial value x_0 is

$$\|\mathcal{Y}\|_{L^2}^2 := \mathbb{E} \int_0^\infty \mathcal{Y}^T(t)\mathcal{Y}(t)dt = x_0^T \mathbb{E} \int_0^\infty \Phi^T(t)C^T C\Phi(t)dt\, x_0 = x_0^T Q x_0, \tag{23}$$

where we set $Q := \mathbb{E} \int_0^\infty \Phi^T(s)C^T C\Phi(s)ds$. As in Benner, Damm [4], Q takes the part of the observability Gramian of the stochastic system with output equation (19). We call a state x_0 unobservable if it is in the kernel of Q. Otherwise it is said to be observable. We say that a system is completely observable if the kernel of Q is trivial.

4 Balanced truncation for stochastic systems

For obtaining a reduced order model for a deterministic LTI system, balanced truncation is a method of major importance. For the procedure of balanced truncation in the deterministic case, see Antoulas [1], Benner et al. [5] and Obinata, Anderson [19]. In this section we want to generalize this method for stochastic linear systems, which are influenced by Levy noise.

4.1 Procedure

We assume A, $\Psi^k \in \mathbb{R}^{n\times n}$ $(k = 1, \ldots, q)$, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times n}$ and consider the following stochastic system:

$$\begin{aligned} dX(t) &= [AX(t) + Bu(t)]dt + \sum_{k=1}^{q} \Psi^k X(t-)dM_k(t), \quad t \geq 0,\ X(0) = x_0, \\ \mathcal{Y}(t) &= CX(t), \end{aligned} \tag{24}$$

where the noise processes M_k $(k = 1, \ldots, q)$ are uncorrelated real-valued and square integrable Levy processes with mean zero. We assume the homogenous solution Y_{y_0}, which fulfills

$$dY(t) = AY(t)dt + \sum_{k=1}^{q} \Psi^k Y(t-)dM_k(t), \quad t \geq 0,\ Y(0) = y_0,$$

to be mean square asymptotically stable, which means that $\mathbb{E}\left\|Y_{y_0}(t)\right\|_2^2 \to 0$ for $t \to \infty$ and arbitrary initial condition $y_0 \in \mathbb{R}^n$. In addition, we require that the system (24) is completely reachable and observable, which is equivalent to P and Q being positive definite.

Let $T \in \mathbb{R}^{n\times n}$ be a regular matrix. If we transform the states in the following way

$$\hat{X}(t) = TX(t),$$

we obtain the following system:

$$d\hat{X}(t) = [\tilde{A}\hat{X}(t) + \tilde{B}u(t)]dt + \sum_{k=1}^{q} \tilde{\Psi}^k \hat{X}(t-)dM_k(t), \quad \hat{X}(0) = Tx_0, \tag{25}$$
$$\mathcal{Y}(t) = \tilde{C}\hat{X}(t), \quad t \geq 0,$$

where $\tilde{A} = TAT^{-1}$, $\tilde{\Psi}^k = T\Psi^k T^{-1}$, $\tilde{B} = TB$ and $\tilde{C} = CT^{-1}$. For an arbitrary fixed input, the transformated system (25) has always the same output as the system (24). The reachability Gramian $P := \int_0^\infty \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right] ds$ of system (24) fulfills

$$-BB^T = AP + PA^T + \sum_{k=1}^{q} \Psi^k P(\Psi^k)^T \cdot c_k$$

where $c_k = \mathbb{E}\left[M_k(1)^2\right]$. By multiplying T from the left and T^T from the right hand side we obtain

$$\begin{aligned} -\tilde{B}\tilde{B}^T &= TAPT^T + TPA^TT^T + \sum_{k=1}^{q} T\Psi^k P(\Psi^k)^T T^T \cdot c_k \\ &= \tilde{A}TPT^T + TPT^T\tilde{A}^T + \sum_{k=1}^{q} \tilde{\Psi}^k TPT^T (\tilde{\Psi}^k)^T \cdot c_k. \end{aligned}$$

Hence, the reachability Gramian of the transformed system (25) is given by $\tilde{P} = TPT^T$. For the observability Gramian of the transformed system it holds $\tilde{Q} = T^{-T}QT^{-1}$, where $Q := \int_0^\infty \mathbb{E}\left[\Phi^T(s)C^TC\Phi(s)\right] ds$ is the observability Gramian of the original system. Hence,

$$-\tilde{C}^T\tilde{C} = \tilde{A}^T\tilde{Q} + \tilde{Q}\tilde{A} + \sum_{k=1}^{q} (\tilde{\Psi}^k)^T \tilde{Q}\tilde{\Psi}^k \cdot c_k.$$

In addition, it is easy to verify that the generalized Hankel singular values $\sigma_1 \geq \ldots \geq \sigma_n > 0$ of (24), which are the square roots of the eigenvalues of PQ, are equal to those of (25).

Like in the deterministic case (see [1] and [19]) we choose T such that $\tilde{Q}$ and $\tilde{P}$ are equal and diagonal. A system with equal and diagonal Gramians we call balanced system. The corresponding balancing T is given by

$$T = \Sigma^{\frac{1}{2}} K^T U^{-1} \text{ and } T^{-1} = UK\Sigma^{-\frac{1}{2}}, \tag{26}$$

where $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$, U comes from the Cholesky decompostion of $P = UU^T$ and K is an orthogonal matrix corresponding to the EVD (SVD respectively) of $U^TQU = K\Sigma^2K^T$. So, we obtain

$$\tilde{Q} = \tilde{P} = \Sigma.$$

Our aim is to truncate the average states that are difficult to observe and difficult to reach, which are those producing least observation energy and causing the most energy to reach, respectively. By equation (23) we can say that the states which are difficult to observe are contained in the space spanned by the eigenvectors corresponding to the small eigenvalues of Q. Analogously, the average states that are difficult to reach are contained in the space spanned by the eigenvectors corresponding to the small eigenvalues of P (or to the large eigenvalues of P^{-1}, respectively) considering the remarks below Proposition 3.8. In a balanced system the dominant reachable and observable states are the same.

We consider the following partitions:

$$T = \begin{bmatrix} W^T \\ T_2^T \end{bmatrix}, \; T^{-1} = \begin{bmatrix} V & T_1 \end{bmatrix} \text{ and } \hat{X} = \begin{pmatrix} \tilde{X} \\ X_1 \end{pmatrix},$$

where $W^T \in \mathbb{R}^{r\times n}, V \in \mathbb{R}^{n\times r}$ and $\tilde{X}$ takes values in $\mathbb{R}^r$ $(r < n)$. Hence, we have

$$\begin{aligned} \begin{pmatrix} d\tilde{X}(t) \\ dX_1(t) \end{pmatrix} &= \left(\begin{bmatrix} W^TAV & W^TAT_1 \\ T_2^TAV & T_2^TAT_1 \end{bmatrix} \begin{pmatrix} \tilde{X}(t) \\ X_1(t) \end{pmatrix} + \begin{bmatrix} W^TB \\ T_2^TB \end{bmatrix} u(t) \right) dt \\ &+ \sum_{k=1}^{q} \begin{bmatrix} W^T\Psi^kV & W^T\Psi^kT_1 \\ T_2^T\Psi^kV & T_2^T\Psi^kT_1 \end{bmatrix} \begin{pmatrix} \tilde{X}(t-) \\ X_1(t-) \end{pmatrix} dM_k(t) \end{aligned} \tag{27}$$

and

$$\mathcal{Y}(t) = \begin{bmatrix} CV & CT_1 \end{bmatrix} \begin{pmatrix} \tilde{X}(t) \\ X_1(t) \end{pmatrix}.$$

By truncating the system and neglecting the X_1 terms, the approximating reduced order model is given by

$$\begin{aligned} d\tilde{X}(t) &= [W^TAV\tilde{X}(t) + W^TBu(t)]dt + \sum_{k=1}^{q} W^T\Psi^kV\tilde{X}(t-)dM_k(t), \\ \hat{\mathcal{Y}}(t) &= CV\tilde{X}(t). \end{aligned} \tag{28}$$

What we will show now is that the homogenous solution $\tilde{Y}_{y_0}$ of the reduced system (28), which fulfills

$$dY(t) = W^TAVY(t)dt + \sum_{k=1}^{q} W^T\Psi^kVY(t-)dM_k(t), \quad Y(0) = y_0,$$

is mean square stable in general and still asymptotically mean square stable under certain conditions.

Theorem 4.1. *Let $\tilde{Y}_{y_0}$ be the homogenous solution of the system (28) with initial condition $y_0 \in \mathbb{R}^r$. Then,*

$$\mathbb{E}\left\|\tilde{Y}_{y_0}(t)\right\|_2^2 \leq \frac{\sigma_1}{\sigma_r}\left\|y_0\right\|_2^2, \quad t \geq 0. \tag{29}$$

If furthermore

$$\ker CV \cap \ker T_2^T\Psi^1V \cap \ldots \cap \ker T_2^T\Psi^qV = \{0\} \tag{30}$$

holds, then there exists a $\lambda_r < 0$ such that we have

$$\mathbb{E}\left\|\tilde{Y}_{y_0}(t)\right\|_2^2 \leq \frac{\sigma_1}{\sigma_r}\left\|y_0\right\|_2^2 \mathrm{e}^{\frac{\lambda_r}{\sigma_1}t}, \quad t \geq 0.$$

Proof. In equation (27) we block-wise set

$$\begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix} := \begin{bmatrix} W^TAV & W^TAT_1 \\ T_2^TAV & T_2^TAT_1 \end{bmatrix} \text{ and } \begin{bmatrix} \tilde{\Psi}_{11}^k & \tilde{\Psi}_{12}^k \\ \tilde{\Psi}_{21}^k & \tilde{\Psi}_{22}^k \end{bmatrix} := \begin{bmatrix} W^T\Psi^kV & W^T\Psi^kT_1 \\ T_2^T\Psi^kV & T_2^T\Psi^kT_1 \end{bmatrix}.$$

In the corresponding output equation we block-wise define

$$\begin{bmatrix}\tilde{C}_1 & \tilde{C}_2\end{bmatrix} := \begin{bmatrix}CV & CT_1\end{bmatrix}.$$

We know

$$\begin{bmatrix}\tilde{A}_{11}^T & \tilde{A}_{21}^T \\ \tilde{A}_{12}^T & \tilde{A}_{22}^T\end{bmatrix}\begin{bmatrix}\Sigma_1 & \\ & \Sigma_2\end{bmatrix} + \begin{bmatrix}\Sigma_1 & \\ & \Sigma_2\end{bmatrix}\begin{bmatrix}\tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22}\end{bmatrix} + \sum_{k=1}^{q}\begin{bmatrix}(\tilde{\Psi}_{11}^k)^T & (\tilde{\Psi}_{21}^k)^T \\ (\tilde{\Psi}_{12}^k)^T & (\tilde{\Psi}_{22}^k)^T\end{bmatrix}\begin{bmatrix}\Sigma_1 & \\ & \Sigma_2\end{bmatrix}\begin{bmatrix}\tilde{\Psi}_{11}^k & \tilde{\Psi}_{12}^k \\ \tilde{\Psi}_{21}^k & \tilde{\Psi}_{22}^k\end{bmatrix}\cdot c_k$$
$$= -\begin{bmatrix}\tilde{C}_1^T\tilde{C}_1 & \tilde{C}_1^T\tilde{C}_2 \\ \tilde{C}_2^T\tilde{C}_1 & \tilde{C}_2^T\tilde{C}_2\end{bmatrix},$$

where $\Sigma_1 = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$, $\Sigma_2 = \operatorname{diag}(\sigma_{r+1}, \ldots, \sigma_n)$ and $c_k = \mathbb{E}[M_k(1)^2]$. Considering the left upper block we obtain

$$\tilde{A}_{11}^T\Sigma_1 + \Sigma_1\tilde{A}_{11} + \sum_{k=1}^{q}(\tilde{\Psi}_{11}^k)^T\Sigma_1\tilde{\Psi}_{11}^k\cdot c_k = -\left(\sum_{k=1}^{q}(\tilde{\Psi}_{21}^k)^T\Sigma_2\tilde{\Psi}_{21}^k\cdot c_k + \tilde{C}_1^T\tilde{C}_1\right) =: L.$$

Below, we operate with similar arguments like in the proof of Theorem 1.5.3 ($(iv) \Rightarrow (ii)$) in Damm [9]. From equation (20) we can conclude that

$$\mathbb{E}\left[\tilde{Y}_{y_0}(t)^T\Sigma_1\tilde{Y}_{y_0}(t)\right] = y_0^T\Sigma_1 y_0 + \mathbb{E}\left[\int_0^t \tilde{Y}_{y_0}^T(s)\Sigma_1\tilde{A}_{11}\tilde{Y}_{y_0}(s)ds\right] + \mathbb{E}\left[\int_0^t \tilde{Y}_{y_0}^T(s)\tilde{A}_{11}^T\Sigma_1\tilde{Y}_{y_0}(s)ds\right]$$
$$+\,\mathbb{E}\left[\int_0^t \tilde{Y}_{y_0}(s)^T\sum_{k=1}^{q}(\tilde{\Psi}_{11}^k)^T\Sigma_1\tilde{\Psi}_{11}^k c_k\tilde{Y}_{y_0}(s)ds\right].$$

Thus,

$$g(t) := \mathbb{E}\left[\tilde{Y}_{y_0}(t)^T\Sigma_1\tilde{Y}_{y_0}(t)\right] = y_0^T\Sigma_1 y_0 + \mathbb{E}\left[\int_0^t \tilde{Y}_{y_0}^T(s)L\tilde{Y}_{y_0}(s)ds\right].$$

We differentiate both sides of the equation. This yields

$$\dot{g}(t) = \mathbb{E}\left[\tilde{Y}_{y_0}^T(t)L\tilde{Y}_{y_0}(t)\right].$$

With property (30) we conclude that L is symmetric and negative definite, since from $v^TLv = 0$ it follows that $\tilde{\Psi}_{21}^1 v = 0, \ldots, \tilde{\Psi}_{21}^q v = 0$ and $\tilde{C}_1 v = 0$, which means by assumption that $v = 0$. Hence, we can choose an orthonormal basis of $\mathbb{R}^r$ consisting of eigenvectors $(l_i)_{i=1,\ldots,r}$ of L. The corresponding eigenvalues we denote by $\lambda_1 \leq \ldots \leq \lambda_r < 0$. For $v \in \mathbb{R}^r$, there are constants $a_1, \ldots, a_r$ such that $x = \sum_{j=1}^r a_j l_j$. So, we have

$$v^TLv = \sum_{j=1}^{r} a_j l_j^T \sum_{i=1}^{r} a_i\lambda_i l_i = \sum_{i=1}^{r} a_i^2\lambda_i \leq \lambda_r v^Tv.$$

Hence,

$$\dot{g}(t) \leq \lambda_r\mathbb{E}\left[\tilde{Y}_{y_0}^T(t)\tilde{Y}_{y_0}(t)\right].$$

On the other hand, we have $\sigma_r v^Tv \leq v^T\Sigma_1 v \leq \sigma_1 v^Tv$. Thus,

$$\dot{g}(t) \leq \frac{\lambda_r}{\sigma_1}\mathbb{E}\left[\tilde{Y}_{y_0}^T(t)\Sigma_1\tilde{Y}_{y_0}(t)\right] = \frac{\lambda_r}{\sigma_1}g(t).$$

Due to the continuity of the function g, we can apply the Lemma of Gronwall. This provides

$$g(t) \leq g(0)\, \mathrm{e}^{\frac{\lambda_r}{\sigma_1} t}$$

for $t \geq 0$. By inserting the function g, we obtain

$$\mathbb{E}\left[\tilde{Y}_{y_0}(t)^T \Sigma_1 \tilde{Y}_{y_0}(t)\right] \leq y_0^T \Sigma_1 y_0\, \mathrm{e}^{\frac{\lambda_r}{\sigma_1} t}.$$

Finally, we have

$$\mathbb{E}\left[\tilde{Y}_{y_0}(t)^T \tilde{Y}_{y_0}(t)\right] \leq \frac{\sigma_1}{\sigma_r} y_0^T y_0\, \mathrm{e}^{\frac{\lambda_r}{\sigma_1} t}, \quad t \geq 0. \tag{31}$$

In general, L is negative semidefinite, such that $\lambda_r = 0$ in inequality (31). So, the mean square stability (29) follows. □

Remark. *One persisting problem is to find an explicit structure of the Gramians of the reduced model. As we will see in an example below, the reduced order model is not balanced, that means the Gramians are neither diagonal nor equal. In addition, the Hankel singular values are different from those of the original system.*
Also, it is not clear if the preservation of asymptotic mean square stability in is given in general. In fact, it remains to show that

$$I_r \otimes A_{11} + A_{11} \otimes I_r + \sum_{k=1}^{q} \Psi_{11}^k \otimes \Psi_{11}^k \cdot \mathbb{E}\left[M_k(1)^2\right]$$

is invertible, where $A_{11} := W^T A V$ and $\Psi_{11}^k := W^T \Psi^k V$.

Example 4.2. *We consider the case, where $q = 1$ and the noise process is a Wiener process W. So, the system we focus on is*

$$\begin{aligned} dX(t) &= [AX(t) + Bu(t)]dt + \Psi X(t) dW(t) \\ \mathcal{Y}(t) &= CX(t). \end{aligned} \tag{32}$$

The following matrices provide a balanced and asymptotically mean square stable system:

$$A = \begin{pmatrix} -4.4353 & 3.9992 & -0.3287 \\ 2.9337 & -11.0285 & -0.4319 \\ -0.0591 & -0.1303 & -11.5362 \end{pmatrix}, B = \begin{pmatrix} -3.4648 & -1.9391 & -3.6790 \\ 5.7925 & 4.1379 & 2.3036 \\ -0.3258 & 1.1359 & 2.8972 \end{pmatrix},$$
$$\Psi = \begin{pmatrix} -1.4886 & 2.8510 & -0.2429 \\ 0.4720 & 0.5803 & 3.1152 \\ -1.6123 & -0.8082 & -0.0917 \end{pmatrix}, C = \begin{pmatrix} -3.0588 & 0.4275 & 0.2630 \\ -4.8686 & 1.2886 & 1.0769 \\ -4.3349 & 0.6747 & -0.1734 \end{pmatrix}.$$

The Gramians are given by

$$P = Q = \Sigma = \begin{pmatrix} 8.4788 & 0 & 0 \\ 0 & 3.3232 & 0 \\ 0 & 0 & 1.4726 \end{pmatrix}.$$

The reduced order model ($r = 2$) is asymptotically mean square stable and has the following Gramians:

$$P_R = \begin{pmatrix} 7.7470 & -0.3562 \\ -0.3562 & 2.5496 \end{pmatrix} \text{ and } Q_R = \begin{pmatrix} 7.7495 & -0.2074 \\ -0.2074 & 2.8980 \end{pmatrix}.$$

The Hankel singular values of the reduced order model are 7.6633 *and* 2.7001.

At the end of this section we provide a short example that shows that the reduced order model need not be completely observable and reachable even if the original system is completely observable and reachable:

Example 4.3. *We consider the equations (32) with the matrices*

$$(A, B, \Psi, C) = \left(\begin{pmatrix} -0.25 & 1 \\ 1 & -9 \end{pmatrix}, \begin{pmatrix} 0 \\ \sqrt{7} \end{pmatrix}, \Psi = \begin{pmatrix} 0 & 1 \\ 1 & -3 \end{pmatrix}, C = \begin{pmatrix} 0 & \sqrt{7} \end{pmatrix} \right)$$

and obtain a balanced and asymptotically mean square stable system being completely reachable and observable. The Hankel singular values are 2 and 1. Truncating yields a system with coefficients $(A_{11}, B_1, \Psi_{11}, C_1) = (-0.25, 0, 0, 0)$ *having Gramians* $P_R = Q_R = 0$.

4.2 Error bound for balanced truncation

Let $\left(A, \Psi^k, B, C\right)$ $(k = 1, \ldots, q)$ be a realization of system (24). Furthermore, we assume the inital condition of the system to be zero. We introduce the following partitions:

$$TAT^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \; T\Psi^k T^{-1} = \begin{bmatrix} \Psi_{11}^k & \Psi_{12}^k \\ \Psi_{21}^k & \Psi_{22}^k \end{bmatrix}, \; TB = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \text{ and } CT^{-1} = \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \tag{33}$$

where T is the balancing transformation defined in (26) and $\left(A_{11}, \Psi_{11}^k, B_1, C_1\right)$ are the coefficients of the reduced order model. The output of the reduced (truncated) system is given by

$$\hat{\mathcal{Y}}(t) = C_1 \tilde{X}(t) = C_1 \int_0^t \tilde{\Phi}(t, s) B_1 u(s) ds,$$

where $\tilde{\Phi}$ is the fundamental matrix of the truncated system. In addition, we assume that the homogenous equation of the reduced system is still asymptotically mean square stable. The asymptotic mean square stability of the reduced model we have for example if $\ker C_1 \cap \ker \Psi_{21}^1 \cap \ldots \cap \ker \Psi_{21}^q = \{0\}$ as shown in Theorem 4.1. In addition, we know

$$\mathcal{Y}(t) = CX(t) = C \int_0^t \Phi(t, s) B u(s) ds.$$

It is our goal to steer the average state via the control u and to truncate the average states that are difficult to reach for obtaining a reduced order model. Therefore, it is a meaningful criterion to consider the worst case mean error of $\hat{\mathcal{Y}}(t)$ and $\mathcal{Y}(t)$. Below, we give a bound for that kind of error:

$$\begin{aligned} \mathbb{E} \left\| \hat{\mathcal{Y}}(t) - \mathcal{Y}(t) \right\|_2 &= \mathbb{E} \left\| C \int_0^t \Phi(t, s) B u(s) ds - C_1 \int_0^t \tilde{\Phi}(t, s) B_1 u(s) ds \right\|_2 \\ &\leq \mathbb{E} \int_0^t \left\| \left(C\Phi(t, s) B - C_1 \tilde{\Phi}(t, s) B_1 \right) u(s) \right\|_2 ds \\ &\leq \mathbb{E} \int_0^t \left\| C\Phi(t, s) B - C_1 \tilde{\Phi}(t, s) B_1 \right\|_F \left\| u(s) \right\|_2 ds \end{aligned}$$

and by Hölder inequality it holds

$$\mathbb{E} \left\| \hat{\mathcal{Y}}(t) - \mathcal{Y}(t) \right\|_2 \leq \left(\mathbb{E} \int_0^t \left\| C\Phi(t, s) B - C_1 \tilde{\Phi}(t, s) B_1 \right\|_F^2 ds \right)^{\frac{1}{2}} \left(\mathbb{E} \int_0^t \left\| u(s) \right\|_2^2 ds \right)^{\frac{1}{2}}.$$

Now,

$$
\begin{aligned}
&\mathbb{E}\int_0^t \left\|C\Phi(t,s)B - C_1\tilde{\Phi}(t,s)B_1\right\|_F^2 ds \\
&= \mathbb{E}\int_0^t \|C\Phi(t,s)B\|_F^2 + \left\|C_1\tilde{\Phi}(t,s)B_1\right\|_F^2 - 2\left\langle C\Phi(t,s)B, C_1\tilde{\Phi}(t,s)B_1\right\rangle_F ds \\
&= \mathbb{E}\int_0^t \operatorname{tr}\left(C\Phi(t,s)BB^T\Phi^T(t,s)C^T\right) ds + \mathbb{E}\int_0^t \operatorname{tr}\left(C_1\tilde{\Phi}(t,s)B_1B_1^T\tilde{\Phi}^T(t,s)C_1^T\right) ds \\
&\quad - 2\mathbb{E}\int_0^t \operatorname{tr}\left(C\Phi(t,s)BB_1^T\tilde{\Phi}^T(t,s)C_1^T\right) ds \\
&= \operatorname{tr}\left(C\int_0^t \mathbb{E}\left[\Phi(t,s)BB^T\Phi^T(t,s)\right] ds\, C^T\right) + \operatorname{tr}\left(C_1\int_0^t \mathbb{E}\left[\tilde{\Phi}(t,s)B_1B_1^T\tilde{\Phi}^T(t,s)\right] ds\, C_1^T\right) \\
&\quad - 2\ \operatorname{tr}\left(C\int_0^t \mathbb{E}\left[\Phi(t,s)BB_1^T\tilde{\Phi}^T(t,s)\right] ds\, C_1^T\right).
\end{aligned} \tag{34}
$$

Due to the remarks we give above Proposition 3.4 we have

$$
\begin{aligned}
\mathbb{E}\left[\Phi(t,s)BB^T\Phi^T(t,s)\right] &= \mathbb{E}\left[\Phi(t-s)BB^T\Phi^T(t-s)\right] \text{ and} \\
\mathbb{E}\left[\tilde{\Phi}(t,s)B_1B_1^T\tilde{\Phi}^T(t,s)\right] &= \mathbb{E}\left[\tilde{\Phi}(t-s)B_1B_1^T\tilde{\Phi}^T(t-s)\right]
\end{aligned}
$$

for $0 \leq s \leq t$. Furthermore, we want to analyze the term in (34). Therefor, we need the following Proposition:

Proposition 4.4. *The $\mathbb{R}^{n\times r}$-valued function $\mathbb{E}\left[\Phi(t)BB_1^T\tilde{\Phi}^T(t)\right]$, $t \geq 0$, is the solution of the following differential equation:*

$$
\dot{\mathbb{Y}}(t) = \mathbb{Y}(t)A_{11}^T + A\mathbb{Y}(t) + \sum_{k=1}^q \Psi^k\mathbb{Y}(t)(\Psi_{11}^k)^T\,\mathbb{E}\left[M_k(1)^2\right], \quad \mathbb{Y}(0) = BB_1^T. \tag{35}
$$

Proof. With $B = [b_1, \ldots, b_m]$ and $B_1 = \left[\tilde{b}_1, \ldots, \tilde{b}_m\right]$, we obtain

$$
\Phi(t)BB_1^T\tilde{\Phi}^T(t) = \Phi(t)b_1\tilde{b}_1^T\tilde{\Phi}^T(t) + \ldots + \Phi(t)b_m\tilde{b}_m^T\tilde{\Phi}^T(t) \tag{36}
$$

By applying the Ito product formula from Corollary 2.5, we have

$$
\Phi(t)b_l\tilde{b}_l^T\tilde{\Phi}^T(t) = b_l\tilde{b}_l^T + \int_0^t d(\Phi(s)b_l)\tilde{b}_l^T\tilde{\Phi}^T(s-) + \int_0^t \Phi(s-)b_l d(\tilde{b}_l^T\tilde{\Phi}^T(s)) + \left(\left[e_i^T\Phi b_l, e_j^T\tilde{\Phi}\tilde{b}_l\right]_t\right)_{\substack{i=1,\ldots,n\\ j=1,\ldots,r}}.
$$

From (8) we know that

$$
\mathbb{E}\left[e_i^T\Phi b_l, e_j^T\tilde{\Phi}\tilde{b}_l\right]_t = \sum_{k=1}^q \mathbb{E}\left[\int_0^t e_i^T\Psi^k\Phi(s)b_l\tilde{b}_1^T\tilde{\Phi}^T(s)(\Psi_{11}^k)^T e_j ds\right]\mathbb{E}\left[M_k(1)^2\right].
$$

With Theorem 2.11 (i) we obtain

$$
\begin{aligned}
\mathbb{E}\left[\Phi(t)b_l\tilde{b}_l^T\tilde{\Phi}^T(t)\right] = b_l\tilde{b}_l^T &+ \mathbb{E}\left[\int_0^t \Phi(s)b_l\tilde{b}_l^T\tilde{\Phi}^T(s)ds\right]\ A_{11}^T + A\mathbb{E}\left[\int_0^t \Phi(s)b_l\tilde{b}_l^T\tilde{\Phi}^T(s)ds\right] \\
&+ \sum_{k=1}^q \Psi^k\mathbb{E}\left[\int_0^t \Phi(s)b_l\tilde{b}_l^T\tilde{\Phi}^T(s)ds\right]\ (\Psi_{11}^k)^T\mathbb{E}\left[M_k(1)^2\right]
\end{aligned}
$$

using that the trajectories of Φ and $\tilde{\Phi}$ only have jumps on Lebesgue zero sets. By equation (36), we have

$$\begin{aligned}\mathbb{E}\left[\Phi(t)BB_1^T\tilde{\Phi}^T(t)\right] = BB_1^T &+ \mathbb{E}\left[\int_0^t \Phi(s)BB_1^T\tilde{\Phi}^T(s)ds\right]\ A_{11}^T \\ &+ A\mathbb{E}\left[\int_0^t \Phi(s)BB_1^T\tilde{\Phi}^T(s)ds\right] \\ &+ \sum_{k=1}^q \Psi^k\mathbb{E}\left[\int_0^t \Phi(s)BB_1^T\tilde{\Phi}^T(s)ds\right]\ (\Psi_{11}^k)^T\mathbb{E}\left[M_k(1)^2\right]\end{aligned}\tag{37}$$

which provides the result. □

By Proposition 4.4 we can conclude that the function $\mathbb{E}\left[\Phi(t-\tau)BB_1^T\tilde{\Phi}^T(t-\tau)\right]$, $t \geq \tau \geq 0$, is the solution of the equation

$$\dot{\mathbb{Y}}(t) = \mathbb{Y}(t)A_{11}^T + A\mathbb{Y}(t) + \sum_{k=1}^q \Psi^k\mathbb{Y}(t)(\Psi_{11}^k)^T\ \mathbb{E}\left[M_k(1)^2\right], \quad \mathbb{Y}(\tau) = BB_1^T,\tag{38}$$

for all $t \geq \tau \geq 0$. Analogous to Proposition 4.4 we can conclude that $\mathbb{E}\left[\Phi(t,\tau)BB_1^T\tilde{\Phi}^T(t,\tau)\right]$ is also a solution of equation (38), which yields

$$\mathbb{E}\left[\Phi(t,\tau)BB_1^T\tilde{\Phi}^T(t,\tau)\right] = \mathbb{E}\left[\Phi(t-\tau)BB_1^T\tilde{\Phi}^T(t-\tau)\right]\tag{39}$$

for all $t \geq \tau \geq 0$. Using equation (39) we have

$$\begin{aligned}\mathbb{E}\int_0^t \left\|C\Phi(t,s)B - C_1\tilde{\Phi}(t,s)B_1\right\|_F^2 ds = &\operatorname{tr}\left(C\int_0^t \mathbb{E}\left[\Phi(t-s)BB^T\Phi^T(t-s)\right]ds\, C^T\right) \\ &+ \operatorname{tr}\left(C_1\int_0^t \mathbb{E}\left[\tilde{\Phi}(t-s)B_1B_1^T\tilde{\Phi}^T(t-s)\right]ds\, C_1^T\right) \\ &- 2\ \operatorname{tr}\left(C\int_0^t \mathbb{E}\left[\Phi(t-s)BB_1^T\tilde{\Phi}^T(t-s)\right]ds\, C_1^T\right).\end{aligned}$$

By substitution, we obtain

$$\begin{aligned}\mathbb{E}\int_0^t \left\|C\Phi(t,s)B - C_1\tilde{\Phi}(t,s)B_1\right\|_F^2 ds = &\ \operatorname{tr}\left(C\int_0^t \mathbb{E}\left[\Phi(s)BB^T\Phi^T(s)\right]ds\, C^T\right) \\ &+ \operatorname{tr}\left(C_1\int_0^t \mathbb{E}\left[\tilde{\Phi}(s)B_1B_1^T\tilde{\Phi}^T(s)\right]ds\, C_1^T\right) \\ &- 2\ \operatorname{tr}\left(C\int_0^t \mathbb{E}\left[\Phi(s)BB_1^T\tilde{\Phi}^T(s)\right]ds\, C_1^T\right).\end{aligned}$$

Provided the homogenous equation of the truncated system is still asymptotically mean square stable it holds

$$\begin{aligned}\mathbb{E}\left\|\hat{\mathcal{Y}}(t) - \mathcal{Y}(t)\right\|_2 &\leq \left(\mathbb{E}\int_0^\infty \left\|C\Phi(t,s)B - C_1\tilde{\Phi}(t,s)B_1\right\|_F^2 ds\right)^{\frac{1}{2}} \left(\mathbb{E}\int_0^t \|u(s)\|_2^2\, ds\right)^{\frac{1}{2}} \\ &= \left(\operatorname{tr}\left(CPC^T\right) + \operatorname{tr}\left(C_1P_RC_1^T\right) - 2\ \operatorname{tr}\left(CP_MC_1^T\right)\right)^{\frac{1}{2}}\|u\|_{L_t^2},\end{aligned}$$

where $P = \mathbb{E}\int_0^\infty \Phi(\tau)BB^T\Phi^T(\tau)d\tau$ is the reachability Gramian of the original system, $P_R = \mathbb{E}\int_0^\infty \tilde{\Phi}(\tau)B_1B_1^T\tilde{\Phi}^T(\tau)d\tau \in \mathbb{R}^{r\times r}$ the reachability Gramian of the approximating system and $P_M =$

$\mathbb{E}\int_0^\infty \Phi(\tau)BB_1^T\tilde{\Phi}^T(\tau)d\tau \in \mathbb{R}^{n\times r}$ a matrix that fulfills the following equation:

$$0 = BB_1^T + P_M A_{11}^T + AP_M + \sum_{k=1}^{q} \Psi^k P_M (\Psi_{11}^k)^T \,\mathbb{E}\left[M_k(1)^2\right], \tag{40}$$

which we get by taking the limit $t \to \infty$ on both sides of equation (37). We summarize these results in the following theorem:

Theorem 4.5. *Let $\left(A, \Psi^k, B, C\right)$ be a realization of system (24). Suppose that the reduced order model with the coefficients $\left(A_{11}, \Psi_{11}^k, B_1, C_1\right)$ defined in (33) is asymptotically mean square stable, then*

$$\sup_{t\in[0,T]} \mathbb{E}\left\|\hat{\mathcal{Y}}(t) - \mathcal{Y}(t)\right\|_2 \leq \left(\operatorname{tr}\left(CPC^T\right) + \operatorname{tr}\left(C_1P_RC_1^T\right) - 2\ \operatorname{tr}\left(CP_MC_1^T\right)\right)^{\frac{1}{2}} \|u\|_{L_T^2} \tag{41}$$

for every $T > 0$, where $\mathcal{Y}$ is the output of the original and $\hat{\mathcal{Y}}$ the output of the reduced system. Here, P denotes the reachability Gramian of system (24), P_R denotes the reachability Gramian of reduced system and P_M satisfies equation (40).

Remark. *If $u \in L^2$ we can replace $\|\cdot\|_{L_T^2}$ by $\|\cdot\|_{L^2}$ and $[0,T]$ by $\mathbb{R}_+$ in inequality (41).*

We want to specify the error bound for a particular case.

Proposition 4.6. *If the realization (A, Ψ^k, B, C) is balanced and $\Psi_{12}^k = \Psi_{21}^k = 0$ for $k = 1, \ldots, q$, then*

$$\operatorname{tr}\left(CPC^T\right) + \operatorname{tr}\left(C_1P_RC_1^T\right) - 2\ \operatorname{tr}\left(CP_MC_1^T\right) = \operatorname{tr}((C_2^TC_2 + 2P_{M,2}A_{21}^T)\Sigma_2),$$

where $P_{M,2}$ are the last $n-r$ rows of P_M and $\Sigma_2 = \operatorname{diag}(\sigma_{r+1}, \ldots, \sigma_n)$.

Proof. We have

$$\begin{bmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{bmatrix} \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} + \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \sum_{k=1}^{q} \begin{bmatrix} (\Psi_{11}^k)^T & 0 \\ 0 & (\Psi_{22}^k)^T \end{bmatrix} \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} \Psi_{11}^k & 0 \\ 0 & \Psi_{22}^k \end{bmatrix} \cdot c_k$$
$$= -\begin{bmatrix} C_1^TC_1 & C_1^TC_2 \\ C_2^TC_1 & C_2^TC_2 \end{bmatrix},$$

where $c_k = \mathbb{E}\left[M_k(1)^2\right]$. Hence,

$$A_{11}^T\Sigma_1 + \Sigma_1A_{11} + \sum_{k=1}^{q} (\Psi_{11}^k)^T\Sigma_1\Psi_{11}^kc_k = -C_1^TC_1,$$

such that Σ_1 is the observability Gramian of the reduced order model. In addition, it is also easy to check that Σ_1 represents the reachability Gramian of the reduced system. We define $\mathcal{E} := \left(\operatorname{tr}\left(CPC^T\right) + \operatorname{tr}\left(C_1P_RC_1^T\right) - 2\ \operatorname{tr}\left(CP_MC_1^T\right)\right)^{\frac{1}{2}}$ and obtain

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

It holds

$$-C_1^TC_2 = A_{21}^T\Sigma_2 + \Sigma_1A_{12}.$$

This yields

$$\mathcal{E}^2 = \operatorname{tr}(C_2\Sigma_2C_2^T + 2C_1\Sigma_1C_1^T - 2C_1P_{M,1}C_1^T) + 2\operatorname{tr}(A_{21}^T\Sigma_2P_{M,2} + \Sigma_1A_{12}P_{M,2}).$$

We have

$$-B_1B_1^T = P_{M,1}A_{11}^T + A_{11}P_{M,1} + A_{12}P_{M,2} + \sum_{k=1}^{q}\Psi_{11}^kP_{M,1}(\Psi_{11}^k)^Tc_k.$$

Thus,

$$\Sigma_1A_{12}P_{M,2} = -\Sigma_1(B_1B_1^T + P_{M,1}A_{11}^T + A_{11}P_{M,1} + \sum_{k=1}^{q}\Psi_{11}^kP_{M,1}(\Psi_{11}^k)^Tc_k).$$

Since

$$\begin{aligned}
\operatorname{tr}(\Sigma_1A_{12}P_{M,2}) &= -\operatorname{tr}(\Sigma_1(B_1B_1^T + P_{M,1}A_{11}^T + A_{11}P_{M,1} + \sum_{k=1}^{q}\Psi_{11}^kP_{M,1}(\Psi_{11}^k)^Tc_k)) \\
&= -\operatorname{tr}(\Sigma_1B_1B_1^T + A_{11}^T\Sigma_1P_{M,1} + \Sigma_1A_{11}P_{M,1} + \sum_{k=1}^{q}(\Psi_{11}^k)^T\Sigma_1\Psi_{11}^kc_kP_{M,1}) \\
&= -\operatorname{tr}(B_1^T\Sigma_1B_1) + \operatorname{tr}(C_1^TC_1P_{M,1}),
\end{aligned}$$

we have

$$\begin{aligned}
\mathcal{E}^2 &= \operatorname{tr}(C_2\Sigma_2C_2^T + 2C_1\Sigma_1C_1^T - 2C_1P_{M,1}C_1^T) - \operatorname{tr}(2B_1^T\Sigma_1B_1) + \operatorname{tr}(2C_1^TC_1P_{M,1}) \\
&\quad + 2\operatorname{tr}(A_{21}^T\Sigma_2P_{M,2}) \\
&= \operatorname{tr}(C_2\Sigma_2C_2^T) + 2\operatorname{tr}(A_{21}^T\Sigma_2P_{M,2}) = \operatorname{tr}((C_2^TC_2 + 2P_{M,2}A_{21}^T)\Sigma_2).
\end{aligned}$$

□

The error bound we obtained in Proposition 4.6 has the same structure as the $\mathcal{H}^2$ error bound in the deterministic case, which can be found in Section 7.2.2 in Antoulas [1].

5 Applications

In order to demonstrate the use of the model reduction method introduced in Section 4 we apply it in the context of the numerical solution of linear controlled evolution equations with Levy noise. Therefor, we apply the Galerkin scheme on that evolution equation and end up with a sequence of ordinary stochastic differential equations. Then, we use balanced truncation for reducing the dimension of the Galerkin solution. Finally, we compute the error bounds and exact errors for the example considered here.

5.1 Finite dimensional approximations for stochastic evolution equations

In this section we deal with an infinite dimensional system, where the noise process is denoted by M. We suppose that M is a Levy process with values in a separable Hilbert space U. Additionally, we assume that M is square integrable with zero mean. The most important properties regarding this process and the definition of an integral with respect to M one can find in the book of Peszat, Zabczyk [20].
Suppose $A : D(A) \to H$ is a densely defined linear operator being self adjoint and negative definite

such that we have an orthonormal basis $(h_k)_{k\in\mathbb{N}}$ of H consisting of eigenvectors of A:

$$Ah_k = -\lambda_k h_k,$$

where $0 \leq \lambda_1 \leq \lambda_2 \leq \dots$ are the corresponding eigenvalues. Furthermore, the linear operator A generates a contraction C_0-semigroup $(S(t))_{t\geq 0}$ defined by

$$S(t)x = \sum_{k=1}^{\infty} \mathrm{e}^{-\lambda_k t} \langle x, h_k \rangle h_k$$

for $x \in H$ being exponentially stable for the case $0 < \lambda_1$. By Q we denote the covariance operator of M which is a symmetric and positive definite trace class operator that is characterized by

$$\mathbb{E} \langle M(t), x\rangle_U \langle M(s), y\rangle_U = \min\{t, s\} \langle Qx, y\rangle_U$$

for $x, y \in U$ and $s, t \geq 0$. We can choose an orthonormal basis of U consisting of eigenvectors $(u_k)_{k\in\mathbb{N}}$ of Q.[6] The corresponding eigenvalues we denote by $(\mu_k)_{k\in\mathbb{N}}$ such that

$$Qu_k = \mu_k u_k.$$

We consider the following stochastic differential equation:

$$\begin{aligned} dX(t) &= [AX(t) + Bu(t)]\, dt + \Psi(X(t-))dM(t), \quad X(0) = x_0 \in H, \\ Y(t) &= CX(t), \quad t \geq 0. \end{aligned} \tag{42}$$

We make the following assumptions:

- Ψ is a linear mapping on H with values in the set of all linear operators from U to H such that $\Psi(h)Q^{\frac{1}{2}}$ is a Hilbert Schmidt operator for every $h \in H$. In addition,

$$\left\| \Psi(h)Q^{\frac{1}{2}} \right\|_{L_{HS}(U,H)} \leq \tilde{M} \|h\|_H \tag{43}$$

  holds for $\tilde{M} > 0$.

- The process $u : \mathbb{R}_+ \times \Omega \to \mathbb{R}^m$ is $(\mathcal{F}_t)_{t\geq 0}$-adapted with

$$\int_0^T \mathbb{E} \|u(s)\|_2^2 \, ds < \infty$$

  for each $T > 0$.

- B is a linear and bounded operator on $\mathbb{R}^m$ with values in H and $C \in L(H, \mathbb{R}^p)$.

Definition 5.1. *A cadlag process $(X(t))_{t\geq 0}$ with values in H is called mild solution of (42) if $\mathbb{P}$-almost surely*

$$X(t) = S(t)x_0 + \int_0^t S(t-s)Bu(s)ds + \int_0^t S(t-s)\Psi(X(s-))dM(s) \tag{44}$$

holds for all $t \geq 0$.

Remark. *Since the operator A generates a contraction semigroup, the stochastic convolution in equation (44) has a cadlag modification (Theorem 9.24 in [20]), which enables us to construct a cadlag mild solution of equation (42). This solution is unique for every fixed u considering Theorem 9.29 in [20].*

[6] By Theorem VI.21 in Reed, Simon [22] Q is a compact operator such that this property follows by the spectral theorem.

We will now approximate the mild solution of the infinite dimensional equation (42). We use the Galerkin method for a finite dimensional approximation that one can for example find in Grecksch, Kloeden [10]. There they dealt with strong solutions of stochastic evolution equations with scalar Wiener noise.

We construct a sequence $(X_n)_{n\in\mathbb{N}}$ of finite dimensional cadlag processes with values in $H_n = \operatorname{span}\{h_1, \dots, h_n\}$ given by

$$\begin{aligned} dX_n(t) &= [A_n X_n(t) + B_n u(t)]\,dt + \Psi_n(X_n(t-))dM_n(t), \quad t \geq 0, \\ X_n(0) &= x_{0,n}, \end{aligned} \tag{45}$$

where

- $M_n(t) = \sum_{k=1}^n \langle M(t), u_k\rangle_U u_k$, $t \geq 0$, is a $\operatorname{span}\{u_1, \dots, u_n\}$-valued Levy process,
- $A_n x = \sum_{k=1}^n \langle Ax, h_k\rangle_H h_k \in H_n$ holds for all $x \in D(A)$,
- $B_n x = \sum_{k=1}^n \langle Bx, h_k\rangle_H h_k \in H_n$ holds for all $x \in H$,
- $\Psi_n(x)y = \sum_{k=1}^n \langle \Psi(x)y, h_k\rangle_H h_k \in H_n$ holds for all $y \in U$ and $x \in H$,
- $x_{0,n} = \sum_{k=1}^n \langle x_0, h_k\rangle_H h_k \in H_n$.

We know that A_n generates a C_0-semigroup $(S_n(t))_{t\geq 0}$ on H_n which is defined by $S_n(t)x = \sum_{k=1}^n \langle S(t)x, h_k\rangle_H h_k$ for all $x \in H$ such that the mild solution of equation (45) is given by

$$X_n(t) = S_n(t)x_{0,n} + \int_0^t S_n(t-s)B_n u(s)ds + \int_0^t S_n(t-s)\Psi_n(X_n(s-))dM_n(s)$$

for $t \geq 0$. Furthermore, we consider the p dimensional approximating output

$$Y_n(t) = CX_n(t), \quad t \geq 0.$$

With similar arguments like in the proof of Theorem 1 in Grecksch, Kloeden [10] one can show the follwing theorem and hence

$$\mathbb{E}\,\|Y_n(t) - Y(t)\|_2^2 \to 0$$

is true for $n \to \infty$ and $t \geq 0$:

Theorem 5.2. *It holds*

$$\mathbb{E}\,\|X_n(t) - X(t)\|_H^2 \to 0$$

for $n \to \infty$ and $t \geq 0$.

Remark. *If $U = \mathbb{R}^q$ one has to replace M_n by M in equation (45) and Theorem 5.2 holds for this case as well.*

We now determine the components of Y_n. They are given by

$$Y_n^l(t) = \langle Y_n(t), e_l\rangle_{\mathbb{R}^p} = \langle CX_n(t), e_l\rangle_{\mathbb{R}^p} = \sum_{k=1}^n \langle Ch_k, e_l\rangle_{\mathbb{R}^p} \langle X_n(t), h_k\rangle_H$$

for $l = 1, \dots, p$, where e_l is the lth unit vector in $\mathbb{R}^p$. We set

$$\mathcal{X}(t) = (\langle X_n(t), h_1\rangle_H, \dots, \langle X_n(t), h_n\rangle_H)^T \text{ and } \mathcal{C} = (\langle Ch_k, e_l\rangle_{\mathbb{R}^p})_{\substack{l=1,\dots,p \\ k=1,\dots,n}}$$

and obtain

$$Y_n(t) = \mathcal{C}\mathcal{X}(t), \quad t \geq 0.$$

The components of $\mathcal{X}$ fulfill the following:

$$\begin{aligned}\langle X_n(t), h_k \rangle_H = \langle S_n(t)x_{0,n}, h_k \rangle_H &+ \int_0^t \langle S_n(t-s)B_n u(s), h_k \rangle_H \, ds \\ &+ \left\langle \int_0^t S_n(t-s)\Psi_n(X_n(s-))dM_n(s), h_k \right\rangle_H .\end{aligned}$$

Considering the representation $S_n(t)x = \sum_{i=1}^n \mathrm{e}^{-\lambda_i t} \langle x, h_i \rangle_H h_i$ ($x \in H$) we have

$$\langle S_n(t)x_{0,n}, h_k \rangle_H = \mathrm{e}^{-\lambda_k t} \langle x_{0,n}, h_k \rangle_H = \mathrm{e}^{-\lambda_k t} \langle x_0, h_k \rangle_H ,$$

and

$$\begin{aligned}\langle S_n(t-s)B_n u(s), h_k \rangle_H &= \mathrm{e}^{-\lambda_k(t-s)} \langle B_n u(s), h_k \rangle_H = \mathrm{e}^{-\lambda_k(t-s)} \langle Bu(s), h_k \rangle_H \\ &= \sum_{l=1}^m \mathrm{e}^{-\lambda_k(t-s)} \langle Be_l, h_k \rangle_H \langle u(s), e_l \rangle_{\mathbb{R}^m}\end{aligned}$$

for $k = 1, \ldots, n$, where e_l is the lth unit vector in $\mathbb{R}^m$. Furthermore,

$$\begin{aligned}&\left\langle \int_0^t S_n(t-s)\Psi_n(X_n(s-))dM_n(s), h_k \right\rangle_H = \sum_{j=1}^n \int_0^t \langle S_n(t-s)\Psi_n(X_n(s-))u_j, h_k \rangle_H \, d\langle M(s), u_j \rangle_U \\ &= \sum_{j=1}^n \sum_{i=1}^n \int_0^t \langle S_n(t-s)\Psi_n(h_i)u_j, h_k \rangle_H \langle X_n(t-), h_i \rangle_H \, d\langle M(s), u_j \rangle_U \\ &= \sum_{j=1}^n \sum_{i=1}^n \int_0^t \mathrm{e}^{-\lambda_k(t-s)} \langle \Psi(h_i)u_j, h_k \rangle_H \langle X_n(t-), h_i \rangle_H \, d\langle M(s), u_j \rangle_U .\end{aligned}$$

Hence, in compact form $\mathcal{X}$ is given by

$$\mathcal{X}(t) = \mathrm{e}^{\mathcal{A}t} \mathcal{X}_0 + \int_0^t \mathrm{e}^{\mathcal{A}(t-s)} \mathcal{B}u(s)ds + \sum_{j=1}^n \int_0^t \mathrm{e}^{\mathcal{A}(t-s)} \mathcal{N}^j \mathcal{X}(s-)dM^j(s), \tag{46}$$

where

- $\mathcal{A} = \operatorname{diag}(-\lambda_1, \ldots, -\lambda_n)$, $\mathcal{B} = (\langle Be_i, h_k \rangle_H)_{\substack{k=1,\ldots,n \\ i=1,\ldots,m}}$, $\mathcal{N}^j = \left(\langle \Psi(h_i)u_j, h_k \rangle_H\right)_{k,i=1,\ldots,n}$,
- $\mathcal{X}_0 = (\langle x_0, h_1 \rangle_H, \ldots, \langle x_0, h_n \rangle_H)^T$ and $M^j(s) = \langle M(s), u_j \rangle_U$.

The processes M^j are uncorrelated real-valued Levy processes with $\mathbb{E}\left|M^j(t)\right|^2 = t\mu_j$, $t \geq 0$, and zero mean. Below, we show that the solution of equation (46) fulfills the strong solution equation as well. Therefor, we set

$$f(t) := \mathcal{X}_0 + \int_0^t \mathrm{e}^{-\mathcal{A}s} \mathcal{B}u(s)ds + \sum_{j=1}^n \int_0^t \mathrm{e}^{-\mathcal{A}s} \mathcal{N}^j \mathcal{X}(s-)dM^j(s), \quad t \geq 0,$$

and determine the stochastic differential of $\mathrm{e}^{\mathcal{A}t} f(t)$ via the Ito product formula in Corollary 2.4:

$$\begin{aligned}
e_i^T \mathcal{X}(t) &= e_i^T \mathrm{e}^{\mathcal{A}t} f(t) = e_i^T f(0) + \int_0^t d\left(e_i^T \mathrm{e}^{\mathcal{A}s}\right) f(s-) + \int_0^t e_i^T \mathrm{e}^{\mathcal{A}s} df(s) \\
&= e_i^T \left(\mathcal{X}_0 + \int_0^t \mathcal{A}\mathrm{e}^{\mathcal{A}s} f(s)ds + \int_0^t \mathcal{B}u(s)ds + \sum_{j=1}^n \int_0^t \mathcal{N}^j \mathcal{X}(s-)dM^j(s) \right),
\end{aligned}$$

where e_i is the ith unit vector of $\mathbb{R}^n$ and the quadratic covariation terms are zero, since $t \mapsto e_i^T \mathrm{e}^{\mathcal{A}t}$ is a continuous semimartingale with a martingale part of zero. Hence,

$$\mathcal{X}(t) = \mathcal{X}_0 + \int_0^t \mathcal{A}\mathcal{X}(s) + \mathcal{B}u(s)ds + \sum_{j=1}^n \int_0^t \mathcal{N}^j \mathcal{X}(s-)dM^j(s), \quad t \geq 0.$$

Example 5.3. *We consider a bar of lenght π, which is heated on $[0, \frac{\pi}{2}]$. The temperature of the bar is described by the following stochastic partial differential equation:*

$$\begin{aligned}
\frac{\partial X(t,\zeta)}{\partial t} &= \frac{\partial^2}{\partial \zeta^2} X(t,\zeta) + 1_{[0,\frac{\pi}{2}]}(\zeta)u(t) + aX(t-,\zeta)\frac{\partial M(t)}{\partial t} \\
X(t,0) &= 0 = X(t,\pi), \\
X(0,\zeta) &= x_0(\zeta)
\end{aligned} \tag{47}$$

for $t \geq 0$ and $\zeta \in [0,\pi]$. Thereby, we assumed that M is a scalar square integrable Levy process with zero mean, $H = L^2([0,\pi])$, $U = \mathbb{R}$, $m = 1$, $A = \frac{\partial^2}{\partial \zeta^2}$ Furthermore, we set $B = 1_{[0,\frac{\pi}{2}]}(\cdot)$ and $\Psi(x) = ax$ for $x \in L^2([0,\pi])$. Additionally, we assume $\mathbb{E}\left[M(1)^2\right] a^2 < 2$, which is equivalent to the homogeneous solution fulfills

$$\mathbb{E}\left\|X^h(t,\cdot)\right\|_H^2 \leq c\,\mathrm{e}^{-\alpha t} \left\|x_0(\cdot)\right\|_H^2 \tag{48}$$

for $c, \alpha > 0$. This equivalence is a consequence of Theorem 3.1 in Ichikawa [13] and Theorem 5 in Haussmann [11]. For further information regarding the exponential mean square stability condition (48) see Section 5 in Curtain [7].[7] *It is a well known fact that here the eigenvalues of the second derivative are given by $-\lambda_k = -k^2$ and the corresponding eigenvectors which represent an orthonormal basis are $h_k = \sqrt{\frac{2}{\pi}} \sin(k\cdot)$. We are interested in the avarage temparature of the bar on $[\frac{\pi}{2}, \pi]$ such that the scalar output of the system is*

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

where $Cx = \frac{2}{\pi} \int_{\frac{\pi}{2}}^{\pi} x(\zeta)d\zeta$ for $x \in L^2([0,\pi])$. We approximate Y via

$$Y_n(t) = \mathcal{C}\mathcal{X}(t),$$

$\mathcal{C}^T = (Ch_k)_{k=1,\dots,n} = \left(\left(\frac{2}{\pi}\right)^{\frac{3}{2}} \frac{1}{k} \left[\cos(\frac{k\pi}{2}) - \cos(k\pi)\right]\right)_{k=1,\dots,n}$.

$\mathcal{X}$ is given by

$$\mathcal{X}(t) = \mathcal{X}_0 + \int_0^t \mathcal{A}\mathcal{X}(s) + \mathcal{B}u(s)ds + \int_0^t \mathcal{N}\mathcal{X}(s-)dM(s), \tag{49}$$

where

[7] Curtain, Ichikawa and Haussmann stated these conditions for exponential mean square stability for the Wiener case, which can be easily generalized for the case of square integrable Levy process with mean zero.

- $\mathcal{A} = \operatorname{diag}\left(-1, -4, \ldots, -n^2\right)$,
- $\mathcal{N} = \left(\langle \Psi(h_i), h_k\rangle_H\right)_{k,i=1,\ldots,n} = \left(\langle a h_i, h_k\rangle_H\right)_{k,i=1,\ldots,n} = aI_n$,
- $\mathcal{B} = \left(\langle B, h_k\rangle_H\right)_{k=1,\ldots,n} = \left(\left\langle 1_{[0,\frac{\pi}{2}]}(\cdot), h_k\right\rangle_H\right)_{k=1,\ldots,n} = \left(\left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{1}{k}\left[1-\cos(\frac{k\pi}{2})\right]\right)_{k=1,\ldots,n}$.

Since we now choose $x_0 \equiv 0$ for simplicity, we additionally have $\mathcal{X}_0 = 0$.

Next, we consider a more complex example with a two dimensional state variable:

Example 5.4. *We determine the Galerkin solution of the following controlled stochastic partial differential equation:*

$$\begin{aligned}
\frac{\partial X(t,\zeta)}{\partial t} &= \Delta X(t,\zeta) + 1_{[\frac{\pi}{4},\frac{3\pi}{4}]^2}(\zeta)u(t) + \mathrm{e}^{-\left|\zeta_1-\frac{\pi}{2}\right|-\zeta_2} X(t-,\zeta)\frac{\partial M(t)}{\partial t}, \quad t\geq 0,\ \zeta\in[0,\pi]^2, \quad (50)\\
\frac{\partial X(t,\zeta)}{\partial \mathbf{n}} &= 0, \quad t\geq 0,\ \zeta \in \partial[0,\pi]^2,\\
X(0,\zeta) &\equiv 0.
\end{aligned}$$

Again, M is a scalar square integrable Levy process with zero mean, $H = L^2([0,\pi]^2)$, $U = \mathbb{R}$, $m = 1$, A is the Laplace operator, $B = 1_{[\frac{\pi}{4},\frac{3\pi}{4}]^2}(\cdot)$ and $\Psi(x) = \mathrm{e}^{-\left|\cdot-\frac{\pi}{2}\right|-\cdot}\, x$ for $x \in L^2([0,\pi]^2)$. The eigenvalues of the Laplacian on $[0,\pi]^2$ are given by $-\lambda_{ij} = -(i^2+j^2)$ and the corresponding eigenvectors which represent an orthonormal basis are $h_{ij} = \frac{2}{\pi}\cos(i\cdot)\cos(j\cdot)$. For simplicity we write $-\lambda_k$ for the kth largest eigenvalue and the corresponding eigenvector we denote by h_k. The scalar output of the system is

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

where $Cx = \frac{4}{3\pi^2}\int_{[0,\pi]^2\setminus[\frac{\pi}{4},\frac{3\pi}{4}]^2} x(\zeta)d\zeta$ for $x \in L^2([0,\pi]^2)$. The output of the Galerkin system is

$$Y_n(t) = \mathcal{C}\mathcal{X}(t),$$

$\mathcal{C}^T = (Ch_k)_{k=1,\ldots,n}$. The Galerkin solutions $\mathcal{X}$ satisfies

$$\mathcal{X}(t) = \int_0^t \mathcal{A}\mathcal{X}(s) + \mathcal{B}u(s)ds + \int_0^t \mathcal{N}\mathcal{X}(s-)dM(s), \quad (51)$$

where $\mathcal{A} = \operatorname{diag}(0,-1,-1,-2,\ldots)$, $\mathcal{N} = \left(\langle \Psi(h_i), h_k\rangle_H\right)_{k,i=1,\ldots,n}$, $\mathcal{B} = \left(\langle B, h_k\rangle_H\right)_{k=1,\ldots,n}$.

5.2 Error bounds of the examples

We consider the system from Example 5.3. Using Theorem 3.3 the uncontrolled equation (49) is asymptotically mean square stable if and only if the Kronecker matrix

$$I_n\otimes\mathcal{A} + \mathcal{A}\otimes I_n + \mathcal{N}\otimes\mathcal{N}\cdot\mathbb{E}\left[M(1)^2\right] = I_n\otimes\mathcal{A} + (\mathcal{A} + \mathbb{E}\left[M(1)^2\right]a^2 I_n)\otimes I_n$$

is Hurwitz. From Section 2.6 in Steeb [24] we can conclude that the largest eigenvalue of the Kronecker matrix is $-2+\mathbb{E}\left[M(1)^2\right]a^2$. Thus, the homogeneous solution of system (49) is asymptotically mean square stable if and only if $\mathbb{E}\left[M(1)^2\right]a^2 < 2$, which is fulfilled by (48).
We want to obtain a reduced order model via balanced truncation. We choose $a = \mathbb{E}\left[M(1)^2\right] = 1$ and additionally let $n = 1000$. It turns out that the system is neither completely observable nor completely reachable since the Gramians do not have full rank. So, we need an alternative method to determine the reduced order model. We use a method for non minimal systems that is known from the deterministic case and which is for example described in Section 1.4.2 in Benner et al.

[5]. In this algorithm we do not compute the full transformation matrix T. So, we obtain the matrices of the reduced order model by

$$\tilde{\mathcal{A}} = W^T \operatorname{diag}(-1, \ldots, -n^2)V, \quad \tilde{\mathcal{N}} = W^T I_n V = I_r, \quad \tilde{\mathcal{B}} = W^T \mathcal{B}, \quad \tilde{\mathcal{C}} = \mathcal{C}V.$$

Thereby,

$$W^T = \Sigma_1^{-\frac{1}{2}} V_1^T R \text{ and } V = S^T U_1 \Sigma_1^{-\frac{1}{2}},$$

where V_1 and U_1 are from the SVD of SR^T:

$$SR^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix},$$

where $Q = R^T R$ and $P = S^T S$. Reducing the model yields the following error bounds:

Dimension of the reduced order model	$\left(\operatorname{tr}\left(\mathcal{C}P\mathcal{C}^T\right) + \operatorname{tr}\left(\tilde{\mathcal{C}}P_R\tilde{\mathcal{C}}^T\right) - 2 \operatorname{tr}\left(\mathcal{C}P_M\tilde{\mathcal{C}}^T\right)\right)^{\frac{1}{2}}$
8	$4.5514 \cdot 10^{-6}$
4	$2.3130 \cdot 10^{-4}$
2	$1.7691 \cdot 10^{-3}$
1	0.0879

Below, we reduce the Galerkin solution of Example 5.4 with dimension $n = 1000$ and $\mathbb{E}\left[M(1)^2\right] = 1$. Here, the matrix $\mathcal{A} = \operatorname{diag}(0, -1, -1, -2, \ldots)$ is not stable, such that we need to stabilize system (51) before using balanced truncation. Inserting the feedback control $u(t) = -2e_1^T \mathcal{X}(t)$, $t \geq 0$, where e_1 is the first unit vector in $\mathbb{R}^n$, yields a asymptotically mean square stable system, since the following sufficient condition holds (see Corollary 3.6.3 in [9] and Theorem 5 in [11]): $\mathcal{A}_S = \mathcal{A} - 2\mathcal{B}e_1^T$ is stable and

$$\left\| \int_0^\infty \mathrm{e}^{\mathcal{A}_S^T t} \mathcal{N}^T \mathcal{N} \, \mathrm{e}^{\mathcal{A}_S t} \, dt \right\| = 0.0658 < 1.$$

We repeat the procedure from above and obtain

Dimension of the reduced order model	$\left(\operatorname{tr}\left(\mathcal{C}P\mathcal{C}^T\right) + \operatorname{tr}\left(\tilde{\mathcal{C}}P_R\tilde{\mathcal{C}}^T\right) - 2 \operatorname{tr}\left(\mathcal{C}P_M\tilde{\mathcal{C}}^T\right)\right)^{\frac{1}{2}}$
8	$3.7545 \cdot 10^{-6}$
4	$6.4323 \cdot 10^{-4}$
2	$3.1416 \cdot 10^{-3}$
1	0.0333

for the the stabilized system (51) meaning that we replaced $\mathcal{A}$ by $\mathcal{A}_S$.

5.3 Comparison between exact error and error bound

Since equations (49) and (51) do not have an explicit solution in general we need to discretize in time for computing the exact error of the estimation given here. For simplicity, we assume that $n = 80$ and M is a scalar Wiener process and use the Euler-Maruyama scheme[8] for approximating the original system:

$$\mathcal{X}_{k+1} = \mathcal{X}_k + \left(\mathcal{A}\mathcal{X}_k + \mathcal{B}u(t_k)\right) h + \mathcal{N}\mathcal{X}_k \Delta M_k$$

[8]The theory regarding this method one can find in Kloeden, Platen [15].

and the reduced order model:

$$\tilde{\mathcal{X}}_{k+1} = \tilde{\mathcal{X}}_k + \left(\tilde{\mathcal{A}}\tilde{\mathcal{X}}_k + \tilde{\mathcal{B}}u(t_k)\right) h + \tilde{\mathcal{N}}\tilde{\mathcal{X}}_k \Delta M_k,$$

where we consider these equations on the time interval $[0, \pi]$. Furthermore, we choose $\mathcal{X}_0 = 0$, $h = \frac{\pi}{10000}$ and $t_k = kh$ for $k = 0, 1, \ldots, 10000$, $\Delta M_k = M(t_{k+1}) - M(t_k)$.

For system (49) we insert the normalized control functions $u_1(t) = \frac{\sqrt{2}}{\pi} M(t)$, $u_2(t) = \sqrt{\frac{2}{\pi}} \cos(t)$, $u_3(t) = \sqrt{\frac{2}{1-\mathrm{e}^{-2\pi}}}\,\mathrm{e}^{-t}$, $t \in [0, \pi]$ and obtain $\mathcal{D} := \max_{k=1,\ldots,10000} \mathbb{E}\left|\mathcal{C}X_k - \tilde{\mathcal{C}}\tilde{X}_k\right|$ for different dimensions of the reduced order model (ROM) and different inputs:

Dimension of the ROM	$\mathcal{D}$ with $u = u_1$	$\mathcal{D}$ with $u = u_2$	$\mathcal{D}$ with $u = u_3$	$\mathcal{EB}$
8	$9.0615 \cdot 10^{-9}$	$7.8832 \cdot 10^{-8}$	$1.3987 \cdot 10^{-7}$	$1.4813 \cdot 10^{-6}$
4	$3.8702 \cdot 10^{-6}$	$6.4204 \cdot 10^{-6}$	$1.1353 \cdot 10^{-5}$	$2.2706 \cdot 10^{-4}$
2	$6.8932 \cdot 10^{-5}$	$1.1195 \cdot 10^{-4}$	$1.9549 \cdot 10^{-4}$	$1.7671 \cdot 10^{-3}$
1	0.0141	0.0243	0.0354	0.0879

,

where $\mathcal{EB} := \left(\mathrm{tr}\left(\mathcal{C}P\mathcal{C}^T\right) + \mathrm{tr}\left(\tilde{\mathcal{C}}P_R\tilde{\mathcal{C}}^T\right) - 2\ \mathrm{tr}\left(\mathcal{C}P_M\tilde{\mathcal{C}}^T\right)\right)^{\frac{1}{2}}$.

For system (51) we use the inputs $\tilde{u}_i(t) = -2e_1^T\mathcal{X}(t) + u_i(t)$, $t \geq 0$, $i = 1, 2, 3$ and obtain

Dimension of the ROM	$\mathcal{D}$ with $u = \tilde{u}_1$	$\mathcal{D}$ with $u = \tilde{u}_2$	$\mathcal{D}$ with $u = \tilde{u}_3$	$\mathcal{EB}$
8	$5.6162 \cdot 10^{-7}$	$5.5374 \cdot 10^{-7}$	$6.5699 \cdot 10^{-7}$	$3.5376 \cdot 10^{-6}$
4	$4.7245 \cdot 10^{-5}$	$5.2722 \cdot 10^{-5}$	$6.8758 \cdot 10^{-5}$	$3.1487 \cdot 10^{-4}$
2	$5.1270 \cdot 10^{-4}$	$4.6627 \cdot 10^{-4}$	$6.2103 \cdot 10^{-4}$	$2.4164 \cdot 10^{-3}$
1	$3.7520 \cdot 10^{-3}$	0.0118	$9.9629 \cdot 10^{-3}$	0.0327

.

6 Conclusion

We generalized balanced truncation for stochastic system with noise processes having jumps. In particular, we focused on a linear controlled state equation driven by uncorrelated Levy processes which is asymptotically mean square stable and equipped with an output equation. We showed that the Gramians we defined are solutions of generalized Lyapunov equations and proofed that the reachable and observable states and the corresponding energy are characterized by these Gramians. We showed that the reduced order model (ROM) is mean square stable, not balanced, the Hankel singular values (HV) of the ROM are not a subset of the HVs of the original system and one can lose complete observability and reachability. Furthermore, we provided an error bound for balanced truncation of the Levy driven system assuming mean square asymptotic stability of the ROM. Finally, we demonstrated the use of balanced truncation for stochastic systems. We applied it in the context of the numerical solution of linear controlled evolution equations with Levy noise and computed the error bounds and exact errors for the example considered here.

Acknowledgements

The authors would like to thank Tobias Damm for his comments and advice and Tobias Breiten for providing Example 4.3.

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] David Applebaum. *Lévy Processes and Stochastic Calculus. 2nd ed.* Cambridge Studies in Advanced Mathematics 116. Cambridge: Cambridge University Press, 2009.

[3] Ludwig Arnold. *Stochastische Differentialgleichungen. Theorie und Anwendungen.* München-Wien: R. Oldenbourg Verlag, 1973.

[4] 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.

[5] Peter Benner, Volker Mehrmann, and Danny C. Sorensen (eds.). *Dimension reduction of large-scale systems. Proceedings of a workshop, Oberwolfach, Germany, October 19–25, 2003.* Lecture Notes in Computational Science and Engineering 45. Berlin: Springer, 2005.

[6] Jean Bertoin. *Lévy Processes.* Cambridge Tracts in Mathematics. 121. Cambridge: Cambridge Univ. Press, 1998.

[7] Ruth F. Curtain. Linear stochastic Ito equations in Hilbert space. Stochastic control theory and stochastic differential systems, Proc. Workshop, Bad Honnef 1979, Lect. Notes Control Inf. Sci. 16, 61-84 , 1979.

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

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

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

[11] Ulrich G. Haussmann. Asymptotic Stability of the Linear Ito Equation in Infinite Dimensions. *J. Math. Anal. Appl.*, 65:219–235, 1978.

[12] Liqian Zhang; James Lam; Biao Huang and Guang-Hong Yang. On gramians and balanced truncation of discrete-time bilinear systems. *Int. J. Control*, 76(4):414–427, 2003.

[13] Akira Ichikawa. Dynamic programming approach to stochastic evolution equations. *SIAM J. Control Optimization*, 17:152–174, 1979.

[14] Jean Jacod and Albert N. Shiryaev. *Limit Theorems for Stochastic Processes. 2nd ed.* Grundlehren der Mathematischen Wissenschaften. 288. Berlin: Springer, 2003.

[15] Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations.* Applications of Mathematics 23. Berlin: Springer, 2010.

[16] Hui-Hsiung Kuo. *Introduction to Stochastic Integration.* Universitext. New York, NJ: Springer, 2006.

[17] Michel Metivier. *Semimartingales: A Course on Stochastic Processes.* De Gruyter Studies in Mathematics, 2. Berlin - New York: de Gruyter, 1982.

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

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

[20] 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.

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

[22] Michael Reed and Barry Simon. *Methods of modern mathematical physics. I: Functional analysis.* New York etc.: Academic Press, A Subsidiary of Harcourt Brace Jovanovich, Publishers, 1980.

[23] Ken-Iti Sato. *Lévy Processes and Infinitely Divisible Distributions.* Cambridge Studies in Advanced Mathematics. 68. Cambridge: Cambridge University Press, 1999.

[24] Willi-Hans Steeb. *Matrix calculus and the Kronecker product with applications and C++ programs. With the collaboration of Tan Kiat Shi.* Singapore: World Scientific, 1997.