The missing link between the output and the $\mathcal{H}_2$-norm of bilinear systems

Martin Redmann*

November 6, 2019

Abstract

In this paper, we prove several new results that give new insights to bilinear systems. We show under which condition a bilinear system is asymptotically stable. Moreover, we provide a global characterization of reachability in bilinear systems based on a certain Gramian. Reachability energy estimates using the same Gramian have only been local so far. The main result of this paper, however, is the link between the output error and the $\mathcal{H}_2$-error of two bilinear systems. This result has several consequences in the field of model order reduction. It explains why $\mathcal{H}_2$-optimal model order reduction leads to good approximations in terms of the output error. Moreover, output errors based on the $\mathcal{H}_2$-norm can now be proved for balancing related model order reduction schemes. These are given in this paper. All these new results are based on a connection between bilinear equations and a linear stochastic differential equations that is established here.

Keywords: Model order reduction, bilinear systems, $\mathcal{H}_2$-error bounds, asymptotic stability, reachability, stochastic systems.
MSC classification: 93A15, 93B05, 93C10, 93D20, 93E03.

1 Introduction

In this paper, we consider bilinear control systems that have applications in various fields [9, 19, 26]. These are of the form

$$\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + \sum_{k=1}^{m} N_k x(t) u_k(t), \quad x(0) = x_0, \\ y(t) = Cx(t), \quad t \geq 0, \end{cases} \tag{1}$$

where $A, N_k \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{p\times n}$ are constant matrices.The vectors x, u and y denote the state, the control input and the quantity of interest (output vector),

*Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin, Germany, Email: martin.redmann@wias-berlin.de. Financial support by the DFG via Research Unit FOR 2402 is gratefully acknowledged.

respectively. The solution of the state equation in (1) is denoted by $x(t, x_0, B)$. The dependence on the initial condition x_0 and the input matrix B is also indicated this way for the other state variables appearing in this paper. We also assume that the matrix A is Hurwitz, meaning that $\sigma(A) \subset \mathbb{C}_- = \{z \in \mathbb{C} : \quad \Re(z) < 0\}$, where $\sigma(\cdot)$ denotes the spectrum of a matrix and $\Re(\cdot)$ represents the real part of a complex number. Furthermore, let $u \in L^2$, i.e.,

$$\|u\|_{L^2}^2 := \int_0^\infty \|u(s)\|_2^2 \, ds = \int_0^\infty u^T(s) u(s) ds < \infty.$$

System (1) can, e.g., represent a spatially discretized partial differential equation. Then, n is usually large and solving (1) becomes computationally expensive, in particular if the system has to be evaluated for many controls u. Therefore, model order reduction (MOR) is vital aiming to replace the original large scale system by a system of small order in order to reduce complexity. We introduce such a reduced order model (ROM) as follows:

$$\hat{\Sigma} : \begin{cases} \dot{\hat{x}}(t) = \hat{A}\hat{x}(t) + \hat{B}u(t) + \sum_{k=1}^{m} \hat{N}_k \hat{x}(t) u_k(t), \quad \hat{x}(0) = \hat{x}_0, \\ \hat{y}(t) = \hat{C}\hat{x}(t), \quad t \geq 0, \end{cases} \tag{2}$$

where $\hat{A}, \hat{N}_k \in \mathbb{R}^{r \times r}$, $\hat{B} \in \mathbb{R}^{r \times m}$, and $\hat{C} \in \mathbb{R}^{p \times r}$ with $r \ll n$. In order to determine the quality of the reduction, it is essential to find a bound $\mathcal{E} \geq 0$ which estimates the error between y and $\hat{y}$, e.g., as follows

$$\sup_{t \geq 0} \|y(t) - \hat{y}(t)\|_2 \leq \mathcal{E} f(u) \tag{3}$$

assuming zero initial conditions, where f is a suitable function. Let us introduce a stochastic equation associated to (1) to find such an error bound. This system is obtained by replacing the control components in the bilinearity of the state equation by the "derivatives" of independent standard Brownian motions $w_1, \dots, w_m$. This results in the following Ito stochastic differential equation:

$$dz(t) = [Az(t) + Bu(t)]dt + \sum_{k=1}^{m} N_k z(t) dw_k(t), \quad z(0) = x_0, \quad t \geq 0. \tag{4}$$

The idea is to approximate the state variable in (1) by z. Establishing this connection, the output error analysis for (1) can be reduced to analyzing output errors for (4). Error bounds for outputs of the form $y_z(t) = Cz(t)$ are already well studied. Relevant bounds for stochastic systems can be found in [8, 22, 24, 25].

In this paper, we first find an estimate for the state variable in (1) which is based on the solution to (4). This link between both states has enormous consequences. On the one hand, we will be able to characterize asymptotic stability for bilinear systems. On the other hand, we will prove that there is an f such that $\mathcal{E} = \left\| \Sigma - \hat{\Sigma} \right\|_{\mathcal{H}_2}$ in (3), i.e., the output error between (1) and (2) can be bounded by an expression depending on the $\mathcal{H}_2$-error between both bilinear systems. This connection has been an open problem for a long time that is extremely important since it finally answers the question why $\mathcal{H}_2$-optimal MOR techniques like the bilinear iterative rational Krylov algorithm [4, 13] lead to good approximations. Moreover, it is possible to find output error bounds based on the $\mathcal{H}_2$-error for balancing related MOR schemes like balanced truncation (BT) [1, 5] and singular perturbation approximation (SPA) [16] which could not be established so far. We will provide these bounds here which allow us to point out the situations in which balancing related methods perform well for bilinear systems. In the case of BT, a different kind of bound can has already been found in [3]. Using a different reachability Gramian in comparison to the work in [1, 5, 16], error bounds for BT and SPA could be achieved [20, 23]. We will also discuss reachability in bilinear systems. Using the output error bound established in this paper, an estimate based on the reachability Gramian proposed in [1] is derived. This estimate allows us to identify states in the system dynamics that require a larger amount of energy to be reached. Previous characterizations of reachability based on the same Gramian have only been local so far [5, 14].

2 Fundamental solutions and solution representations for bilinear and stochastic systems

We briefly recall the concept of fundamental solutions to stochastic linear and deterministic bilinear systems. Furthermore, we state their solution representations. Those will be essential for the error analysis for bilinear systems. We introduce the fundamental solution $\Phi_u(t,s)$, $s \leq t$, of the state equation in (1) as a matrix-valued function solving

$$\Phi_u(t,s) = I + \int_s^t A\Phi_u(\tau,s)d\tau + \sum_{k=1}^m \int_s^t N_k\Phi_u(\tau,s)u_k(\tau)d\tau. \tag{5}$$

For the stochastic differential equation (4), it is a matrix-valued stochastic process $\Phi(t,s)$, $s \leq t$, satisfying

$$\Phi(t,s) = I + \int_s^t A\Phi(\tau,s)d\tau + \sum_{k=1}^m \int_s^t N_k\Phi(\tau,s)dw(\tau). \tag{6}$$

For initial time $s = 0$, we simply write $\Phi_u(t) := \Phi_u(t,0)$ and $\Phi(t) := \Phi(t,0)$. In order to distinguish between state variables with different initial times, we introduce x_s and z_s as the solutions to the state equation in (1) and (4), respectively, with initial time s. For $s = 0$ the index is usually omitted. By multiplying (5) and (6) with x_0 from the right, it can be seen that the solutions to the homogeneous systems ($B = 0$) with initial time $s \geq 0$ are

$$\begin{aligned} x_s(t,x_0,0) &= \Phi_u(t,s)x_0, \quad x_s(s,x_0,0) = x_0, \\ z_s(t,x_0,0) &= \Phi(t,s)x_0, \quad z_s(s,x_0,0) = x_0. \end{aligned}$$

The fundamental solutions moreover allow us to find explicit expressions for the solutions to the bilinear state equation in (1) and to equation (4) for general B. Those are given in the next theorem.

Theorem 2.1. *Let $x(t, x_0, B)$ and $z(t, x_0, B)$, $t \geq 0$, be the solutions to the state equation in (1) and to (4), respectively. Then, they have the following representations:*

$$x(t, x_0, B) = \Phi_u(t)x_0 + \int_0^t \Phi_u(t, s)Bu(s)ds,$$
$$z(t, x_0, B) = \Phi(t)x_0 + \int_0^t \Phi(t, s)Bu(s)ds,$$

where $t \geq 0$, Φ_u is fundamental solution to the bilinear system and Φ the one of the stochastic system.

Proof. The identity for x is obtained by applying the product rule to $\Phi_u(t)g_1(t)$, where $g_1(t) := x_0 + \int_0^t \Phi_u^{-1}(s)Bu(s)ds$ exploiting that $\Phi_u(t, s) = \Phi_u(t)\Phi_u^{-1}(s)$. The result for z is derived in a similar manner. The only difference is that the Ito product rule in Lemma A.1 is used for $\Phi(t)g_2(t)$, where $g_2(t) := x_0 + \int_0^t \Phi^{-1}(s)Bu(s)ds$. In this particular case the Ito formula coincides with the standard calculus since g_2 is an Ito process with a zero diffusion term. Again, it is exploited that $\Phi(t, s) = \Phi(t)\Phi^{-1}(s)$ in the stochastic case, too. We refer, e.g., to [8, 11] for a detailed proof. □

The fundamental solution Φ has an interesting property that is very helpful in the error bound analysis. This property is formulated in the next theorem.

Theorem 2.2. *Let Φ be the fundamental solution of the stochastic system defined in (6) and K be a symmetric postive semidefinite matrix. Then, the following relation holds:*

$$\mathbb{E}\left[\Phi(t, s)K\Phi^T(t, s)\right] = \mathbb{E}\left[\Phi(t - s)K\Phi^T(t - s)\right]. \tag{7}$$

Proof. The result is proved by showing that both sides in (7) satisfy the same uniquely solvable differential equation. We refer to [8] for more details. □

Since Φ_u depends on the control u, a property as in Theorem 2.2 cannot be achieved for Φ_u in general. This makes the error bound analysis for bilinear systems way harder than the one for stochastic systems. That is why we will prove a connection between $x_s(t, x_0, 0)$ and $z_s(t, x_0, 0)$ in the next section. We are then able to bound Φ_u by Φ. Exploiting Theorems 2.1 and 2.2, we subsequently find the desired error bound for bilinear systems.

3 An $\mathcal{H}_2$-bound and asymptotic stability for bilinear systems

In this section, we establish the connection between the state variable in (1) and the solution to (4). We find an output bound depending on the $\mathcal{H}_2$-norm, present an improved characterization of reachability and analyze asymptotic stability for bilinear systems through this new relation.

Given two matrices K and L, we write $K \leq L$ below if $L - K$ is a symmetric positive semidefinite matrix. Moreover, we introduce the vector of control components with a non-zero N_k, $k \in \{1, \ldots, m\}$, by

$$u^0 = (u_1^0, \ldots, u_m^0)^T \quad \text{with} \quad u_k^0 \equiv \begin{cases} 0, & \text{if } N_k = 0 \\ u_k, & \text{else.} \end{cases} \tag{8}$$

We start with three vital Lemmas since they are the basis for every result given here. Their proofs are moved to the appendix in order to improve the readability of the paper.

Lemma 3.1. *Let $x_s(t, x_0, 0)$, $t \geq s \geq 0$, denote the solution to the homogeneous bilinear equation*

$$\dot{x}(t) = Ax(t) + \sum_{k=1}^{m} N_k x(t) u_k(t), \quad x(s) = x_0. \tag{9}$$

Then, the function $x_s(t, x_0, 0)x_s^T(t, x_0, 0)$, $t \geq s \geq 0$, satisfies the following matrix differential inequality:

$$\dot{X}(t) \leq AX(t) + X(t)A^T + \sum_{k=1}^{m} N_k X(t) N_k^T + X(t) \left\| u^0(t) \right\|_2^2, \tag{10}$$

where $X(s) = x_0 x_0^T$.

Proof. See Appendix C.1. □

Lemma 3.2. *Let $z_s(t, x_0, 0)$, $t \geq s \geq 0$, be the solution to the homogeneous stochastic linear equation*

$$dz(t) = Az(t)dt + \sum_{k=1}^{m} N_k z(t) dw_k(t), \quad z(s) = x_0. \tag{11}$$

Then, the function $\exp\left\{\int_s^t \left\| u^0(s) \right\|_2^2 ds\right\} \mathbb{E}\left[z_s(t, x_0, 0) z_s^T(t, x_0, 0)\right]$, $t \geq s \geq 0$, satisfies the following matrix differential equation:

$$\dot{Z}(t) = AZ(t) + Z(t)A^T + \sum_{k=1}^{m} N_k Z(t) N_k^T + Z(t) \left\| u^0(t) \right\|_2^2, \tag{12}$$

where $Z(s) = x_0 x_0^T$.

Proof. See Appendix C.2. □

Lemma 3.3. *Let the matrix-valued function $X(t)$, $t \geq s \geq 0$, satisfy (10) and let $Z(t)$, $t \geq s \geq 0$, be the solution to the matrix differential equation (12). If $X(s) \leq Z(s)$, we have that $X(t) \leq Z(t)$ for all $t \geq s \geq 0$.*

Proof. See Appendix C.3. □

From Lemmas 3.1, 3.2 and 3.3 it immediately follows that

$$x_s(t, x_0, 0)x_s^T(t, x_0, 0) \leq \exp\left\{\int_s^t \left\|u^0(\tau)\right\|_2^2 d\tau\right\} \mathbb{E}\left[z_s(t, x_0, 0)z_s^T(t, x_0, 0)\right] \tag{13}$$

for all $t \geq s \geq 0$. This is exactly the connection we need to establish a bound for the bilinear system. Before we formulate this result, we state another lemma that characterizes mean square asymptotic stability of (4).

Lemma 3.4. *The following two statements are equivalent:*

- *The homogeneous equation (11) with initial time $s = 0$ is exponentially mean square stable, that is, there exist $k_1, k_2 > 0$, such that*

$$\mathbb{E}\left\|z(t, x_0, 0)\right\|_2^2 \leq \left\|x_0\right\|_2^2 k_1 \,\mathrm{e}^{-k_2 t}. \tag{14}$$

- *The eigenvalues of $A \otimes I + I \otimes A + \sum_{k=1}^m N_k \otimes N_k$ have negative real parts only, i.e.,*

$$\sigma(A \otimes I + I \otimes A + \sum_{k=1}^m N_k \otimes N_k) \subset \mathbb{C}_-. \tag{15}$$

Moreover, $\sigma(A) \subset \mathbb{C}_-$ and

$$\left\| \int_0^\infty \mathrm{e}^{At} \left(\sum_{k=1}^m N_k N_k^T\right) \mathrm{e}^{A^T t}\, dt \right\|_2 < 1 \tag{16}$$

imply (15).

Proof. The equivalence of (14) and (15) is a well-known result for stochastic systems. A proof can, e.g., be found in [11, 18, 22]. That $\sigma(A) \subset \mathbb{C}_-$ and (16) are sufficient for exponentially mean square stability is proved in [11, 17]. □

We are now able to establish one of our main results of this paper.

Theorem 3.5. *Let y be the output of system (1) with $x_0 = 0$ and suppose that*

$$\sigma(A \otimes I + I \otimes A + \sum_{k=1}^m N_k \otimes N_k) \subset \mathbb{C}_-. \tag{17}$$

Then, it holds that

$$\sup_{t\geq 0} \|y(t)\|_2 \leq \left(\mathrm{tr}(CPC^T)\right)^{\frac{1}{2}} \exp\left\{0.5 \left\|u^0\right\|_{L^2}^2\right\} \|u\|_{L^2},$$

where $P := \mathbb{E}\int_0^\infty \Phi(t)BB^T\Phi^T(t)dt$. *Moreover,* P *is the solution to*

$$AP + PA^T + \sum_{k=1}^m N_k P N_k^T = -BB^T. \tag{18}$$

Proof. Let $y(t) = Cx(t, 0, B)$ be the output of (1) with zero initial state. Then, plugging in the representation from Theorem 2.1 yields

$$\begin{aligned} \|y(t)\|_2 &= \left\|C\int_0^t \Phi_u(t,s)Bu(s)ds\right\|_2 \leq \int_0^t \|C\Phi_u(t,s)Bu(s)\|_2\, ds \\ &\leq \int_0^t \|C\Phi_u(t,s)B\|_F \|u(s)\|_2\, ds \leq \left(\int_0^t \|C\Phi_u(t,s)B\|_F^2\, ds\right)^{\frac{1}{2}} \left(\int_0^t \|u(s)\|_2^2\, ds\right)^{\frac{1}{2}}. \end{aligned} \tag{19}$$

We further analyze the term $\int_0^t \|C\Phi_u(t,s)B\|_F^2\, ds$. We partition the input matrix by $B = [b_1, b_2, \dots, b_m]$, where b_k is the kth column of B. We have

$$\begin{aligned} \Phi_u(t,s)B &= [\Phi_u(t,s)b_1, \Phi_u(t,s)b_2, \dots, \Phi_u(t,s)b_m] \\ &= [x_s(t, b_1, 0), x_s(t, b_2, 0), \dots, x_s(t, b_m, 0)]. \end{aligned} \tag{20}$$

As mentioned before, Lemmas 3.1, 3.2 and 3.3 provide (13). Through (20) and (13), we obtain

$$\begin{aligned} \Phi_u(t,s)BB^T\Phi_u^T(t,s) &= \sum_{k=1}^m x_s(t, b_k, 0)x_s^T(t, b_k, 0) \\ &\leq \exp\left\{\int_s^t \left\|u^0(\tau)\right\|_2^2 d\tau\right\} \sum_{k=1}^m \mathbb{E}\left[z_s(t, b_k, 0)z_s^T(t, b_k, 0)\right] \\ &= \exp\left\{\int_s^t \left\|u^0(\tau)\right\|_2^2 d\tau\right\} \sum_{k=1}^m \mathbb{E}\left[\Phi(t,s)b_k b_k^T\Phi^T(t,s)\right] \\ &= \exp\left\{\int_s^t \left\|u^0(\tau)\right\|_2^2 d\tau\right\} \mathbb{E}\left[\Phi(t,s)BB^T\Phi^T(t,s)\right]. \end{aligned}$$

With Theorem 2.2 and $\int_s^t \left\|u^0(\tau)\right\|_2^2 d\tau \leq \int_0^t \left\|u^0(\tau)\right\|_2^2 d\tau$, we find

$$\Phi_u(t,s)BB^T\Phi_u^T(t,s) \leq \exp\left\{\int_0^t \left\|u^0(\tau)\right\|_2^2 d\tau\right\} \mathbb{E}\left[\Phi(t-s)BB^T\Phi^T(t-s)\right].$$

This estimated leads to

$$\begin{aligned}
\int_0^t \|C\Phi_u(t,s)B\|_F^2\, ds &= \int_0^t \operatorname{tr}(C\Phi_u(t,s)BB^T\Phi_u^T(t,s)C^T)ds \\
&\leq \exp\left\{\int_0^t \left\|u^0(s)\right\|_2^2 ds\right\} \mathbb{E}\int_0^t \operatorname{tr}(C\Phi(t-s)BB^T\Phi^T(t-s)C^T)ds \\
&= \exp\left\{\int_0^t \left\|u^0(s)\right\|_2^2 ds\right\} \operatorname{tr}(C\,\mathbb{E}\int_0^t \Phi(s)BB^T\Phi^T(s)ds\, C^T)
\end{aligned}$$

using the linearity of the trace operator and substitution $s \mapsto t-s$. We define $P_t := \mathbb{E}\int_0^t \Phi(s)BB^T\Phi^T(s)ds$ and insert the above result into (19) which yields

$$\|y(t)\|_2 \leq \exp\left\{0.5\int_0^t \left\|u^0(s)\right\|_2^2 ds\right\} \left(\operatorname{tr}(CP_tC^T)\right)^{\frac{1}{2}} \left(\int_0^t \|u(s)\|_2^2\, ds\right)^{\frac{1}{2}}. \tag{21}$$

Due to (17) using Lemma 3.4, the fundamental solution Φ decays exponentially such that $P = \mathbb{E}\int_0^\infty \Phi(s)BB^T\Phi^T(s)ds$ exists. That P is the solution to (18) is a well-known result, see, e.g., [8, 11]. Now, taking the supremum on both sides of (21), the claim follows. □

Remark 1. *It was shown in [28] that the term entering the bound in Theorem 3.5 is nothing but the Gramian based representation of the* $\mathcal{H}_2-$ *norm of system (1), i.e.,* $\|\Sigma\|_{\mathcal{H}_2}^2 = \operatorname{tr}(CPC^T)$. *When we choose* $N_k = 0$ *for all* $k = 1,\ldots,m$, *then the exponential term in the bound becomes* 1 *and hence we obtain the well-known relation between the output and the* $\mathcal{H}_2$*-norm in the linear case [15].*

Condition (17) is needed to guarantee the existence of the Gramian P. *However, it can be weakened to* $\sigma(A) \subset \mathbb{C}_-$, *since the bilinear state equation can be equivalently rewritten as*

$$\dot{x}(t) = Ax(t) + [\frac{1}{\gamma}B][\gamma u(t)] + \sum_{k=1}^m [\frac{1}{\gamma}N_k]x(t)[\gamma u_k(t)], \tag{22}$$

see also[5, 10], where the weighted matrices $\tilde{N}_k = \frac{1}{\gamma}N_k$ *can be made arbitrary small with a sufficiently large constant* $\gamma > 0$. *Now, we see that we have*

$$\tilde{f}(A,\tilde{N}_k) := \left\|\int_0^\infty \mathrm{e}^{At}\left(\sum_{k=1}^m \tilde{N}_k\tilde{N}_k^T\right)\mathrm{e}^{A^Tt}\,dt\right\|_2 \leq \frac{1}{\gamma^2}\sum_{k=1}^m \|N_k\|_2^2 \int_0^\infty \left\|\mathrm{e}^{At}\right\|_2^2 dt.$$

$\mathcal{I} := \int_0^\infty \left\|\mathrm{e}^{At}\right\|_2^2 dt$ *is finite because* A *is Hurwitz, such that choosing* $\gamma > \sqrt{\sum_{k=1}^m \|N_k\|_2^2\,\mathcal{I}}$ *leads to* $\tilde{f}(A,\tilde{N}_k) < 1$. *This, by Lemma 3.4, implies*

$$\sigma(A\otimes I + I\otimes A + \sum_{k=1}^m \tilde{N}_k\otimes\tilde{N}_k) \subset \mathbb{C}_-. \tag{23}$$

Then, Theorem 3.5 can be applied to (22) and we get

$$\sup_{t\geq 0} \|y(t)\|_2 \leq \gamma \left(\operatorname{tr}(CP_\gamma C^T)\right)^{\frac{1}{2}} \exp\left\{0.5\gamma^2 \left\|u^0\right\|_{L^2}^2\right\} \|u\|_{L^2},$$

where P_γ solves

$$AP_\gamma + P_\gamma A^T + \frac{1}{\gamma^2}\sum_{k=1}^{m} N_k P_\gamma N_k^T = -\frac{1}{\gamma^2}BB^T.$$

We see that the rescaling makes the bound potentially large for $\gamma \gg 1$ due to the exponential term.

We can derive an inequality from Theorem 3.5 that can be used to characterize reachability in the bilinear system. It leads to an improved characterization in comparison to [5, 14], where energy estimates are shown that hold only in a small neighborhood of zero.

Corollary 3.6. *Let $x(t, 0, B)$, $t \geq 0$, be the solution to the state equation in (1) with $x_0 = 0$ and suppose that (17) holds. Let P be the solution to (18) and $(p_k)_{k=1,\dots,n}$ be an orthonormal basis of eigenvectors of P with corresponding eigenvalues $(\lambda_k)_{k=1,\dots,n}$. Then, it holds that*

$$\sup_{t\geq 0} |\langle x(t, 0, B), p_k\rangle_2| \leq \lambda_k^{\frac{1}{2}} \exp\left\{0.5 \left\|u^0\right\|_{L^2}^2\right\} \|u\|_{L^2}.$$

Proof. We set $C = p_k^T$ in Theorem 3.5. We then obtain $y(t) = Cx(t, 0, B) = \langle x(t, 0, B), p_k\rangle_2$ and $\operatorname{tr}(CPC^T) = \lambda_k$ since the eigenvectors are orthonormal. □

Using an orthonormal basis $(p_k)_{k=1,\dots,n}$ of eigenvectors of P, we can write

$$x(t, 0, B) = \sum_{k=1}^{n} \langle x(t, 0, B), p_k\rangle_2 \, p_k.$$

If the control u is not too large and if the eigenvalue λ_k corresponding to p_k is small, then the Fourier coefficient $\langle x(\cdot, 0, B), p_k\rangle_2$ is close to zero according to Corollary 3.6. This means that the state variable takes only very small values in the direction of p_k. States with a larger component in this direction are not relevant in this case. In order to reach a state with a large component in the eigenspaces of P belonging to the small eigenvalues, a larger control needs to be used. A similar estimate as in Corollary 3.6 has already been obtained for a different reachability Gramian [20, 21].

Based on the result in Theorem 3.5, a bound for the output error between systems (1) and (2) is derived now.

Corollary 3.7. *Let y be the output of system (1) with $x_0 = 0$ satisfying (17). Moreover, let $\hat{y}$ be the output of the reduced system (2) with $\hat{x}_0 = 0$ and*

$$\sigma(\hat{A} \otimes I + I \otimes \hat{A} + \sum_{k=1}^{m} \hat{N}_k \otimes \hat{N}_k) \subset \mathbb{C}_-. \tag{24}$$

Then, we have

$$\sup_{t\geq 0} \|y(t) - \hat{y}(t)\|_2 \leq \left(\operatorname{tr}(CPC^T) + \operatorname{tr}(\hat{C}\hat{P}\hat{C}^T) - 2\operatorname{tr}(CP_g\hat{C}^T)\right)^{\frac{1}{2}} \exp\left\{0.5 \left\|u^0\right\|_{L^2}^2\right\} \|u\|_{L^2}, \tag{25}$$

where P solves (18) and $\hat{P}$, P_g are the solutions to

$$\hat{A}\hat{P} + \hat{P}\hat{A}^T + \sum_{k=1}^{m} \hat{N}_k \hat{P} \hat{N}_k^T = -\hat{B}\hat{B}^T, \tag{26}$$

$$AP_g + P_g\hat{A}^T + \sum_{k=1}^{m} N_k P_g \hat{N}_k^T = -B\hat{B}^T. \tag{27}$$

Proof. We define the error state x^e and the error matrices (A^e, B^e, C^e, N_k^e) as follows:

$$x^e = \begin{bmatrix} x \\ \hat{x} \end{bmatrix}, \; A^e = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix}, \; B^e = \begin{bmatrix} B \\ \hat{B} \end{bmatrix}, \; C^e = \begin{bmatrix} C & -\hat{C} \end{bmatrix}, \; N_k^e = \begin{bmatrix} N_k & 0 \\ 0 & \hat{N}_k \end{bmatrix}.$$

It is not hard to see that x^e satisfies the state equation of system

$$\begin{aligned} \dot{x}^e(t) &= A^e x^e(t) + B^e u(t) + \sum_{k=1}^{m} N_k^e x^e(t) u_k(t), \quad x^e(0) = 0, \\ y^e(t) &= C^e x^e(t), \quad t \geq 0, \end{aligned} \tag{28}$$

and the corresponding output y^e coincides with the output error between (1) and (2), i.e., $y^e = y - \hat{y}$. We need to make sure that the reachability Gramian $P^e = \mathbb{E}\int_0^\infty \Phi^e(t) B^e (B^e)^T (\Phi^e)^T(t) dt$ exists, where $\Phi^e = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}$ is the fundamental solution of the corresponding stochastic system, i.e., it solves

$$\Phi^e(t) = I + \int_0^t A^e \Phi^e(t) ds + \sum_{k=1}^{m} \int_0^t N_k^e \Phi^e(t) dw_k(t). \tag{29}$$

Evaluating the left upper block of (29), we see that $\Phi_{11} = \Phi$ is the fundamental solution of the full system which decays exponentially due to (17). Considering the right upper block of the above identity, the columns of Φ_{12} solve (11) with zero initial state and hence $\Phi_{12} = 0$. With a similar argument, $\Phi_{21} = 0$ is obtained. From the right lower block, we see that $\Phi_{22} = \hat{\Phi}$ is the fundamental solution to the reduced system. (24)

implies the exponential decay of $\hat{\Phi}$ using Lemma 3.4. This gives us the existence of $P^e = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix}$ which according to Theorem 3.5 satisfies

$$A^e P^e + P^e (A^e)^T + \sum_{k=1}^{m} N_k^e P^e (N_k^e)^T = -B^e (B^e)^T. \tag{30}$$

Evaluating the respective blocks of (30), we find $P_{11} = P$, $P_{12} = P_g$ and $P_{22} = \hat{P}$. We can now apply Theorem 3.5 to (28) and obtain

$$\sup_{t\geq 0} \|y^e(t)\|_2 \leq \left(\operatorname{tr}(C^e P^e (C^e)^T)\right)^{\frac{1}{2}} \exp\left\{0.5 \left\|u^0\right\|_{L^2}^2\right\} \|u\|_{L^2}.$$

Using the partitions of C^e and P^e, we have $\operatorname{tr}(C^e P^e (C^e)^T) = \operatorname{tr}(CPC^T) + \operatorname{tr}(\hat{C}\hat{P}\hat{C}^T) - 2\operatorname{tr}(CP_g\hat{C}^T)$. This concludes the proof. □

The trace expression in (25) is a representation of the $\mathcal{H}_2$-error between systems (1) and (2), i.e., $\left\|\Sigma - \hat{\Sigma}\right\|_{\mathcal{H}_2}^2 = \operatorname{tr}(CPC^T) + \operatorname{tr}(\hat{C}\hat{P}\hat{C}^T) - 2\operatorname{tr}(CP_g\hat{C}^T)$. Notice that the bound in (25) can potentially be large due to the exponential term if the control energy is large. This can, e.g., happen if the original system has to be rescaled with a constant γ (see Remark 1) in order to guarantee (17). We can also see from Corollary 3.7 that control components u_k with $N_k \neq 0$ have a much larger impact on the bound because their energy enters exponentially. Later we will discuss balancing related MOR schemes and prove their error bounds based on (25). For those methods (24) usually follows automatically from (17).

We conclude this section with a new result on asymptotic stability for bilinear systems which is a consequence of Lemmas 3.1, 3.2 and 3.3 again.

Theorem 3.8. *Let $x(t, x_0, 0)$, $t \geq 0$, denote the solution to the homogeneous bilinear equation*

$$\dot{x}(t) = Ax(t) + \sum_{k=1}^{m} N_k x(t) u_k(t), \quad x(0) = x_0. \tag{31}$$

If $\sigma(A) \subset \mathbb{C}_-$, then there exit $\gamma, k_1, k_2 > 0$ such that

$$\|x(t, x_0, 0)\|_2^2 \leq \exp\left\{\gamma^2 \left\|u^0\right\|_{L^2}^2\right\} \|x_0\|_2^2 \, k_1 \, \mathrm{e}^{-k_2 t}$$

for all $u \in L^2$, i.e., the bilinear equation is asymptotically stable with exponential decay.

Proof. We can equivalently rewrite equation (31) as

$$\dot{x}(t) = Ax(t) + \sum_{k=1}^{m} [\frac{1}{\gamma} N_k] x(t) [\gamma u_k(t)]$$

as explained in Remark 1. We set $\tilde{N}_k := \frac{1}{\gamma} N_k$. We define the corresponding stochastic equation by

$$d\tilde{z}(t) = A\tilde{z}(t)dt + \sum_{k=1}^{m} \tilde{N}_k \tilde{z}(t) dw_k(t), \quad \tilde{z}(0) = x_0,$$

and denote its solution by $\tilde{z}(t, x_0, 0)$. Lemmas 3.1, 3.2 and 3.3 imply a relation between x and $\tilde{z}$ $(s = 0)$ given in (13), only u is replaced by the rescaled input γu. Applying the trace operator to both sides of (13), the inequality is preserved and we obtain

$$x^T(t, x_0, 0)x(t, x_0, 0) \leq \exp\left\{\gamma^2 \int_0^t \left\|u^0(\tau)\right\|_2^2 d\tau\right\} \mathbb{E}\left[\tilde{z}^T(t, x_0, 0)\tilde{z}(t, x_0, 0)\right].$$

We enlarge the right side of this inequality through replacing $\int_0^t \left\|u^0(\tau)\right\|_2^2 d\tau$ by $\left\|u^0\right\|_{L_2}^2$. Now we can choose $\gamma > \sqrt{\sum_{k=1}^m \|N_k\|_2^2 \int_0^\infty \|\mathrm{e}^{At}\|_2^2 \, dt}$. According to Remark 1 this implies (23). By Lemma 3.4, we therefore know that there exist $k_1, k_2 > 0$, such that

$$\mathbb{E}\left[\tilde{z}^T(t, x_0, 0)\tilde{z}(t, x_0, 0)\right] \leq \|x_0\|_2^2 \, k_1 \, \mathrm{e}^{-k_2 t}.$$

This concludes the proof. □

It is very interesting to notice that adding a bilinearity to an asymptotically stable linear system $\dot{x}(t) = Ax(t)$ preserves this stability condition. The additional bilinear term only enlarges the constant but does not change the decay.

4 Consequences of the $\mathcal{H}_2$-bound in the context of MOR

The bound in Corollary 3.7 has a lot of consequences for MOR schemes applied to bilinear systems. Based on this bound, we are able to explain why $\mathcal{H}_2$-optimal MOR techniques lead to small output errors. Moreover, we can prove output error bounds for both BT and SPA which allow us to point out the cases in which both methods yield a good approximation. These error bounds are a derived from known results in the error analysis for stochastic systems through the link that Corollary 3.7 provides. Throughout this section, we will assume (17). We know that this is achieved if $\sigma(A) \subset \mathbb{C}_-$ and the bilinear system is rescaled according to Remark 1. Moreover, we assume (24) if it is not automatically given through (17). This then guarantees existence of the bound in Corollary 3.7.

4.1 $\mathcal{H}_2$-optimal MOR

The error bound in Corollary 3.7 shows that we can expect a small output error if we find a reduced system that leads to a small $\mathcal{E}^2 := \operatorname{tr}(CPC^T) + \operatorname{tr}(\hat{C}\hat{P}\hat{C}^T) - 2\operatorname{tr}(CP_g\hat{C}^T)$ (in case the control u is not too large). Consequently, one can expect a good ROM when $\mathcal{E}$ is minimized with respect to $\hat{A}, \hat{B}, \hat{C}$ and $\hat{N}_k$. As mentioned in the previous section,

Algorithm 1 Bilinear IRKA

Input: The system matrices: A, B, C, N_k.

Output: The reduced matrices: $\hat{A}, \hat{B}, \hat{C}, \hat{N}_k$.

1: Make an initial guess for the reduced matrices $\hat{A}, \hat{B}, \hat{C}, \hat{N}_k$.

2: **while** not converged **do**

3: Perform the spectral decomposition of $\hat{A}$ and define:
$D = S\hat{A}S^{-1},\ \tilde{B} = S\hat{B},\ \tilde{C} = \hat{C}S^{-1},\ \tilde{N}_k = S\hat{N}_kS^{-1}$.

4: Solve for V and W:
$-VD - AV - \sum_{k=1}^{m} N_k V \tilde{N}_k^T = B\tilde{B}^T$,
$-WD - A^TW - \sum_{k=1}^{m} N_k^T W \tilde{N}_k = C^T\tilde{C}$.

5: $V = \operatorname{orth}(V)$ and $W = \operatorname{orth}(W)$, where $\operatorname{orth}(\cdot)$ returns an orthonormal basis for the range of a matrix.

6: Determine the reduced matrices:
$\hat{A} = (W^TV)^{-1}W^TAV, \qquad \hat{B} = (W^TV)^{-1}W^TB, \qquad \hat{C} = CV.$

7: **end while**

$\mathcal{E}$ is nothing but the $\mathcal{H}_2$-error between systems (1) and (2). Necessary conditions for a locally optimal $\mathcal{H}_2$-error have already been provided [28]. These are

$$\hat{C}\hat{P} = CP_g, \quad \hat{Q}\hat{B} = Q_gB, \quad \hat{Q}\hat{P} = Q_gP_g, \quad \hat{Q}\hat{N}_k\hat{P} = Q_gN_kP_g, \tag{32}$$

where $\hat{P}, P_g$ are the solutions to (26) and (27). Moreover, $\hat{Q}, Q_g$ satisfy

$$\hat{A}^T\hat{Q} + \hat{Q}\hat{A} + \sum_{k=1}^{m} \hat{N}_k^T\hat{Q}\hat{N}_k = -\hat{C}^T\hat{C},$$
$$\hat{A}^TQ_g + Q_gA + \sum_{k=1}^{m} \hat{N}_k^TQ_gN_k = -\hat{C}^TC.$$

Through Corollary 3.7 it is now clear that choosing reduced systems (2) satisfying (32) is meaningful in terms of the output error. This is new insight since the link between the output and the $\mathcal{H}_2$-error was not previously known. Such $\mathcal{H}_2$-optimal ROMs are, e.g., derived through generalized Sylvester iterations, see Algorithm 1 in [4]. Another very famous representative of $\mathcal{H}_2$-optimal schemes is the bilinear iterative rational Krylov algorithm (IRKA), see Algorithm 1. Due to a reformulation of (32), it could be shown in [4] that bilinear IRKA satisfies the necessary optimality conditions.

4.2 Error bounds for balancing related MOR techniques applied to bilinear systems

We explain the procedure of balancing related MOR first, before we show the error bounds for two particular methods. These are BT and SPA. States that require a lager amount of energy to be reached (hard to reach states) can be identified through the

Gramian P solving (18) using Corollary 3.6. We refer to [5, 14] for alternative characterizations based on P and to [20, 21] for estimates based on an alternative reachability Gramian. States that produce only a small amount of observation energy (hard to observe states) can be found through an observability Gramian Q [5, 14, 20] satisfying

$$A^T Q + QA + \sum_{k=1}^{m} N_k^T Q N_k = -C^T C. \tag{33}$$

The goal is to remove the hard to reach and observe states that are contained in the eigenspaces of P and Q, respectively, corresponding to the small eigenvalues. This is done by simultaneously diagonalizing P and Q such that they are equal and diagonal. Subsequently, the states contributing only very little to the systems dynamics are neglected.

In detail, the procedure works as as follows. Assuming $P, Q > 0$, we choose a regular state space transformation $S \in \mathbb{R}^n$ given by

$$S = \Sigma^{-\frac{1}{2}} U^T L^T \quad \text{and} \quad S^{-1} = KV\Sigma^{-\frac{1}{2}}, \tag{34}$$

where $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n) > 0$ with diagonal entries being the square root of eigenvalues of PQ. These diagonal entries are called Hankel singular values (HSVs). The other ingredients of the transformation S are computed in the following way. Let us factorize $P = KK^T$ and $Q = LL^T$, then a singular value decomposition of $K^T L = V\Sigma U^T$ gives the required matrices. We now introduce a transformed state

$$x_b(t) = Sx(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix},$$

where $x_1(t) \in \mathbb{R}^r$. The transformed state x_b satisfies a bilinear system with the same output as (1) having the following coefficients

$$SAS^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad SB = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad CS^{-1} = \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \quad SN_kS^{-1} = \begin{bmatrix} N_{k,11} & N_{k,12} \\ N_{k,21} & N_{k,22} \end{bmatrix}, \tag{35}$$

where $A_{11} \in \mathbb{R}^{r\times r}$ etc. Using the above partitions, this system is

$$\begin{aligned} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(t) + \sum_{k=1}^{m} \begin{bmatrix} N_{k,11} & N_{k,12} \\ N_{k,21} & N_{k,22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} u_k(t), \\ y(t) &= \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \quad t \geq 0. \end{aligned} \tag{36}$$

The Gramians P_b and Q_b of (36) are

$$P_b = SPS^T = Q_b = S^{-T} Q S^{-1} = \Sigma = \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix}$$

with $\Sigma_1 \in \mathbb{R}^{r\times r}$. $\Sigma_2 = \operatorname{diag}(\sigma_{r+1}, \ldots, \sigma_n)$ contains the $n - r$ smallest HSVs of the systems. The corresponding state variables x_2 are hence less important and can be neglected in the system dynamics since those represent the difficult to reach and observe states in (36). In order to obtain a ROM, the second line in the state equation of (36) is truncated. The remaining x_2 components in the first line of the state equation and in the output equation can now be approximated in two ways. One is setting $x_2(t) = 0$. This method is called BT and the ROM (2) then has coefficients

$$(\hat{A}, \hat{B}, \hat{C}, \hat{N}_k) = (A_{11}, B_1, C_1, N_{k,11}). \tag{37}$$

An alternative method is SPA where one sets $x_2(t) = -A_{22}^{-1}A_{21}x_1(t)$. This results in a ROM with matrices

$$(\hat{A}, \hat{B}, \hat{C}, \hat{N}_k) = (\bar{A}, B_1, \bar{C}, \bar{N}_k), \tag{38}$$

where we define

$$\bar{A} := A_{11} - A_{12}A_{22}^{-1}A_{21}, \quad \bar{C} := C_1 - C_2A_{22}^{-1}A_{21}, \quad \bar{N}_k := N_{k,11} - N_{k,12}A_{22}^{-1}A_{21}.$$

We refer to [16] for more details about SPA. There, the reduced system with matrices as in (38) is derived through an averaging principle. In the following, we present L^∞-error bounds for both BT and SPA. Both results are new and the first ones for balancing related methods based on the $\mathcal{H}_2$-error. However, we want to mention that there is an alternative bound for BT in infinite dimensions [3] and there are L^2-error bounds for BT and SPA based on a different reachability Gramian P_2 [20, 23]. P_2 is defined to be a positive definite solution to

$$A^T P_2^{-1} + P_2^{-1} A + \sum_{k=1}^m N_k^T P_2^{-1} N_k \leq -P_2^{-1} B B^T P_2^{-1}. \tag{39}$$

Replacing P by P_2 is also called type II approach.

For simplicity of the notation, we from now on assume that system (1) is already balanced, i.e., we already applied the balancing transformation in (34) such that $P = Q = \Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_n)$. We formulate the error bound for BT first.

Theorem 4.1 (Error bound BT). *Let (A, B, C, N_k) be a balanced realization of system (1) with partitions as in (35). Suppose that $\hat{y}$ is the output of the reduced system with matrices given in (37). Moreover, we assume that (17) holds. Then, given zero initial conditions for both the full and the reduced system, we have*

$$\sup_{t\geq 0} \|y(t) - \hat{y}(t)\|_2 \leq (\operatorname{tr}(\Sigma_2 K_{BT}))^{\frac{1}{2}} \exp\left\{0.5 \left\|u^0\right\|_{L^2}^2\right\} \|u\|_{L^2}$$

with $\Sigma_2 = \operatorname{diag}(\sigma_{r+1}, \dots, \sigma_n)$ and the weighting matrix given by

$$K_{BT} = B_2 B_2^T + 2P_{g,2}A_{21}^T + \sum_{k=1}^m (2N_{k,22}P_{g,2}N_{k,21}^T + 2N_{k,21}P_{g,1}N_{k,21}^T - N_{k,21}\hat{P}N_{k,21}^T),$$

where $\hat{P}$ and $P_g = \begin{bmatrix} P_{g,1} \\ P_{g,2} \end{bmatrix}$ satisfy (26) and (27), respectively.

Proof. Assumption (17) implies (24) in the case of BT, see [6], i.e., mean square asymptotic stability is preserved in the ROM. Consequently, the bound in Corollary 3.7 exists. So, it holds that

$$\sup_{t\geq 0}\|y(t)-\hat{y}(t)\|_2 \leq \left(\operatorname{tr}(CPC^T)+\operatorname{tr}(\hat{C}\hat{P}\hat{C}^T)-2\operatorname{tr}(CP_g\hat{C}^T)\right)^{\frac{1}{2}} \exp\left\{0.5\left\|u^0\right\|_{L^2}^2\right\}\|u\|_{L^2}.$$

The trace expression in the above estimate has already been analyzed within the error bound analysis of stochastic systems. By [8, Proposition 4.6], we then have

$$\operatorname{tr}(CPC^T)+\operatorname{tr}(\hat{C}\hat{P}\hat{C}^T)-2\operatorname{tr}(CP_g\hat{C}^T)=\operatorname{tr}(\Sigma_2 K_{BT}).$$

□

Setting $N_k = 0$ in Theorem 4.1 leads to the $\mathcal{H}_2$-error bound in the linear case [2]. We now state the bound for SPA.

Theorem 4.2 (Error bound SPA)**.** *Let (A, B, C, N_k) be a balanced realization of system (1) with partitions as in (35). Suppose that $\hat{y}$ is the output of the reduced system with matrices given in (38). Moreover, we assume that (17) and*

$$0 \notin \sigma(\hat{A}\otimes I + I\otimes \hat{A} + \sum_{k=1}^{m}\hat{N}_k\otimes\hat{N}_k) \tag{40}$$

hold. Then, zero initial conditions for both the full and the reduced system yield

$$\sup_{t\geq 0}\|y(t)-\hat{y}(t)\|_2 \leq \left(\operatorname{tr}(\Sigma_2 K_{SPA})\right)^{\frac{1}{2}} \exp\left\{0.5\left\|u^0\right\|_{L^2}^2\right\}\|u\|_{L^2}$$

with $\Sigma_2 = \operatorname{diag}(\sigma_{r+1},\dots,\sigma_n)$ and the weighting matrix given by

$$\begin{aligned} K_{SPA} =& B_2B_2^T - 2(A_{22}P_{g,2}+A_{21}P_{g,1})(A_{22}^{-1}A_{21})^T \\ &+ 2\sum_{k=1}^{m}(N_{k,22}P_{g,2}+N_{k,21}P_{g,1})(N_{k,21}-N_{k,22}A_{22}^{-1}A_{21})^T \\ &- \sum_{k=1}^{m}(N_{k,21}-N_{k,22}A_{22}^{-1}A_{21})\hat{P}(N_{k,21}-N_{k,22}A_{22}^{-1}A_{21})^T, \end{aligned}$$

where $\hat{P}$ and $P_g = \left[\begin{smallmatrix} P_{g,1} \\ P_{g,2}\end{smallmatrix}\right]$ satisfy (26) and (27), respectively.

Proof. Assumptions (17) and (40) yield (24) for SPA, see [25]. This guarantees existence of the bound in Corollary 3.7. Consequently, we have

$$\sup_{t\geq 0}\|y(t)-\hat{y}(t)\|_2 \leq \left(\operatorname{tr}(CPC^T)+\operatorname{tr}(\hat{C}\hat{P}\hat{C}^T)-2\operatorname{tr}(CP_g\hat{C}^T)\right)^{\frac{1}{2}} \exp\left\{0.5\left\|u^0\right\|_{L^2}^2\right\}\|u\|_{L^2}.$$

The above term is known for SPA applied to stochastic systems. By [25, Theorem 4.1], we know that

$$\operatorname{tr}(CPC^T)+\operatorname{tr}(\hat{C}\hat{P}\hat{C}^T)-2\operatorname{tr}(CP_g\hat{C}^T)=\operatorname{tr}(\Sigma_2 K_{SPA}).$$

□

Theorems 4.1 and 4.2 show us in which cases BT and SPA work well. If one only truncates the states corresponding to the small Hankel singular values (hard to reach and observe states), then the diagonal entries $\sigma_{r+1},\dots,\sigma_n$ of Σ_2 and hence the output error are small assuming that the control is not too large.

Remark 2. *Based on the results in [24] and [22] Theorems 4.1 and 4.2 can be formulated the same way if P is replaced by P_2 satisfying (39). In this type II case,(24) automatically follows from (17) for BT and SPA due to [7] and [22]. The reason why (40) has to be assumed above is that stability preservation for SPA based on P is still an open question.*

5 Conclusions

In this paper, we studied bilinear systems in terms of asymptotic stability. Furthermore, we characterized reachability within bilinear equations. Moreover, we found a bound for the output errors of bilinear systems based on the $\mathcal{H}_2$-error. This error bound could finally explain why $\mathcal{H}_2$-optimal model order reduction techniques lead to good approximations. Such a link has been an open question for quite some time. Subsequently, error bounds for balanced truncation and singular perturbation approximation could be derived. These bounds are the first ones for this type of balancing related schemes considered here and they can tell us in which cases the methods perform well.

Acknowledgments

The author thanks Tobias Damm (University of Kaiserslautern) for the fruitful discussion around the proof of Lemma 3.3.

A Ito product rule

Suppose that $w_1,\dots,w_m$ are independent scalar standard Brownian motions and suppose that $(\mathcal{F}_t)_{t\geq 0}$ is the filtration generated by these processes. Let the Brownian motions and all other stochastic processes appearing in this section be defined on the filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$. We introduce vector-valued Ito processes with $t\geq 0$ through

$$\begin{aligned} z(t) &= z(0)+\int_0^t a(s)ds+\sum_{k=1}^m\int_0^t b_k(s)dw_k(s),\\ \tilde{z}(t) &= \tilde{z}(0)+\int_0^t \tilde{a}(s)ds+\sum_{k=1}^m\int_0^t \tilde{b}_k(s)dw_k(s), \end{aligned} \tag{41}$$

where the vector-valued diffusion processes b_k and $\tilde{b}_k$ are $(\mathcal{F}_t)_{t\geq 0}$-adapted and integrable with respect to w_k $(k=1,\dots,m)$. Furthermore, the drift terms a and $\tilde{a}$ are $\mathbb{P}$-almost surely Lebesgue integrable and $(\mathcal{F}_t)_{t\geq 0}$-adapted. Next, we state the well-known Ito product formula for the vector-valued case.

Lemma A.1. *Let z and $\tilde{z}$ be $\mathbb{R}^n$ and $\mathbb{R}^d$-valued Ito processes, respectively, with representations as in (41). Then, we have*

$$d\left(z(t)\tilde{z}^T(t)\right) = dz(t)\tilde{z}^T(t) + z(t)d\tilde{z}^T(t) + dz(t)d\tilde{z}^T(t)$$

for all $t \geq 0$. The compensator process is given through the diffusion terms as follows:

$$dz(t)d\tilde{z}^T(t) = \sum_{k=1}^{m} b_k(t)\tilde{b}_k^T(t)dt.$$

B Resolvent positive operators

Let $(H^n, \langle \cdot, \cdot \rangle_F)$ be the Hilbert space of symmetric $n \times n$ matrices and let H_+^n be the subset of symmetric positive semidefinite matrices. We now define positive and resolvent positive operators on H^n.

Definition B.1. *A linear operator $L : H^n \to H^n$ is called positive if $L(H_+^n) \subset H_+^n$. It is resolvent positive if there is an $\alpha_0 \in \mathbb{R}$ such that for all $\alpha > \alpha_0$ the operator $(\alpha I - L)^{-1}$ is positive.*

The operator $\mathcal{L}(X) := AX + XA^T$ is resolvent positive and $\Pi(X) := \sum_{k=1}^m N_k X N_k^T$ is positive for $A, N_k \in \mathbb{R}^n$. This implies that the generalized Lyapunov operator $\mathcal{L} + \Pi$ is resolvent positive. We refer to [11] for a more detailed discussion and a proof. We now state an equivalent characterization for resolvent positive operators in the following. It can be found in a more general form in [11, 12, 27].

Theorem B.2. *A linear operator $L : H^n \to H^n$ is resolvent positive if and only if $\langle V_1, V_2 \rangle_F = 0$ implies $\langle LV_1, V_2 \rangle_F \geq 0$ for $V_1, V_2 \in H_+^n$.*

C Pending proofs

Notice that the Lemmas 3.1, 3.2 and 3.3 are proved for initial time zero for simplicity of the notation. The proofs are completely analogous for general initial times s.

C.1 Proof of Lemma 3.1

We write $x(t)$ instead of $x(t, x_0, 0)$ to shorten the notation within this proof. We apply the product rule and insert (9). Hence, we obtain

$$\begin{aligned}
\frac{d}{dt}x(t)x^T(t) &= [\frac{d}{dt}x(t)]x^T(t) + x(t)[\frac{d}{dt}x^T(t)] \\
&= \left(Ax(t) + \sum_{k=1}^{m} N_k x(t) u_k(t)\right) x^T(t) + x(t)\left(x^T(t)A^T + \sum_{k=1}^{m} x^T(t) N_k^T u_k(t)\right) \\
&= Ax(t)x^T(t) + x(t)x^T(t)A^T + \sum_{k=1}^{m} \left[N_k x(t)x^T(t)u_k(t) + x(t)x^T(t)N_k^T u_k(t)\right].
\end{aligned}$$

Suppose that $v \in \mathbb{R}^n$ is arbitrary. Then,

$$\sum_{k=1}^{m} v^T \left[N_k x(t) x^T(t) u_k(t) + x(t) x^T(t) N_k^T u_k(t)\right] v = \sum_{k=1}^{m} 2 v^T N_k x(t) x^T(t) u_k(t) v$$
$$\leq \sum_{k=1}^{m} v^T N_k x(t) x^T(t) N_k^T v + v^T x(t) x^T(t) (u_k^0(t))^2 v$$
$$= \sum_{k=1}^{m} \left[v^T N_k x(t) x^T(t) x(t) v\right] + v^T x(t) x^T(t) v \left\|u^0(t)\right\|_2^2.$$

This leads to the claim of this lemma.

C.2 Proof of Lemma 3.2

We write $z(t)$ instead of $z(t, x_0, 0)$ in this proof and moreover define $\tilde{z}(t) := \exp\left\{0.5 \int_0^t \left\|u^0(s)\right\|_2^2 ds\right\} z(t)$. The function $g(t) := \exp\left\{0.5 \int_0^t \left\|u^0(s)\right\|_2^2 ds\right\}$ is an Ito process with diffusion term zero, since it can be written as $g(t) = 1 + 0.5 \int_0^t \left\|u^0(s)\right\|_2^2 g(s) ds$. Consequently, Ito's product rule, see Lemma A.1, is just the classical product rule when applying it to $g(t)z(t)$. We obtain

$$\begin{aligned} d\tilde{z}(t) &= dg(t) z(t) + g(t) dz(t) \\ &= 0.5 \left\|u^0(t)\right\|_2^2 g(t) z(t) dt + g(t) \left(Az(t)dt + \sum_{k=1}^{m} N_k z(t) dw_k(t)\right) \\ &= (A + 0.5 \left\|u^0(t)\right\|_2^2 I) \tilde{z}(t) dt + \sum_{k=1}^{m} N_k \tilde{z}(t) dw_k(t) \end{aligned} \tag{42}$$

inserting equation (11). Knowing the stochastic differential for $\tilde{z}$ now, the Ito product rule in Lemma A.1 is applied to $\tilde{z}(t)\tilde{z}^T(t)$. Hence, we find

$$d\left(\tilde{z}(t)\tilde{z}^T(t)\right) = d\tilde{z}(t)\tilde{z}^T(t) + \tilde{z}(t) d\tilde{z}^T(t) + \sum_{k=1}^{m} N_k \tilde{z}(t)\tilde{z}^T(t) N_k^T dt.$$

We insert (42) and apply the expected value to both sides of the equation. Using the fact that an Ito integral has mean zero, we have

$$\begin{aligned} d\left(\mathbb{E}[\tilde{z}(t)\tilde{z}^T(t)]\right) =& (A + 0.5 \left\|u^0(t)\right\|_2^2 I) \mathbb{E}[\tilde{z}(t)\tilde{z}^T(t)] dt \\ &+ \mathbb{E}[\tilde{z}(t)\tilde{z}^T(t)](A^T + 0.5 \left\|u^0(t)\right\|_2^2 I) dt + \sum_{k=1}^{m} N_k \mathbb{E}[\tilde{z}(t)\tilde{z}^T(t)] N_k^T dt. \end{aligned}$$

This proves the result of this Lemma.

C.3 Proof of Lemma 3.3

We set $Y := Z - X$ and $L(Y(t)) := AY(t) + Y(t)A^T + \sum_{k=1}^m N_k Y(t) N_k^T + Y(t) \left\|u^0(t)\right\|_2^2$. We subtract (10) from (12) and obtain

$$\dot{Y}(t) \geq L(Y(t)).$$

We define the difference function $D(t) := \dot{Y}(t) - L(Y(t)) \geq 0$ and consider the following perturbed differential equation

$$\dot{Y}_\epsilon(t) = L(Y_\epsilon(t)) + D(t) + \epsilon I$$

with parameter $\epsilon \geq 0$ and initial state $Y_\epsilon(0) = Y(0) + \epsilon I$. We see that $Y_0(t) = Y(t)$ for all $t \geq 0$ since $Y_0 - Y$ solves the differential equation $\dot{\tilde{Y}}(t) = L(\tilde{Y}(t))$ with zero initial state. Since Y_ϵ continuously depends on ϵ and the initial data, we have $\lim_{\epsilon\to 0} Y_\epsilon(t) = Y_0(t) = Y(t)$ for all $t \geq 0$.

Let us now assume that Y_ϵ is not positive definite for $\epsilon > 0$. Then, there exist a $\tilde{v} \neq 0$ and a $\tilde{t} > 0$ such that $\tilde{v}^T Y_\epsilon(\tilde{t})\tilde{v} \leq 0$. We know that $f_\epsilon(v,t) := v^T Y_\epsilon(t) v$ is positive at $t = 0$ for all $v \in \mathbb{R}^n \setminus \{0\}$ by assumption. Since f_ϵ is non-positive in some point $(\tilde{v}, \tilde{t})$ and due to the continuity of $t \mapsto Y_\epsilon(t)$, there is a point $t_0 \in (0, \tilde{t}]$ for which

$$v_0^T Y_\epsilon(t_0) v_0 = 0 \quad \text{and} \quad v_0^T Y_\epsilon(t) v_0 > 0, \quad t < t_0, \tag{43}$$

for some $v_0 \neq 0$, whereas $v^T Y_\epsilon(t_0) v \geq 0$ for all other $v \in \mathbb{R}^n$. L is a generalized Lyapunov operator and hence resolvent positive (see Appendix B). The identify $0 = v_0^T Y_\epsilon(t_0) v_0 = \langle Y_\epsilon(t_0), v_0 v_0^T \rangle_F$, by Theorem B.2, then implies $0 \leq \langle L(Y_\epsilon(t_0)), v_0 v_0^T \rangle_F = v_0^T L(Y_\epsilon(t_0)) v_0$. Using these facts, we have

$$\frac{d}{dt} v_0^T Y_\epsilon(t_0) v_0 = v_0^T L(Y_\epsilon(t_0)) v_0 + v_0^T D(t_0) v_0 + \epsilon \left\|v_0\right\|_2^2 > 0.$$

Consequently, we know that there are $t < t_0$ close to t_0 for which $v_0^T Y_\epsilon(t) v_0 < 0$. This contradicts (43) and hence our assumption is wrong such that $Y_\epsilon(t)$ is positive definite for all $t \geq 0$ and $\epsilon > 0$. Taking the limit of $\epsilon \to 0$, we obtain $Y(t) \geq 0$ for all $t \geq 0$ which concludes the proof.

References

[1] S. A. Al-Baiyat and M. Bettayeb. A new model reduction scheme for k–power bilinear systems. *Proceedings of the 32nd IEEE Conference on Decision and Control*, pages 22–27, 1993.

[2] A. C. Antoulas. *Approximation of Large-Scale Dynamical Systems*, volume 6 of *Adv. Des. Control*. SIAM Publications, Philadelphia, PA, 2005.

[3] S. Becker and C. Hartmann. Infinite-dimensional bilinear and stochastic balanced truncation with explicit error bounds. *Mathematics of Control, Signals, and Systems*, 31(2):1–37, 2019.

[4] P. Benner and T. Breiten. Interpolation-based $\mathcal{H}_2$-model reduction of bilinear control systems. *SIAM J. Matrix Anal. Appl.*, 33(3):859–885, 2012.

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

[6] P. Benner, T. Damm, M. Redmann, and Y. R. Rodriguez Cruz. Positive Operators and Stable Truncation. *Linear Algebra and its Applications*, 498:74–87, 2016.

[7] P. Benner, T. Damm, and Y. R. Rodriguez Cruz. Dual pairs of generalized Lyapunov inequalities and balanced truncation of stochastic linear systems. *IEEE Trans. Autom. Contr.*, 62(2):782–791, 2017.

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

[9] C. Bruni, G. DiPillo, and G. Koch. On the mathematical models of bilinear systems. *Automatica*, 2(1):11–26, 1971.

[10] M. Condon and R. Ivanov. Nonlinear systems-algebraic gramians and model reduction. *COMPEL*, 24(1):202–219, 2005.

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

[12] L. Elsner. Quasimonotonie und Ungleichungen in halbgeordneten Räumen. *Linear Algebra and Its Applications*, 8(3):249–261, 1974.

[13] G. Flagg and S. Gugercin. Multipoint Volterra series interpolation and $\mathcal{H}_2$ optimal model reduction of bilinear systems. *SIAM J. Matrix Anal. Appl.*, 36(2):549–579, 2015.

[14] W. S. Gray and J. Mesko. Energy functions and algebraic Gramians for bilinear systems. In *Preprints of the 4th IFAC Nonlinear Control Systems Design Symposium*, pages 103–108, Enschede, The Netherlands, 1998.

[15] S. Gugercin, A. C. Antoulas, and C. A. Beattie. $\mathcal{H}_2$ model reduction for large-scale dynamical systems. *SIAM J. Matrix Anal. Appl.*, 30(2):609–638, 2008.

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

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

[18] R.Z. Khasminskii. Stochastic stability of differential equations. Monographs and Textbooks on Mechanics of Solids and Fluids. Mechanics: Analysis, 7. Alphen aan den Rijn, The Netherlands; Rockville, Maryland, USA: Sijthoff & Noordhoff., 1980.

[19] R. R. Mohler. *Bilinear Control Processes*. Academic Press, New York, 1973.

[20] M. Redmann. Energy estimates and model order reduction for stochastic bilinear systems. *International Journal of Control*, 2018.

[21] M. Redmann. Type II balanced truncation for deterministic bilinear control systems. *SIAM J. Control Optim.*, 56(4):2593–2612, 2018.

[22] M. Redmann. Type II singular perturbation approximation for linear systems with Lévy noise. *SIAM J. Control Optim.*, 56(3):2120–2158., 2018.

[23] M. Redmann. A new type of singular perturbation approximation for stochastic bilinear systems. *arXiv preprint:1903.11600*, 2019.

[24] M. Redmann and P. Benner. An H_2-Type Error Bound for Balancing-Related Model Order Reduction of Linear Systems with Lévy Noise. *Systems and Control Letters*, 105:1–5, 2017.

[25] M. Redmann and P. Benner. Singular Perturbation Approximation for Linear Systems with Lévy Noise. *Stochastics and Dynamics*, 18(4), 2018.

[26] W. J. Rugh. *Nonlinear System Theory*. The Johns Hopkins University Press, Baltimore, MD, 1981.

[27] H. Schneider and M. Vidyasagar. Cross-Positive Matrices. *SIAM Journal on Numerical Analysis*, 7(4):508–519, 1970.

[28] L. Zhang and J. Lam. On H_2 model reduction of bilinear systems. *Automatica*, 38(2):205–216, 2002.