Error bounds in the gap metric for dissipative balanced approximations

Chris Guiver[*†] and Mark R. Opmeer[‡]

Linear Algebra and its Applications (2013)
`http://dx.doi.org/10.1016/j.laa.2013.09.032`

Abstract

We derive an error bound in the gap metric for positive real balanced truncation and positive real singular perturbation approximation. We prove these results by working in the context of dissipative driving–variable systems, as in behavioral and state/signal systems theory. In such a framework no prior distinction is made between inputs and outputs. Dissipativity preserving balanced truncation of dissipative driving–variable systems is addressed and a gap metric error bound is obtained. Bounded real and positive real input–state–output systems are manifestations of a dissipative driving–variable system through particular decompositions of the signal space. Under such decompositions the existing bounded real and positive real balanced truncation schemes can be seen as special cases of dissipative balanced truncation and the new positive real error bounds follow.

Keywords: model reduction, dissipative system, balanced realisation, singular perturbation approximation, gap metric, bounded real, positive real, driving variable system, KYP Lemma.

Mathematics Subject Classification: 93B11, 93C05, 15A24, 46C20, 47B50.

1 Introduction

Model reduction for control systems refers to replacing a system with many degrees of freedom by one with fewer degrees of freedom. Lyapunov balanced truncation is one such model reduction scheme, introduced by Moore [1]. One of the appeals of Lyapunov balanced truncation is the H^∞ error bound

$$\|G - G_n\|_\infty \leq 2 \sum_{k=n+1}^{N} \sigma_k, \tag{1.1}$$

which was independently derived by Enns [2] and Glover [3]. In (1.1), G and G_n are the transfer functions of the original and truncated systems respectively, with respective orders N and n. The σ_k are the singular values of the Hankel operator of G. The bound (1.1), when combined with the trivial lower bound

$$\sigma_{n+1} \leq \|G - G_n\|_\infty,$$

which holds for any reduced order system of dimension n, shows that Lyapunov balanced truncation is close to optimal.

[*]Environment & Sustainability Institute, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Penryn Campus, TR10 9EZ, United Kingdom. `c.guiver@ex.ac.uk`

[†]corresponding author

[‡]Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom. `m.opmeer@maths.bath.ac.uk`

A downside of Lyapunov balanced truncation is that any dissipativity property of the original system is not necessarily retained in the reduced order system. There are two classical notions of dissipativity in control theory. On the one hand there are the systems called impedance passive, passive or positive real and on the other hand there are the systems called scattering passive, contractive or bounded real. It is well known that these systems are related by a transform known as the Cayley transform, Möbius transform or diagonal transform. These respective notions of dissipativity are preserved by positive real balanced truncation and bounded real balanced truncation introduced by Desai & Pal [4] and Opdenacker & Jonckheere [5] respectively.

The error bound (1.1) holds for bounded real balanced truncation, where σ_k now denote the bounded real singular values and G_n the bounded real balanced truncation. However, the bound (1.1) *does not* hold for positive real balanced truncation (where σ_k denote the positive real singular values), see Example 3.14. There are H^∞ type error bounds for positive real balanced truncation [6], which we recall in Theorem 3.12, but they are somewhat more complicated than (1.1); there is also a false bound in [7]. In this article we prove the gap metric error bound

$$\hat{\delta}(J, J_r) \leq 2 \sum_{k=r+1}^{m} \sigma_k, \tag{1.2}$$

where J is a positive real transfer function, J_r is its positive real balanced truncation and σ_k denote the positive real singular values. The bound (1.2) has been independently established by Timo Reis [8] and as with all error bounds, is useful for simulation purposes. Another advantage of an error bound in the gap metric, however, are the robustness estimates (see, for instance, [9, Chapter 17] and specifically [9, Theorem 17.3] which was originally proven in [10]) and hence its use for combining model reduction with controller design.

To prove the bound (1.2) we take a more conceptual view of classical input–state–output systems and work in the framework of a (dissipative) state/signal system [11], [12]. These systems have the property, amongst others, that no *a priori* distinction is made between inputs and outputs. Instead, an external signal is studied that contains all interactions with the outside environment and which may be decomposed into an input and an output in various ways. It is known that classical bounded real and positive real input–state–output systems appear when specific input–output decompositions are chosen in the signal space of a dissipative state/signal system, Figure 1(a). Such a framework explains more naturally than the Möbius transform why bounded real and positive real input–state–output systems are essentially the same system looked at in different ways.

Dissipativity retaining balanced truncation in such a framework has been studied in [13]–[15], but with a different emphasis to ours. For instance, the error bounds provided there are on the H^∞–norm of the difference of the original and the reduced transfer function, and as such depend on the particular input–output decomposition chosen. The gap metric is a natural metric to consider for the balanced truncations of state/signal systems because it is independent of any input–output decomposition. We prove a new gap metric error bound for dissipative balanced truncations of state/signal systems, Theorem 2.1. We also establish a relation between dissipative balanced truncation and positive real balanced truncation, Figure 1(b). Combining these, the error bound (1.2) for positive real balanced truncation readily follows. As a corollary of (1.2) we obtain a new H^∞ error bound for positive real balanced truncation, Corollary 9.8, which is less interesting than the gap metric error bound, however, as it is not an *a priori* bound.

Singular perturbation approximation of bounded real and positive real input–state–output systems has been considered by Muscato *et al.* [16]. There they show that this model reduction scheme also preserves the respective dissipativity properties and that the balanced truncation error bounds translate across. We demonstrate that singular perturbation approximation is often suitable in our framework and as a consequence we obtain the same gap metric error bound for singular perturbation approximation.

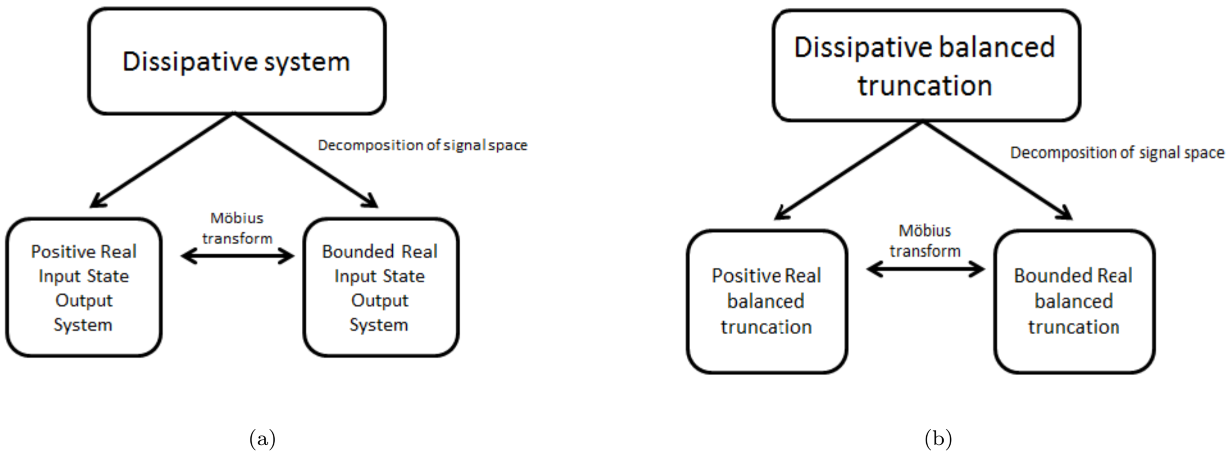

Figure 1.1: Diagram showing relationships between dissipative state/signal systems and classical dissipative input–state–output systems (a) and their respective dissipative balanced truncations (b).

1.1 Organisation of the article and notation

The next section contains an informal overview of the key ideas of the article, statements of our main results and two examples. The technical heart then follows, as we gather the material we require from the three related, but disparate, disciplines of model reduction by balanced truncation, dissipative systems and an input–output free framework. Section 3 briefly reviews model reduction by bounded real and positive real balanced truncation. Many of the concepts in dissipative balanced truncation generalise the notions presented there. Section 4 describes in more detail the systems we consider, including driving–variable systems; the framework in which this paper is based. We use indefinite inner-products to describe dissipative systems, and recap these in Section 5. Section 6 discusses dissipative systems and Section 7 considers dual systems. In Section 8 we show that dissipativity of a driving–variable system and its dual is equivalent to a system of Lur'e equations having a positive, self–adjoint solution on the state space. We formulate this result as the so–called indefinite KYP Lemma. Section 9 contains dissipative balanced truncation and there we combine the material from the previous sections to provide proofs of our main results.

Regarding our notation, we let $\mathbb{R}^+ = [0, \infty)$, $\mathbb{R}^- = (-\infty, 0]$ and for $\alpha \in \mathbb{R}$ we let $\mathbb{C}^+_\alpha$ denote the open right–half complex plane consisting of those s such that $\operatorname{Re} s > \alpha$. For a linear operator $T : \mathscr{X} \to \mathscr{Z}$ between linear spaces, $\operatorname{im} T$, $\ker T$ and $\mathcal{G}(T)$ denote the image, kernel and graph of T respectively. If $\mathscr{X}$ and $\mathscr{Z}$ are Hilbert spaces then T^* denotes the usual Hilbert space adjoint of T. For a self–adjoint operator M on a finite–dimensional Hilbert space, $\sigma_+(M)$ and $\sigma_-(M)$ denote the number of nonnegative and negative eigenvalues of M respectively, counting multiplicities. Other notation is either common or defined as it is introduced.

2 Main results and examples

Mass–spring–damper arrangements are a key ingredient in many mechanical systems (a novel application, for instance, is in renewable wave energy [17]) and are the mechanical equivalent of a resistor–inductor–capacitor (RLC) circuit, ubiquitous in electrical systems. Figure 2.1 demonstrates two masses connected in series. The constants k_i and d_i are the spring and damping coefficients respectively, and m_i is the (point) mass of the i^{th} spring. At time t the i^{th} mass has position $x_i(t)$ and $F^e_i(t)$ denotes any external force applied.

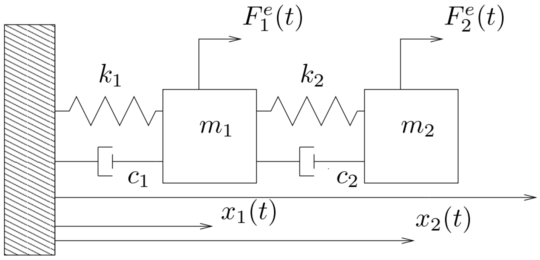

Figure 2.1: Coupled mass–spring–damper system

Now consider N masses connected in series. Elementary physics yields the following dynamics

$$\begin{aligned} m_1\ddot{x}_1 &= -k_1x_1 - d_1\dot{x}_1 + k_2(x_2 - x_1) + d_2(\dot{x}_2 - \dot{x}_1) + F_1^e, \\ m_i\ddot{x}_i &= -k_i(x_i - x_{i-1}) - d_i(\dot{x}_i - \dot{x}_{i-1}) \\ &\quad + k_{i+1}(x_{i+1} - x_i) + d_{i+1}(\dot{x}_{i+1} - \dot{x}_i) + F_i^e, \quad 2 \le i \le N-1, \\ m_N\ddot{x}_N &= -k_N(x_N - x_{N-1}) - d_N(\dot{x}_N - \dot{x}_{N-1}) + F_N^e, \end{aligned} \tag{2.1}$$

where we have suppressed the time dependence (t) for notational convenience. We can write (2.1) in first order form as a $2N \times 2N$ linear system

$$\dot{x}(t) = Ax(t) + Bv(t), \quad x(0) = x_0, \quad t \ge 0, \tag{2.2}$$

in the usual way with

$$x(t) = \begin{bmatrix} x_1(t) & \dots & x_N(t) & \dot{x}_1(t) & \dots & \dot{x}_N(t) \end{bmatrix}^T, \quad v(t) = \begin{bmatrix} F_1^e(t) & \dots & F_N^e(t) \end{bmatrix}^T,$$

and for some initial configuration x_0. Here the superscript ^{T} denotes transpose. We wish to model the external signal, denoted by w, as (possibly not all of) the external forces and the velocities of the masses (and possibly linear combinations thereof), *without* specifying which are inputs and which are outputs in the usual sense. We can do so by writing

$$w(t) = Cx(t) + Dv(t), \quad t \ge 0, \tag{2.3}$$

for some choice of C and D. The motivation for doing so is that it is not always clear where causal input–output relationships exist between the external signals [18]. Loosely speaking, the combination of (2.2) and (2.3) gives rise to a driving–variable system (defined precisely in Section 4), which is an example of a continuous time, time invariant, finite–dimensional state space system. We comment that the so–called driving–variable v is *not* considered as an input, but rather as a latent variable.

The total energy of the mass–spring–damper arrangement at time $t \ge 0$ is given by

$$E(x(t)) := \frac{1}{2}m_1\dot{x}_1^2(t) + \frac{1}{2}k_1x_1^2(t) + \sum_{i=2}^{N}\left[\frac{1}{2}m_ix_i^2(t) + \frac{1}{2}k_i(\dot{x}_i(t) - \dot{x}_{i-1}(t))^2\right], \tag{2.4}$$

the sum of the kinetic energies of the masses and the potential energies of the springs. By differentiating this expression some elementary calculations using the dynamics (2.1) shows that for each $t \ge 0$

$$\int_0^t \sum_{i=1}^{N} F_i^e(s)\dot{x}_i(s)\, ds = E(x(t)) - E(x_0) + d_1\dot{x}_1^2(t) + \sum_{i=2}^{N} d_i(x_i(t) - x_{i-1}(t))^2. \tag{2.5}$$

The left hand side of (2.5) is the integral of a quadratic form of the external signal, and in fact the integrand can be written as an indefinite inner product (recapped in Section 5), which we denote here by $[w(s), w(s)]$. Note that indefinite inner products are in general *not* positive definite, that is $[z, z] < 0$ can occur. To fix ideas in our example we write

$$w = \begin{bmatrix} F_1^e & \dots & F_N^e & \dot{x}_1 & \dots & \dot{x}_N \end{bmatrix}^T,$$

which determines C and D in (2.3) and furthermore

$$2\sum_{i=1}^{N} F_i^e(s)\dot{x}_i(s) = \left\langle \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} w(s), w(s) \right\rangle = [w(s), w(s)], \quad s \geq 0, \tag{2.6}$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product on $\mathbb{C}^{2N}$. When $x_0 = 0$ then we see from (2.5) and (2.6) that for each $t \geq 0$

$$\int_0^t [w(s), w(s)]\, ds \geq 0,$$

a property called signal dissipativity (defined in Section 6) of the driving–variable system. Such a notion encompasses the classical notions of dissipativity (namely scattering or impedance passive) of input–state–output systems. In the context of our example this simply means that energy is dissipated over time.

In Section 8 we prove that, under certain assumptions, signal dissipativity is equivalent to the existence of a positive self–adjoint operator P on the state space $\mathscr{X}$ such that for all signals w and states x of a driving–variable system

$$\int_0^t [w(s), w(s)]\, ds \geq \langle x(t), Px(t) \rangle_{\mathscr{X}} - \langle x_0, Px_0 \rangle_{\mathscr{X}}, \quad \forall\, t \geq 0,\ \forall\, x_0 \in \mathscr{X}. \tag{2.7}$$

In our mass–spring–damper example, from inequality (2.5) we see that the energy E gives rise to one such operator satisfying (2.7).

In general, we prove a so–called indefinite KYP Lemma, formulated as Theorem 8.5, which gives necessary and sufficient conditions for the existence of a P in (2.7) in terms of the solution of a set of Lur'e equations. Of this result both the classical Bounded Real and Positive Real Lemmas can be seen as special cases. There we also prove that there exist 'extremal' positive self–adjoint operators $P_m < P_M$ satisfying (2.7) that are also the unique solutions of the following dissipative optimal control problems

$$\begin{aligned} \langle P_M x_0, x_0 \rangle_{\mathscr{X}} &= \inf_{\substack{w \in L^2(\mathbb{R}^-;\mathscr{W}) \\ x(0)=x_0}} \int_{\mathbb{R}^-} [w(s), w(s)]_{\mathscr{W}}\, ds, \\ -\langle P_m x_0, x_0 \rangle_{\mathscr{X}} &= \inf_{\substack{w \in L^2(\mathbb{R}^+;\mathscr{W}) \\ x(0)=x_0}} \int_{\mathbb{R}^+} [w(s), w(s)]_{\mathscr{W}}\, ds. \end{aligned} \tag{2.8}$$

The above minimisation problems are subject to the driving–variable system (2.2)–(2.3) over $\mathbb{R}^-$ and $\mathbb{R}^+$ respectively. Dissipativity of a driving-variable system (particularly condition (2.7)) shows that the second infimum in (2.8) is finite. It can be demonstrated that the first infimum in (2.8) is finite when the dual system is dissipative, which for driving–variable systems is a property that need not necessarily follow from dissipativity of the original system. A dissipative driving–variable system with dissipative dual is called jointly dissipative.

For model reduction we obtain a dissipative balanced realisation of (2.2)–(2.3) by balancing P_m and P_M^{-1}. The balanced system is then truncated according to the size of the dissipative singular values. Section 9 contains the details. In addition to facilitating our proof of the gap metric error bound for positive real balanced truncation, considering model reduction in an input–output free framework is of separate interest. The behavioral approach, [19] and especially [18], argues that the choice of inputs and outputs of a given system is often artificial. As already mentioned, the gap metric error bound for dissipative balanced truncation is independent of any input–output decomposition of the external signal.

The main result of this paper is the following theorem; precise definitions of the notions involved are given later in the article.

Theorem 2.1. *Given a minimal jointly dissipative driving–variable system Σ let $(\sigma_i)_{i=1}^m$ denote the dissipative singular values and for $r < m$ let Σ_r denote the dissipative balanced truncation of Σ. The* *following bound holds*

$$\hat{\delta}(\Sigma, \Sigma_r) \leq 2 \sum_{i=r+1}^{m} \sigma_i. \tag{2.9}$$

A corollary of Theorem 2.1 is a new gap metric error bound for classical positive real balanced truncation.

Corollary 2.2. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a positive real rational transfer function with positive real singular values $(\sigma_i)_{i=1}^m$ and for $r < m$ let J_r denote the positive real balanced truncation. The following bound holds*

$$\hat{\delta}(J, J_r) \leq 2 \sum_{i=r+1}^{m} \sigma_i. \tag{2.10}$$

This section is concluded with two worked examples.

Example 2.3. Consider the mass–spring–damper arrangement with $N = 10$ springs and dynamics given by (2.1), with therefore 20 states. We assume that an external force is applied to the first and last spring, and so choose the external signal

$$w = \begin{bmatrix} F_1^e & F_{10}^e & \dot{x}_1 & \dot{x}_{10} \end{bmatrix}^T.$$

The spring constants are

$$m_i = 1, \quad k_i = \frac{1}{2}, \quad d_i = \frac{1}{4}, \quad \forall\, i \in \{1, 2, \ldots, 10\}.$$

It can be shown that all of the conditions of Theorem 2.1 are satisfied and so we consider dissipative balanced approximations of the resulting jointly dissipative driving–variable system. Figure 2.2 plots the (log of the) errors in the gap metric and the (log of the) error bounds against the order of the dissipative balanced truncation, as computed in Matlab. We note that for $n \geq 13$ the error bounds are smaller than the errors, which is a consequence of the inaccuracy of the `gapmetric` function in MATLAB, which has a maximal tolerance of 10^{-5}. For $n \geq 13$ this tolerance is attained by the error. This suggests that the gap metric error bound is tight and is a better approximation of the actual error than the error computed by the function `gapmetric` for $n \geq 13$.

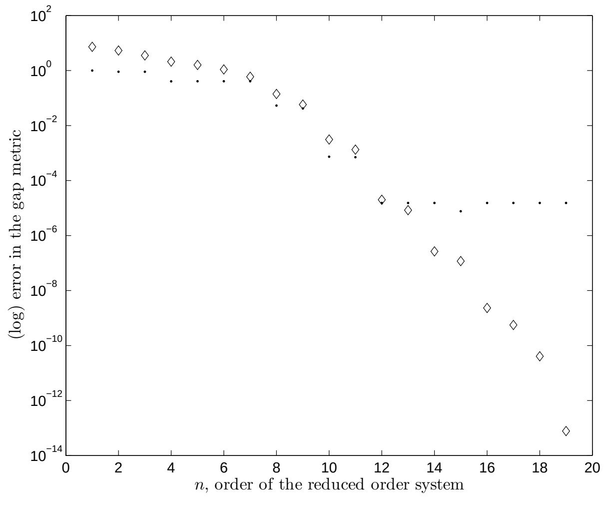

Figure 2.2: Dissipative balanced truncations of the mass–spring–damper arrangement of Example 2.3. The dotted $(\cdot)$ marks denote the distance in the gap metric between the original and truncated dissipative systems, and the error bounds (2.9) are plotted with diamond $(\diamond)$ marks.

Example 2.4. Consider the 1D Euler–Bernoulli beam with Kelvin–Voigt and viscous damping

$$EIz_{\xi\xi\xi\xi}(t,\xi) + \rho z_{tt}(t,\xi) + cz_t(t,\xi) + dz_{\xi\xi\xi\xi t}(t,\xi) = 0, \quad \xi \in [0,1],\ t \geq 0, \tag{2.11}$$

where $EI, \rho > 0$ are beam constants and $c, d > 0$ are the damping constants. The beam is a cantilever beam so that

$$z(t,0) = z_\xi(t,0) = 0, \quad t \geq 0. \tag{2.12}$$

The above PDE (2.11)–(2.12) is an input–state–output system under the application of the collocated boundary control and observation

$$\left.\begin{aligned} (EIz_{\xi\xi} + dz_{\xi\xi t})(t,1) = 0,& \\ u(t) := -(EIz_{\xi\xi\xi} + dz_{\xi\xi\xi t})(t,1),& \\ y(t) := z_t(t,1) + u(t),& \end{aligned}\right\} \quad t \geq 0, \tag{2.13}$$

and furthermore is (strictly) positive real. A finite element discretisation of (2.11)–(2.13), partitioning $[0,1]$ into $N \in \mathbb{N}$ intervals, with $2N$ cubic Hermite polynomial elements gives rise to the ODE

$$K\mathbf{z}(t) + M\ddot{\mathbf{z}}(t) + D\dot{\mathbf{z}}(t) - Ju(t) = 0, \quad \forall\, t \geq 0. \tag{2.14}$$

Here $M = M^* > 0$ is the stiffness matrix so that we can rewrite (2.14) in first order as

$$\begin{aligned} \dot{x}(t) &= \underbrace{\begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}D \end{bmatrix}}_{=:A} x(t) + \underbrace{\begin{bmatrix} 0 \\ M^{-1}J \end{bmatrix}}_{=:B} u(t), \\ y(t) &= \underbrace{\begin{bmatrix} 0 & J^T \end{bmatrix}}_{=:C} x(t) + u(t), \end{aligned} \tag{2.15}$$

where $x = \left[\begin{smallmatrix} \mathbf{z} \\ \dot{\mathbf{z}} \end{smallmatrix}\right]$ and ^{T} denote matrix transposition. The input–state–output system (2.15) is positive real and has order $4N$. For simulations we take $N = 5$ and beam parameters as in [20, Table 11]

$$EI = 43.95, \quad \rho = 1.02, \quad c = 2, \quad d = 0.82.$$

Figure 2.3 plots the (log of the) errors in the gap metric and the (log of the) error bounds against the order of the positive real balanced truncation, as computed in Matlab. We see that for $n \geq 7$ the error in the gap metric is larger than the error bound, which as we explain in Example 2.3, can be attributed to the maximal tolerance 10^{-5} of the MATLAB function `gapmetric`.

3 Bounded real and positive real balanced truncation for input–state–output systems

We review model reduction by bounded real and positive real balanced truncation, introduced in [5] and [4] respectively. The survey article by Gugercin & Antoulas [6], as well as Antoulas [21] also include summaries of the material, but with a somewhat different emphasis.

Let $\mathscr{U}$ and $\mathscr{Y}$ denote finite–dimensional Hilbert spaces and $H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ the space of bounded analytic $B(\mathscr{U}, \mathscr{Y})$–valued functions on the open right–half complex plane. We recall that the proper rational $H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ functions are precisely the transfer functions of usual stable input–state–output systems, with input and output spaces $\mathscr{U}$ and $\mathscr{Y}$ respectively. We denote a realisation of such a function by the quadruple $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, which for now we always assume is minimal, and so controllable and observable (and hence A is Hurwitz).

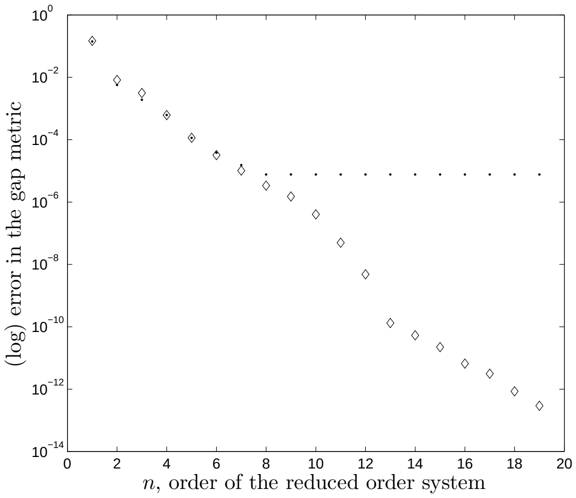

Figure 2.3: Positive real balanced truncations of a FE approximation of a damped Euler–Bernoulli beam (2.11)–(2.13) from Example 2.4. The dotted $(\cdot)$ marks denote the distance in the gap metric between the original system and its positive real balanced truncation, and the error bounds (2.9) are plotted with diamond $(\diamond)$ marks.

3.1 Bounded real functions

We first recall the definition of bounded real.

Definition 3.1. Let $\mathscr{U}$ and $\mathscr{Y}$ denote Banach spaces. A function $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ is said to be bounded real if

$$\|G\|_\infty \leq 1, \tag{3.1}$$

and we say that such a G is strictly bounded real if the above inequality is strict.

Remark 3.2. (i) Synonymously with the term 'bounded real' the terms Schur, contractive and scattering passive are used. In the model reduction literature [21] the term 'bounded real balanced truncation' seems to have become standard and therefore we use this terminology.

(ii) Note that, in spite of the terminology, there is no realness assumption in Definition 3.1. However, if such an assumption is made about the original system, then realness of the reduced order system can be concluded.

Bounded real balanced truncation makes use of the well–known Bounded Real Lemma, see Anderson & Vongpanitlerd [22], which gives a state space characterisation of rational bounded real functions. Since we shall make frequent use of this result, we recall it below.

Proposition 3.3 (Bounded Real Lemma)**.** *Given rational $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$, with $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ a minimal input–state–output realisation of G, the following are equivalent.*

(i) G is bounded real.

(ii) For input $u \in L^2(\mathbb{R}^+; \mathscr{U})$ and output $y \in L^2(\mathbb{R}^+; \mathscr{Y})$ with initial condition $x_0 = 0$

$$\int_0^t \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}} \, ds \geq 0, \quad \forall\, t \geq 0.$$

(iii) There exists a positive, self–adjoint operator P on $\mathscr{X}$ such that for input $u \in L^2(\mathbb{R}^+; \mathscr{U})$ with output $y \in L^2(\mathbb{R}^+; \mathscr{Y})$ and initial state $x_0 \in \mathscr{X}$

$$\int_0^t \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}} \, ds \geq \langle Px(t), x(t)\rangle - \langle Px_0, x_0\rangle, \quad \forall\, t \geq 0.$$

(iv) There exists a triple of operators (P, K, W) with

$$P : \mathscr{X} \to \mathscr{X}, \quad K : \mathscr{X} \to \mathscr{U}, \quad W : \mathscr{U} \to \mathscr{U},$$

and P positive and self–adjoint satisfying the bounded real Lur'e equations

$$A^*P + PA + C^*C = -K^*K, \tag{3.2a}$$
$$PB + C^*D = -K^*W, \tag{3.2b}$$
$$I - D^*D = W^*W. \tag{3.2c}$$

Moreover, if any of (i) – (iv) hold then there are positive self–adjoint solutions P_m, P_M to (3.2) such that for any positive, self–adjoint solution P of (3.2) we have

$$0 < P_m \leq P \leq P_M. \tag{3.3}$$

The extremal operators P_m, P_M are the optimal cost operators of the bounded real optimal control problems, namely:

$$\langle P_M x_0, x_0 \rangle_{\mathscr{X}} = \inf_{u \in L^2(\mathbb{R}^-;\mathscr{U})} \int_{\mathbb{R}^-} \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}} \, ds, \tag{3.4a}$$
$$-\langle P_m x_0, x_0 \rangle_{\mathscr{X}} = \inf_{u \in L^2(\mathbb{R}^+;\mathscr{U})} \int_{\mathbb{R}^+} \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}} \, ds. \tag{3.4b}$$

The minimisation problems (3.4) are subject to the minimal input–state–output realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$.

Proof. A proof of the equivalence of (i) and (iv) is given in [22]. The authors assume that $\dim \mathscr{U} = \dim \mathscr{Y}$, but the result is true in general. A short series of calculations gives the implications $(iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i)$. $\square$

If $P = P^* > 0$ is a solution of (3.2), for some K, W then an elementary calculation shows that $P^{-1} > 0$ solves the dual bounded real Lur'e equations,

$$AQ + QA^* + BB^* = -LL^*, \tag{3.5a}$$
$$QC^* + BD^* = -LX^*, \tag{3.5b}$$
$$I - DD^* = XX^*, \tag{3.5c}$$

for some operators $L : \mathscr{Y} \to \mathscr{X}$, $X : \mathscr{Y} \to \mathscr{Y}$. By the Bounded Real Lemma, there are extremal self–adjoint solutions Q_m, Q_M to (3.5) such that for any self–adjoint solution Q to (3.5); $0 < Q_m \leq Q \leq Q_M$. In particular, it is not difficult to see that

$$P_m = Q_M^{-1}, \quad \text{and} \quad P_M = Q_m^{-1}. \tag{3.6}$$

Remark 3.4. Solutions of the bounded real Lur'e equations are generally not unique. By this we mean that firstly, there are in general many different nonnegative self–adjoint operators (the operator P in (3.2)) solving (3.2). Secondly, given a solution (P, K, W) of (3.2), the operator P does not uniquely determine K and W. For example, for $U : \mathscr{U} \to \mathscr{U}$ unitary, we have that (P, UK, UW) is also a solution of (3.2). Similar statements apply for the dual equations.

3.2 Bounded real balanced truncation

All balanced truncation schemes approximate an input–output relationship, typically the transfer function, by removing states from a state space realisation that are unimportant in some sense. Truncation in the state space is of course dependent on the particular realisation which is chosen. Therefore it is crucial to quantify what unimportant in some sense means. For example, by identifying quantities associated with the system that are not realisation dependent. Bounded real balanced truncation makes use of the self–adjoint, positive optimal cost operators P_M and P_m from (3.4).

Definition 3.5. A minimal realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ of a rational bounded real $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ is bounded real balanced, or in bounded real balanced co–ordinates, if

$$P_m = P_M^{-1} =: \Pi. \tag{3.7}$$

The bounded real singular values, which we denote by $(\sigma_k)_{k=1}^m$, are the nonnegative square roots of the eigenvalues of the product $P_m P_M^{-1}$. The bounded real singular values are ordered such that $\sigma_k > \sigma_{k+1} > 0$ for each k and we let r_k denote the (geometric) multiplicity of σ_k.

Remark 3.6. (i) Condition (3.3) implies that the bounded real singular value are all less than or equal to one. Furthermore, equality (3.6) implies that the bounded real singular values are equal to the square roots of the eigenvalues of $P_m Q_m$. In practise it is sometimes easier to compute Q_m than P_M^{-1}.

(ii) Regarding the terminology bounded real singular value; some authors use the terminology characteristic value (for example [23]), but singular value is also prevalent in the literature [21], and so we keep this convention. It is also true that the σ_i are the singular values of a related Hankel operator (see for example [24]) further supporting this terminology. At any rate, the σ_i are the nonnegative squareroots of the eigenvalues of $P_m P_M^{-1}$ which are realisation independent and depend only on G.

(iii) It follows from [21, Lemma 7.3], that given a minimal stable realisation of a bounded real transfer function there always exists a similarity transformation such that the transformed realisation is bounded real balanced.

Suppose that the realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is bounded real balanced. Since Π is self–adjoint it is diagonalisable, and so we can decompose the state space $\mathscr{X}$ into an orthogonal sum of eigenspaces of Π. For $r < m$ let $\mathscr{X}_r$ and $\mathscr{Z}_r$ denote the sum of the first r and last $m - r$ eigenspaces of Π respectively, with respective orthogonal projections $P_{\mathscr{X}_r}$ and $P_{\mathscr{Z}_r}$. Then with respect to the orthogonal decomposition $\mathscr{X} = \mathscr{X}_r \oplus \mathscr{Z}_r$, the operators A, B, C and Π split as

$$\Pi = \begin{bmatrix} P_{\mathscr{X}_r}\Pi|_{\mathscr{X}_r} & 0 \\ 0 & P_{\mathscr{Z}_r}\Pi|_{\mathscr{Z}_r} \end{bmatrix} = \begin{bmatrix} \Pi_1 & 0 \\ 0 & \Pi_2 \end{bmatrix}, \qquad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix},$$
$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}.$$

The dimension of $\mathscr{X}_r$ is $\sum_{j=1}^r r_j$, the sum of the geometric multiplicities of the first r bounded real singular values. The truncated system with realisation $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$ is called the bounded real balanced truncation and its transfer function denoted by G_r is called the *reduced order transfer function obtained by bounded real balanced truncation.*

The main result for bounded real balanced truncation is stated below.

Theorem 3.7. *Given rational $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ bounded real, let $(\sigma_j)_{j=1}^m$ denote the bounded real singular values, with multiplicities r_j. For $r < m$ let G_r denote the reduced order transfer obtained by bounded real balanced truncation. Then G_r is bounded real and the following error bound holds*

$$\|G - G_r\|_\infty \leq 2 \sum_{j=r+1}^{m} \sigma_j. \tag{3.8}$$

Let $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ denote a minimal, bounded real balanced realisation of G. Then in the bounded real balanced truncation $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$, A_{11} is Hurwitz. If additionally G is strictly bounded real, then G_r has MacMillan degree $\sum_{j=1}^r r_j$ and $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$ is minimal and bounded real balanced.

Proof. See Theorem 2 and Section IV of [5]. The assumption there that G is strictly bounded real is not needed to prove that G_r is bounded real and that A_{11} is Hurwitz. The authors also assume throughout that $\mathscr{U} = \mathscr{Y}$, but this is not needed and the proof for the general case is essentially the same. □

3.3 Positive real functions

We recall the definition of a positive real function.

Definition 3.8. An operator–valued analytic function $J : \mathbb{C}_0^+ \to B(\mathscr{U})$, where $\mathscr{U}$ is a Hilbert space, is positive real if

$$J(s) + [J(s)]^* \geq 0, \quad \forall\, s \in \mathbb{C}_0^+. \tag{3.9}$$

We say that the analytic function $J : \mathbb{C}_0^+ \to B(\mathscr{U})$ is strictly positive real if there exists $\eta > 0$ such that

$$J(s) + [J(s)]^* \geq \eta I, \quad \forall\, s \in \mathbb{C}_0^+. \tag{3.10}$$

Remark 3.9. (i) The term strictly positive real is used for various slightly different concepts in the literature, as described in, for example, Wen [25]. The condition (3.10) is equivalent to the concept sometimes called extended strictly positive real, as in Sun *et al.* [26, Definition 2.1].

(ii) Note that positive real functions need not belong to H^∞ as they need not be proper. The rational function $s \mapsto s$ is a counter–example. Furthermore, proper rational positive real functions need not belong to H^∞ as they may have simple poles on the imaginary axis, such as $s \mapsto \frac{1}{s}$.

(iii) We do not assume that a positive real function is real on the real axis as is sometimes done in the literature.

(iv) Synonymously with the term 'positive real function' the terms impedance passive function, Weyl function, Titchmarsh–Weyl function and Caratheodory–Nevanlinna function are used (see, for example, Staffans [27]). In the model reduction literature the term 'positive real balanced truncation' seems to have become standard and therefore we use this terminology.

Positive real balanced truncation is identical in spirit to bounded real balanced truncation and the key ingredient is the Positive Real Lemma, which analogously to the Bounded Real Lemma provides a state space characterisation of rational positive real functions.

Proposition 3.10 (Positive Real Lemma)**.** *Given rational $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$, with $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ a minimal input–state–output realisation of J, the following are equivalent.*

(i) J is positive real.

(ii) For input $u \in L^2(\mathbb{R}^+; \mathscr{U})$ and output $y \in L^2(\mathbb{R}^+; \mathscr{U})$ with initial condition $x_0 = 0$

$$\int_0^t 2\,\mathrm{Re}\,\langle u(s), y(s)\rangle_{\mathscr{U}}\; ds \geq 0, \quad \forall\, t \geq 0.$$

(iii) There exists a positive, self-adjoint operator P on $\mathscr{X}$ such that for input $u \in L^2(\mathbb{R}^+; \mathscr{U})$ with output $y \in L^2(\mathbb{R}^+; \mathscr{U})$ and initial state $x_0 \in \mathscr{X}$

$$\int_0^t 2\,\mathrm{Re}\,\langle u(s), y(s)\rangle_{\mathscr{U}}\; ds \geq \langle Px(t), x(t)\rangle - \langle Px_0, x_0\rangle, \quad \forall\, t \geq 0.$$

(iv) There exists a triple of operators (P, K, W) with

$$P : \mathscr{X} \to \mathscr{X}, \quad K : \mathscr{X} \to \mathscr{U}, \quad W : \mathscr{U} \to \mathscr{U},$$

and P positive and self–adjoint satisfying the positive real Lur'e equations

$$A^*P + PA = -K^*K, \tag{3.11a}$$
$$PB - C^* = -K^*W, \tag{3.11b}$$
$$D + D^* = W^*W. \tag{3.11c}$$

If any of $(i) - (iv)$ hold then there are positive, self–adjoint solutions $\tilde{P}_m, \tilde{P}_M$ to (3.11) such that any positive, self–adjoint solution P to (3.11) satisfies

$$0 < \tilde{P}_m \leq P \leq \tilde{P}_M. \tag{3.12}$$

The extremal operators $\tilde{P}_m$, $\tilde{P}_M$ are the optimal cost operators of the positive real optimal control problems, namely:

$$\langle \tilde{P}_M x_0, x_0 \rangle_{\mathscr{X}} = \inf_{u \in L^2(\mathbb{R}^-;\mathscr{U})} \int_{\mathbb{R}^-} 2\,\mathrm{Re}\,\langle u(s), y(s) \rangle_{\mathscr{U}}\, ds, \tag{3.13a}$$

$$-\langle \tilde{P}_m x_0, x_0 \rangle_{\mathscr{X}} = \inf_{u \in L^2(\mathbb{R}^+;\mathscr{U})} \int_{\mathbb{R}^+} 2\,\mathrm{Re}\,\langle u(s), y(s) \rangle_{\mathscr{U}}\, ds. \tag{3.13b}$$

The minimisation problems (3.13) are subject to the minimal input–state–output realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ of J.

Proof. A proof of the equivalence of (i) and (iv) is given in Section 5.2 of [22]. For the equivalence of (i) and (ii) see Willems [28, Theorem 1]. A short series of calculations shows $(iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i)$. □

An elementary calculation demonstrates that if $P = P^* > 0$ solves (3.11), for some K, W, then P^{-1} solves the dual positive real Lur'e equations

$$AQ + QA^* = -LL^*, \tag{3.14a}$$

$$QC^* - B = -LX^*, \tag{3.14b}$$

$$D + D^* = XX^*, \tag{3.14c}$$

for some operators $L : \mathscr{U} \to \mathscr{X}$, $X : \mathscr{U} \to \mathscr{U}$. By the Positive Real Lemma, there are positive self–adjoint solutions $\tilde{Q}_m$, $\tilde{Q}_M$ to (3.14) such that for any self–adjoint solution Q to (3.14) it follows that $0 < \tilde{Q}_m \leq Q \leq \tilde{Q}_M$. Again, it readily follows that

$$\tilde{P}_m = \tilde{Q}_M^{-1}, \quad \text{and} \quad \tilde{P}_M = \tilde{Q}_m^{-1}. \tag{3.15}$$

3.4 Positive real balanced truncation

Given a rational positive real $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$, a realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ of J is positive real balanced if (3.7) holds with P_M^{-1} and P_m replaced by $\tilde{P}_M^{-1}$ and $\tilde{P}_m$ respectively. The positive real singular values, which we denote by $(\sigma_k)_{k=1}^m$, are the nonnegative square roots of the eigenvalues of the product $\tilde{P}_m \tilde{P}_M^{-1}$, again ordered such that $\sigma_k > \sigma_{k+1} > 0$ for each k, with r_k denoting the (geometric) multiplicity of σ_k. The positive real balanced truncation is defined in the same way as the bounded real balanced truncation. Note from (3.15) that the positive real singular values are equal to the nonnegative square roots of the eigenvalues of $\tilde{P}_m \tilde{Q}_m$.

The main results for positive real balanced truncation are stated below.

Theorem 3.11. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a positive real transfer function and let $(\sigma_j)_{j=1}^m$ denote the positive real singular values, each with multiplicity r_j. For $r < m$, let J_r denote the reduced order transfer obtained by positive real balanced truncation. Then $J_r \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ and J_r is positive real. If $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ denotes a minimal positive real balanced realisation of J then in the positive real balanced truncation $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$, A_{11} is Hurwitz. If additionally J is strictly positive real, then J_r has MacMillan degree $\sum_{j=1}^r r_j$ and $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$ is minimal and positive real balanced.*

Proof. See Harshavardhana *et al.* [29] and the references therein. □

Theorem 3.12. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a strictly positive real transfer function with minimal realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ and let $(\sigma_j)_{j=1}^m$ denote the positive real singular values, each with multiplicity r_j. For $r < m$, let J_r denote the reduced order transfer obtained by positive real balanced truncation. Then the following bounds hold* *(i)* $\left\|(D^*+J)^{-1}-(D^*+J_r)^{-1}\right\|_\infty \le 2\|(D+D^*)^{-1}\|\sum_{j=r+1}^m \sigma_j$,

(ii) $\left\|(D^*+J)^{-1}[J-J_r](D^*+J_r)^{-1}\right\|_\infty \le 2\|(D+D^*)^{-1}\|\sum_{j=r+1}^m \sigma_j$,

(iii) $\left\|(D^*+J_r)^{-1}(J-J_r)\right\|_\infty \le 2\|(D+D^*)^{-1}\|\,\|D^*+J\|_\infty\sum_{j=r+1}^m \sigma_j$.

Proof. See [21, Proposition 7.17] (or [6, Theorem 5]) for a proof of (i). The bound in (ii) is equivalent to that in (i), see [21, Remark 7.5.2]. For a proof of (iii) see [6, Lemma 3]. □

Remark 3.13. The error bound for positive real balanced truncation

$$\|J-J_r\|_\infty \le \|D+D^*\|\sum_{j=r+1}^m \frac{2\sigma_j}{(1-\sigma_j)^2}\left(1+\sum_{l=1}^{j-1}\frac{2\sigma_i}{1-\sigma_i}\right),$$

claimed in [7, Theorem 2] is false. A counter–example is contained in Guiver & Opmeer [30].

Example 3.14. The H^∞ error bound (3.8) *does not* hold for positive real balanced truncation, as the example beneath shows. Consider

$$\mathbb{C}_0^+ \ni s \mapsto J(s) = 1+\frac{s}{s+1} = 2-\frac{1}{s+1}.$$

The function J is positive real as

$$J(s)+[J(s)]^* = 2\mathrm{Re}\,J(s) = 2\left[2-\mathrm{Re}\left(\frac{s+1}{|s+1|^2}\right)\right]\ge 0, \quad \forall\, s\in\mathbb{C}_0^+,$$

and it is easy to see that

$$A=-1, \quad B=1, \quad C=-1, \quad D=2,$$

is a (minimal) realisation of J. In this instance as $D+D^*$ is invertible, the positive real Lur'e equations collapse to the positive real algebraic Riccati equation

$$A^*P+PA+(PB-C^*)(D+D^*)^{-1}(PB-C^*)^*=0,$$

which in this instance is a scalar quadratic equation with extremal solutions

$$0<\tilde{P}_m = 3-2\sqrt{2} < \tilde{P}_M = 3+2\sqrt{2}.$$

The positive real singular value is $\sigma = \tilde{P}_m\tilde{P}_M^{-1} = 17-12\sqrt{2} = 0.0294$. Thus for $r=0$ we have $J_r = D$ and as

$$|J(0)-D| = 1 > 2\sigma,$$

the H^∞ error bound cannot hold. In fact, since proper rational positive real functions need not belong to H^∞, an H^∞ error bound seems less natural and instead Corollary 2.2 provides the gap metric error bound (2.10) for positive real balanced truncation.

4 State space systems

In this section we collect precise definitions of the systems we consider; namely input–state–output, driving–variable and output–nulling systems. These objects and the relations between them form a backbone of this work. They are examples of state space systems as in Willems [31] or state/signal systems as studied in the discrete time infinite–dimensional case by Arov & Staffans [11] and [32]–[34]. More recently state/signal systems have been studied in continuous time by Arov, Staffans & Kurula [12] and [35]–[38].

4.1 Definitions

We begin with a remark on what we mean by the solution of a linear inhomogeneous ODE.

Remark 4.1. Let $\mathscr{U}$, $\mathscr{X}$ denote finite–dimensional Hilbert spaces and let A, B denote operators

$$A : \mathscr{X} \to \mathscr{X}, \quad B : \mathscr{U} \to \mathscr{X}.$$

For $u \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{U})$ by a solution x of the (formal) ordinary differential equation

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad t \geq 0, \tag{4.1}$$

we mean a mild solution (as in [39, Definition 3.1.4]), that is, the continuous function $x \in C(\mathbb{R}^+; \mathscr{X})$ given by the variation of parameters formula

$$\mathbb{R}^+ \ni t \mapsto x(t) = \mathrm{e}^{At} x_0 + \int_0^t \mathrm{e}^{A(t-s)} Bu(s)\, ds. \tag{4.2}$$

In the above e^A denotes the matrix exponential of A. In fact, x given by (4.2) belongs to the Sobolev space $W^{1,2}_{\text{loc}}(\mathbb{R}^+; \mathscr{X})$ and thus x satisfies equation (4.1) for almost all $t \geq 0$.

Definition 4.2. Given $\mathscr{U}$, $\mathscr{X}$ and $\mathscr{Y}$ finite–dimensional Hilbert spaces we define an input–state–output node as an operator

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{U} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \end{bmatrix}, \tag{4.3}$$

with associated formal differential equation

$$\begin{bmatrix} \dot{x}(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \quad x(0) = x_0, \quad t \geq 0. \tag{4.4}$$

The spaces $\mathscr{U}$, $\mathscr{X}$ and $\mathscr{Y}$ are called the input, state and output spaces respectively. We define the set of trajectories $\mathcal{T}$ by

$$\mathcal{T} := \left\{ \begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} \in \begin{bmatrix} C(\mathbb{R}^+; \mathscr{X}) \\ L^2_{\text{loc}}(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]) \end{bmatrix} : \exists\, x_0 \in \mathscr{X} \text{ such that (4.4) holds} \right\}. \tag{4.5}$$

The component x of a trajectory is understood as a solution of (4.4) as described in Remark 4.1. We define the set of trajectories from $x_0 \in \mathscr{X}$, $\mathcal{T}(x_0)$, by

$$\mathcal{T}(x_0) := \left\{ \begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} \in \mathcal{T} \ : \ x(0) = x_0 \right\}, \tag{4.6}$$

and define the set of externally generated trajectories $\mathcal{T}_{\text{ext}}$ by

$$\mathcal{T}_{\text{ext}} := \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \in L^2_{\text{loc}}(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]) : \exists\, x \in C(\mathbb{R}^+; \mathscr{X}) \text{ such that } \begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} \in \mathcal{T}(0) \right\}. \tag{4.7}$$

The set of stable externally generated trajectories $\mathcal{S}$ is given by

$$\mathcal{S} = \mathcal{T}_{\text{ext}} \cap L^2(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]),$$

and we say that the input–state–output system is (input–output) stable if the projection of $\mathcal{S}$ onto $L^2(\mathbb{R}^+; \mathscr{U})$ is all of $L^2(\mathbb{R}^+; \mathscr{U})$. We call the pair consisting of the node (4.3) and set of trajectories (4.5) an input–state–output system, which we denote by $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{iso}}$.

With an input–state–output system we can associate an input/output map $\mathfrak{D}$ and transfer function G in the usual way. In particular, using our notation we have that $\mathfrak{D} : L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{U}) \to L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{Y})$ satisfies

$$y = \mathfrak{D}u, \quad \text{where } u, y \text{ are such that } \begin{bmatrix} u \\ y \end{bmatrix} \in \mathcal{T}_{\text{ext}}. \tag{4.8}$$

Definition 4.3. Given $\mathscr{V}$, $\mathscr{X}$ and $\mathscr{W}$ finite–dimensional Hilbert spaces we define a driving–variable node as an operator

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{V} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{W} \end{bmatrix}, \tag{4.9}$$

where $D : \mathscr{V} \to \mathscr{W}$ is assumed injective, with associated formal differential equation

$$\begin{bmatrix} \dot{x}(t) \\ w(t) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x(t) \\ v(t) \end{bmatrix}, \quad x(0) = x_0, \quad t \geq 0. \tag{4.10}$$

The spaces $\mathscr{V}$, $\mathscr{X}$ and $\mathscr{W}$ are called the driving–variable, state and signal spaces respectively. We define the set of trajectories $\mathcal{T}$ by

$$\mathcal{T} := \left\{ \begin{bmatrix} x \\ w \end{bmatrix} \in \begin{bmatrix} C(\mathbb{R}^+; \mathscr{X}) \\ L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{W}) \end{bmatrix} : \exists\, x_0 \in \mathscr{X},\, v \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{V}) \text{ such that (4.10) holds} \right\}. \tag{4.11}$$

The component x of a trajectory is understood as a solution of (4.10) as described in Remark 4.1. We define the set of trajectories from $x_0 \in \mathscr{X}$, $\mathcal{T}(x_0)$, by

$$\mathcal{T}(x_0) := \left\{ \begin{bmatrix} x \\ w \end{bmatrix} \in \mathcal{T} \;:\; x(0) = x_0 \right\}, \tag{4.12}$$

and define the set of externally generated trajectories $\mathcal{T}_{\text{ext}}$ by

$$\mathcal{T}_{\text{ext}} := \left\{ w \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{W}) : \exists\, x \in C(\mathbb{R}^+; \mathscr{X}) \text{ such that } \begin{bmatrix} x \\ w \end{bmatrix} \in \mathcal{T}(0) \right\}. \tag{4.13}$$

We define the set of stable externally generated trajectories $\mathcal{S}$ by

$$\mathcal{S} = \mathcal{T}_{\text{ext}} \cap L^2(\mathbb{R}^+; \mathscr{W}). \tag{4.14}$$

We call the pair consisting of the node (4.9) and set of trajectories (4.11) a driving–variable system, which we denote by $\left(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T}\right)_{\text{dv}}$.

Remark 4.4. Although a driving–variable system looks like a standard input–state–output system, its interpretation is very different. The external signal w incorporates all the interaction with the external world (so in the standard input–state–output formulation it would contain both the outputs *and* the inputs). The driving–variable v is a latent variable used to mathematically describe the dynamics and may or may not have any physical meaning (much like a state).

Definition 4.5. Given $\mathscr{V}$, $\mathscr{X}$ and $\mathscr{W}$ finite–dimensional Hilbert spaces we define an output–nulling node as an operator

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{W} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{V} \end{bmatrix}, \tag{4.15}$$

where $D : \mathscr{W} \to \mathscr{V}$ is assumed surjective, with associated formal algebraic–differential equation

$$\begin{bmatrix} \dot{x}(t) \\ 0 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x(t) \\ w(t) \end{bmatrix}, \quad x(0) = x_0, \quad t \geq 0. \tag{4.16}$$

The spaces $\mathscr{V}$, $\mathscr{X}$ and $\mathscr{W}$ are called the error, state and signal spaces respectively. We define the set of trajectories $\mathcal{T}$ by

$$\mathcal{T} := \left\{ \begin{bmatrix} x \\ w \end{bmatrix} \in \begin{bmatrix} C(\mathbb{R}^+; \mathscr{X}) \\ L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{W}) \end{bmatrix} : \exists\, x_0 \in \mathscr{X} \text{ such that (4.16) holds} \right\}. \tag{4.17}$$

The component x of a trajectory is understood as a solution of (4.16) as described in Remark 4.1. We define the set of trajectories from $x_0 \in \mathscr{X}$, $\mathcal{T}(x_0)$, by

$$\mathcal{T}(x_0) := \left\{ \begin{bmatrix} x \\ w \end{bmatrix} \in \mathcal{T} \;:\; x(0) = x_0 \right\}, \tag{4.18}$$

and define the set of externally generated trajectories $\mathcal{T}_{\text{ext}}$ by

$$\mathcal{T}_{\text{ext}} := \left\{ w \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{W}) : \exists\, x \in C(\mathbb{R}^+; \mathscr{X}) \text{ such that } \begin{bmatrix} x \\ w \end{bmatrix} \in \mathcal{T}(0) \right\}. \tag{4.19}$$

We define the set of stable externally generated trajectories $\mathcal{S}$ by

$$\mathcal{S} = \mathcal{T}_{\text{ext}} \cap L^2(\mathbb{R}^+; \mathscr{W}). \tag{4.20}$$

We call the pair consisting of the node (4.15) and the set of trajectories (4.17) an output–nulling system, which we denote by $([\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}], \mathcal{T})_{\text{on}}$.

Remark 4.6. The surjectivity of D in Definition 4.5 implies that for every $x_0 \in \mathscr{X}$ the corresponding set of trajectories from x_0, $\mathcal{T}(x_0)$, is non–empty. This can be proven directly, but it is also a consequence of Theorem 4.15.

4.2 Admissible decompositions

In this section we investigate when given a driving–variable or output–nulling system, it is possible to decompose the original signal space into an input space $\mathscr{U}$ and output space $\mathscr{Y}$ such that the trajectories of the driving–variable or output–nulling system are the trajectories of an input–state–output system. First of all we introduce some notation that we shall make frequent use of.

Remark 4.7. Let $\mathscr{U}, \mathscr{Y}$ denote a direct sum decomposition of a finite–dimensional Hilbert space $\mathscr{W}$, which we denote by $\mathscr{W} = \mathscr{U} \oplus \mathscr{Y}$ (and recall that $\mathscr{U}$ and $\mathscr{Y}$ are termed complementary subspaces when such a decomposition holds). We understand $\mathscr{W} = \mathscr{U} \oplus \mathscr{Y}$ as $\mathscr{W} = [\begin{smallmatrix} \mathscr{U} \\ 0 \end{smallmatrix}] \oplus [\begin{smallmatrix} 0 \\ \mathscr{Y} \end{smallmatrix}]$ so that we identify $u \in \mathscr{U}$ with $[\begin{smallmatrix} u \\ 0 \end{smallmatrix}] \in [\begin{smallmatrix} \mathscr{U} \\ 0 \end{smallmatrix}]$, etc. and thus $w = u + y = [\begin{smallmatrix} u \\ y \end{smallmatrix}]$. We let $\pi^{\mathscr{Y}}_{\mathscr{U}}$ ($\pi^{\mathscr{U}}_{\mathscr{Y}}$) denote the projection of $\mathscr{W}$ onto $\mathscr{U}$ ($\mathscr{Y}$) along $\mathscr{Y}$ ($\mathscr{U}$) and given an operator

$$T : \mathscr{X} \to \mathscr{W},$$

for some linear space $\mathscr{X}$ we write

$$T_{\mathscr{U}} = \pi^{\mathscr{Y}}_{\mathscr{U}} T, \quad T_{\mathscr{Y}} = \pi^{\mathscr{U}}_{\mathscr{Y}} T. \tag{4.21}$$

Definition 4.8. Given a driving–variable system with node $[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}]$, let $\mathscr{U}$ and $\mathscr{Y}$ denote complementary subspaces of the signal space $\mathscr{W}$. We say that the pair $\mathscr{U}, \mathscr{Y}$ is admissible for $\mathscr{W}$, if

$$(\pi^{\mathscr{Y}}_{\mathscr{U}} D)^{-1} : \mathscr{U} \to \mathscr{V}, \quad \text{exists.}$$

Given an output–nulling system with node $\left[\begin{smallmatrix} A' & B' \\ C' & D' \end{smallmatrix}\right]$ and a surjective operator $E \in B(\mathscr{W})$, we say that the pair $\mathscr{U}$, $\mathscr{Y}$ of complementary subspaces is E–admissible for $\mathscr{W}$ if

$$(D' E|_{\mathscr{Y}})^{-1} : \mathscr{V} \to \mathscr{Y}, \quad \text{exists.}$$

If $E = I$, the identity on $\mathscr{W}$, then we say that the pair $\mathscr{U}, \mathscr{Y}$ is admissible instead of I–admissible.

Remark 4.9. In this article we choose to work in the framework of driving–variable systems. We could have defined an E–admissible pair for driving–variable systems, where E is now injective, but have no need for this. We need the notion of an E–admissible pair for output–nulling systems for a sensible notion of duality of systems, which we address in Section 7. We comment that many of the following results are stated and proven from the point of view of driving–variable systems and the corresponding output–nulling versions have been omitted.

The next lemma demonstrates that under our assumptions there always exists (at least one) admissible pair for the state space systems we consider.

Lemma 4.10. *The space $\mathscr{U} := \operatorname{im} D$ and any complementary subspace is always admissible for a driving–variable system with node $[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}]$. Similarly, for an output–nulling system with node $\left[\begin{smallmatrix} A' & B' \\ C' & D' \end{smallmatrix}\right]$, the subspace of $\mathscr{W}$ denoted by $\mathscr{Y}$ that is naturally isomorphic to the quotient space $\mathscr{W}/_{\ker D'}$, and any complementary subspace is always admissible (that is, I–admissible).*

Proof. The first claim follows by injectivity of D and the fact that any map surjects onto its image. The second claim follows from surjectivity of D' and the First Isomorphism Theorem. $\square$

Remark 4.11. For driving–variable systems the space $\mathscr{U} := \operatorname{im} D$ is referred to in [11] as the canonical input space.

The next definitions construct the input–state–output systems that arise from admissible pairs of a signal space.

Definition 4.12. Given a driving–variable system with node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ and an admissible pair $\mathscr{U}, \mathscr{Y}$ we define the derived $(\mathscr{U}, \mathscr{Y})$ input–state–output node by

$$\begin{bmatrix} A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}} & BD_{\mathscr{U}}^{-1} \\ C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{U} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \end{bmatrix}, \tag{4.22}$$

which we denote by $\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right]$. Recall that $C_{\mathscr{U}}, C_{\mathscr{Y}}, D_{\mathscr{U}}$ and $D_{\mathscr{Y}}$ are as in (4.21). We call the corresponding input–state–output system the derived $(\mathscr{U}, \mathscr{Y})$ input–state–output system.

Definition 4.13. Given an output–nulling system with node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, a surjective operator $E \in B(\mathscr{W})$ and an E–admissible pair $\mathscr{U}, \mathscr{Y}$, we define the E-derived input–state–output node by

$$\begin{bmatrix} A - (BE)|_{\mathscr{Y}}(DE)|_{\mathscr{Y}}^{-1}C & (BE)|_{\mathscr{U}} - (BE)|_{\mathscr{Y}}(DE)|_{\mathscr{Y}}^{-1}(DE)|_{\mathscr{U}} \\ (DE)|_{\mathscr{Y}}^{-1}C & (DE)|_{\mathscr{Y}}^{-1}(DE)|_{\mathscr{U}} \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{U} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \end{bmatrix}, \tag{4.23}$$

which we denote by $\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right]$. We call the corresponding input–state–output system the E-derived $(\mathscr{U}, \mathscr{Y})$ input–state–output system. If $E = I$, the identity on $\mathscr{W}$, then we call the I-derived $(\mathscr{U}, \mathscr{Y})$ system the derived $(\mathscr{U}, \mathscr{Y})$ system instead.

Remark 4.14.
(i) The terms admissible and derived system have two meanings, one for driving–variable systems and one for output–nulling systems. In what follows it will be made clear which meaning is being used (though it is also often clear from the context).

(ii) A driving–variable or output–nulling system may have many possible derived systems, but once we fix an admissible pair $\mathscr{U}, \mathscr{Y}$ (and where appropriate $E \in B(\mathscr{W})$), then the derived $(\mathscr{U}, \mathscr{Y})$ system is uniquely specified by its node as in Definition 4.12 or 4.13 respectively.

The following result is crucial in obtaining input–state–output systems from driving–variable and output–nulling systems as it states that the trajectories of a derived input–state–output system are the same as those of the original system.

Theorem 4.15. *Given a driving–variable system with set of trajectories $\mathcal{T}_{\mathrm{dv}}$ and an admissible pair $\mathscr{U}, \mathscr{Y}$, let $\mathcal{T}_{\mathrm{iso}}$ denote the set of trajectories of the derived $(\mathscr{U}, \mathscr{Y})$ input–state–output system. Then $\left[\begin{smallmatrix} x \\ \hline u \\ y \end{smallmatrix}\right]$ is a trajectory in $\mathcal{T}_{\mathrm{dv}}$ if, and only if, $\left[\begin{smallmatrix} x \\ \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \end{smallmatrix}\right]$ is a trajectory in $\mathcal{T}_{\mathrm{iso}}$.*

Given an output–nulling system with set of trajectories $\mathcal{T}_{\mathrm{on}}$, a surjective operator $E \in B(\mathscr{W})$ and an E–admissible pair $\mathscr{U}, \mathscr{Y}$, let $\mathcal{T}'_{\mathrm{iso}}$ denote the set of trajectories of the E-derived $(\mathscr{U}, \mathscr{Y})$ input–state–output system. Then $\left[\begin{smallmatrix} x \\ \hline u \\ y \end{smallmatrix}\right]$ is a trajectory in $\mathcal{T}_{\mathrm{on}}$ if, and only if, $\left[\begin{smallmatrix} x \\ E\left[\begin{smallmatrix} u \\ -y \end{smallmatrix}\right] \end{smallmatrix}\right]$ is a trajectory in $\mathcal{T}'_{\mathrm{iso}}$.

Remark 4.16. The conclusions of Theorem 4.15 can equivalently be expressed as $\mathcal{T}_{\mathrm{dv}}$ and $\mathcal{T}_{\mathrm{iso}}$ are isomorphic and that $\mathcal{T}_{\mathrm{on}}$ and $\mathcal{T}'_{\mathrm{iso}}$ are isomorphic. In what follows we will say that these sets are equal in the sense that

$$\begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} \in \mathcal{T}_{\mathrm{dv}} \iff \begin{bmatrix} x \\ \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \end{bmatrix} \in \mathcal{T}_{\mathrm{iso}}, \quad \text{and} \quad \begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} \in \mathcal{T}_{\mathrm{on}} \iff \begin{bmatrix} x \\ E\left[\begin{smallmatrix} u \\ -y \end{smallmatrix}\right] \end{bmatrix} \in \mathcal{T}'_{\mathrm{iso}}. \tag{4.24}$$

Proof of Theorem 4.15: We only prove the driving–variable case, as the output–nulling case is similar. Suppose that the driving–variable system has node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$. The direct sum decomposition $\mathscr{W} = \mathscr{U} \oplus \mathscr{Y}$ implies that every $w \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{W})$ can be written as $w = u + y = \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$, where $u \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{U})$ and $y \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{Y})$. As such, if x and w are the components of a trajectory in $\mathcal{T}_{\text{dv}}$ then by definition there exists a $v \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{V})$ such that

$$\dot{x} = Ax + Bv, \quad \begin{bmatrix} u \\ y \end{bmatrix} = Cx + Dv = \begin{bmatrix} C_{\mathscr{U}} x \\ C_{\mathscr{Y}} x \end{bmatrix} + \begin{bmatrix} D_{\mathscr{U}} v \\ D_{\mathscr{Y}} v \end{bmatrix}. \tag{4.25}$$

Admissibility of the pair $\mathscr{U}, \mathscr{Y}$ in the driving–variable case means that the operator $D_{\mathscr{U}}$ is invertible and hence we can eliminate v from (4.25) and obtain

$$\begin{aligned} \dot{x} &= (A - BD_{\mathscr{U}}^{-1} C_{\mathscr{U}})x + BD_{\mathscr{U}}^{-1} u, \\ y &= (C_{\mathscr{Y}} - D_{\mathscr{Y}} D_{\mathscr{U}}^{-1} C_{\mathscr{U}})x + D_{\mathscr{Y}} D_{\mathscr{U}}^{-1} u, \end{aligned} \tag{4.26}$$

so that x and $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ are the components of a trajectory in $\mathcal{T}_{\text{iso}}$. Conversely, if x and $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ are the components of a trajectory in $\mathcal{T}_{\text{iso}}$ then defining

$$v := D_{\mathscr{U}}^{-1} u - D_{\mathscr{U}}^{-1} C_{\mathscr{U}} x \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{V}),$$

and substituting back into (4.26) we recover (4.25). As such, x and $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ are the components of a trajectory in $\mathcal{T}_{\text{dv}}$, completing the proof. □

Corollary 4.17. *The set of stable externally generated trajectories of a driving–variable system with signal space $\mathscr{W}$ is a closed subspace of $L^2(\mathbb{R}^+; \mathscr{W})$.*

Proof. Let Σ and $\mathcal{S}$ denote the driving–variable system and its set of stable externally generated trajectories respectively. Let $\mathfrak{D}$ denote the input-output map of the derived $(\mathscr{U}, \mathscr{Y})$ system of Σ, for some choice of admissible pair $\mathscr{U}$, $\mathscr{Y}$ (which always exists by Lemma 4.10). A consequence of Theorem 4.15 is that

$$\mathcal{S} = \left\{ \begin{bmatrix} u \\ \mathfrak{D}u \end{bmatrix} \;:\; u \in L^2(\mathbb{R}^+; \mathscr{U}) \quad \text{such that} \quad \mathfrak{D}u \in L^2(\mathbb{R}^+; \mathscr{Y}) \right\}. \tag{4.27}$$

It is well–known from input–state–output theory that the set on the right hand side of (4.27) is closed and hence so is $\mathcal{S}$. □

It will sometimes be helpful later in this work to obtain a driving–variable system from an input–state–output system and we describe how we do so in the next lemma.

Lemma 4.18. *Every input–state–output system $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{iso}}$ with input, state and output spaces $\mathscr{U}, \mathscr{X}$ and $\mathscr{Y}$ respectively gives rise to a driving–variable system with $\mathscr{V} = \mathscr{U}$, $\mathscr{W} = \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$, driving–variable node*

$$\left[\begin{array}{c|c} A & B \\ \hline 0 & I \\ C & D \end{array}\right] : \begin{bmatrix} \mathscr{X} \\ \mathscr{V} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{W} \end{bmatrix}, \tag{4.28}$$

and set of trajectories $\mathcal{T}$.

Proof. This is immediate from the definitions, noting that the operator $\left[\begin{smallmatrix} I \\ D \end{smallmatrix}\right] : \mathscr{V} \to \mathscr{W}$ is always injective. Note that the set of trajectories of (4.28) is equal to $\mathcal{T}$ in the sense of (4.24) from Remark 4.16. □

The following lemma characterises admissibility of a direct sum decomposition of the signal space of a driving–variable system and is based on [11, Lemma 5.7]. The corollary that follows is useful in relating admissible pairs. A proof of both results can be found in Guiver [40, Lemma 3.1.23].

Lemma 4.19. *Given a driving–variable system $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{dv}}$, let $\mathscr{U}, \mathscr{Y}$ denote a direct sum decomposition of the signal space $\mathscr{W}$. The following are equivalent.*

(i) The pair $\mathscr{U}, \mathscr{Y}$ is admissible.

(ii) There exists a map $\tilde{D} : \mathscr{U} \to \mathscr{Y}$ such that $\operatorname{im} D$ *has the graph representation*

$$\operatorname{im} D = \mathcal{G}(\tilde{D}) = \left\{ \begin{bmatrix} u \\ \tilde{D}u \end{bmatrix} \, : \, u \in \mathscr{U} \right\}.$$

Corollary 4.20. *Given a driving–variable system with signal space $\mathscr{W}$, let $\mathscr{U}, \mathscr{Y}$ denote an admissible pair. Any other direct sum decomposition $\mathscr{U}_1, \mathscr{Y}_1$ of $\mathscr{W}$ is admissible if and only if* $\dim \mathscr{U} = \dim \mathscr{U}_1$.

Remark 4.21. In Willems & Trentelman [41, p. 55] the input cardinality $\mathtt{m}(\mathfrak{B})$ of a behaviour $\mathfrak{B}$ is defined as the maximal number of unconstrained components of $w \in \mathfrak{B}$. It is stated [41, p. 60] that the input cardinality $\mathtt{m}(\mathfrak{B})$ is equal to the dimension of the input space of any input–state–output representation of $\mathfrak{B}$. Similarly, the output cardinality is defined as $\mathtt{w} - \mathtt{m}(\mathfrak{B})$, where $w \in \mathfrak{B}$ has $\mathtt{w}$ components. Since all linear differential behaviours $\mathfrak{B}$ admit a driving–variable representation, Lemma 4.10 and Lemma 4.19 show that the input cardinality of the behaviour described by a driving–variable system is $\dim(\operatorname{im} D)$. Moreover, by Lemma 4.19 if $\mathscr{U}, \mathscr{Y}$ is an admissible pair then necessarily $\dim \mathscr{U}$ equals the input cardinality.

4.3 Minimality and the gap metric

Here we consider minimality of state space systems and recap the gap metric. For the latter see also Kato [42, p.197].

Definition 4.22. An input–state–output, driving–variable or output–nulling system with state space $\mathscr{X}$ and set of trajectories $\mathcal{T}$, is said to be minimal if supposing $\mathcal{T} = \mathcal{T}'$ for another such system with state space $\mathscr{X}'$ it follows that $\dim \mathscr{X} \leq \dim \mathscr{X}'$.

Remark 4.23. The above definition in the input–state–output case is consistent with the usual definition.

Lemma 4.24. *A driving–variable system is minimal if and only if for every admissible pair $\mathscr{U}, \mathscr{Y}$ the derived $(\mathscr{U}, \mathscr{Y})$ system is minimal.*

Proof. See [40, Lemma 3.1.27]. □

Definition 4.25. For $\mathscr{M}, \mathscr{N}$ non–empty closed subspaces of a Hilbert space $\mathscr{Z}$, the gap is defined as

$$\hat{\delta}(\mathscr{M}, \mathscr{N}) = \|P_{\mathscr{M}} - P_{\mathscr{N}}\|, \tag{4.29}$$

where $P_{\mathscr{M}}, P_{\mathscr{N}}$ are the orthogonal projections of $\mathscr{Z}$ onto $\mathscr{M}$ and $\mathscr{N}$ respectively. For $\mathscr{H}$ another Hilbert space and closed linear operators S, T from $\mathscr{Z}$ to $\mathscr{H}$, the gap between S and T is defined as

$$\hat{\delta}(S, T) := \hat{\delta}(\mathcal{G}(S), \mathcal{G}(T)). \tag{4.30}$$

For S, T bounded operators we recall the bound

$$\hat{\delta}(S, T) \leq \|S - T\|, \tag{4.31}$$

proven in [42, Theorem 2.14], that we shall make frequent use of later. We now define the gap between two driving–variable systems.

Definition 4.26. Let Σ_1 and Σ_2 denote two driving–variable systems each with the same signal space and sets of stable externally generated trajectories $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. We define the gap between Σ_1 and Σ_2 as

$$\hat{\delta}(\Sigma_1, \Sigma_2) := \hat{\delta}(\mathcal{S}_1, \mathcal{S}_2).$$

Remark 4.27. The gap between two driving–variable systems is well–defined as the sets of stable externally generated trajectories are closed subspaces of $L^2(\mathbb{R}^+; \mathscr{W})$ by Corollary 4.17.

We conclude this section with some remarks on other notions of an admissible decomposition. Recall that input–state–output systems necessarily have *proper* rational transfer functions. It is possible to choose decompositions of a signal space into an input and output space such that the resulting transfer function is rational, but not necessarily proper. Such a notion is considered further in [40] under the name *weakly admissible.* In Dai [43, Theorem 2-6.3] it is proven that rational functions are precisely the transfer functions of descriptor systems. Consequently, [40, Proposition 3.6.20] gives that every stable input–output trajectory that is described by rational transfer functions is a trajectory of a finite–dimensional driving–variable or output–nulling system with a weakly admissible decomposition of the signal space. In this article we seek to generalise bounded real and positive real balanced truncation of input–state–output systems for which the present notion of admissible is suitable.

5 Finite–dimensional indefinite inner–product spaces

We use the machinery of indefinite inner–products to describe so–called dissipative state space systems. We demonstrate that indefinite inner–products generalise the well–known scattering passive and impedance passive supply rates for bounded real and positive real input–state–outputs systems respectively. In this section we collect the required results on complex finite–dimensional indefinite inner–product spaces. Three supplementary references for this material are Bognár [44] and Gohberg *et al.* [45], [46].

Definition 5.1. Let $\mathscr{W}$ denote a finite–dimensional linear space. A function $[\cdot,\cdot] : \mathscr{W}\times\mathscr{W} \to \mathbb{C}$ is called an indefinite (non–degenerate) inner–product on $\mathscr{W}$ if the following axioms are satisfied:

(1) Linearity in the second argument

$$[x, \alpha y_1 + \beta y_2] = \alpha[x, y_1] + \beta[x, y_2], \quad \forall\, x, y_1, y_2 \in \mathscr{W} \quad \text{and} \quad \forall\, \alpha, \beta \in \mathbb{C}.$$

(2) Antisymmetry

$$[x, y] = \overline{[y, x]}, \quad \forall\, x, y \in \mathscr{W}.$$

(3) Non–degeneracy; if $[x, y] = 0$ for all $y \in \mathscr{W}$, then $x = 0$.

Note that in contrast to a definite inner–product, $[x, x] < 0$ can occur. We call $\mathscr{W}$ equipped with $[\cdot,\cdot]$ an indefinite inner–product space, denoted by $(\mathscr{W}, [\cdot,\cdot])$ or sometimes just $\mathscr{W}$.

Lemma 5.2. *The space $(\mathscr{W}, [\cdot,\cdot])$ is a finite–dimensional indefinite inner–product space if and only if there exists a unique definite inner–product $\langle\cdot,\cdot\rangle$ on $\mathscr{W}$ and unitary self–adjoint operator E (with respect to $\langle\cdot,\cdot\rangle$) such that*

$$[x, y] = \langle Ex, y\rangle, \quad \forall\, x, y \in \mathscr{W}. \tag{5.1}$$

Given $(\mathscr{W}, [\cdot,\cdot])$ we say that E is the signature operator of $(\mathscr{W}, [\cdot,\cdot])$ and $\langle\cdot,\cdot\rangle$ satisfying (5.1) *is the induced definite inner–product. Conversely, a definite $\langle\cdot,\cdot\rangle$ inner product on $\mathscr{W}$ and a unitary, self–adjoint operator E together induce the indefinite inner–product $[\cdot,\cdot]$ by* (5.1).

Proof. The proof is a straightforward extension of the arguments in [46, p. 8] (see also [40, Lemma 3.2.2]). □

Remark 5.3. In Section 6 we shall consider driving–variable systems where $\mathscr{W}$ is an indefinite inner–product space. We remark that in this instance we use the *definite* inner–product $\langle\cdot,\cdot\rangle$ (induced by the indefinite inner–product $[\cdot,\cdot]$) in the theory of driving–variable systems established in Section 4. In particular, $L^2(\mathbb{R}^+; \mathscr{W})$ is a Hilbert space when $\mathscr{W}$ is equipped with $\langle\cdot,\cdot\rangle$.

Example 5.4. For $\mathscr{U},\mathscr{Y}$ Hilbert spaces define $\mathscr{W} := \left[\begin{smallmatrix}\mathscr{U}\\ \mathscr{Y}\end{smallmatrix}\right]$. By Lemma 5.2, the operator $E = \left[\begin{smallmatrix} I & 0\\ 0 & -I\end{smallmatrix}\right]$ on $\mathscr{W}$ induces an indefinite inner product on $\mathscr{W}$ that satisfies

$$[w,w] = \left[\begin{bmatrix}u\\ y\end{bmatrix},\begin{bmatrix}u\\ y\end{bmatrix}\right] = \left\langle \begin{bmatrix}I & 0\\ 0 & -I\end{bmatrix}\begin{bmatrix}u\\ y\end{bmatrix},\begin{bmatrix}u\\ y\end{bmatrix}\right\rangle = \|u\|^2_{\mathscr{U}} - \|y\|^2_{\mathscr{Y}}. \tag{5.2}$$

Similarly, when $\mathscr{Y}=\mathscr{U}$, the operator $E = \left[\begin{smallmatrix} 0 & I\\ I & 0\end{smallmatrix}\right]$ on $\mathscr{W} := \left[\begin{smallmatrix}\mathscr{U}\\ \mathscr{U}\end{smallmatrix}\right]$ induces the indefinite inner product on $\mathscr{W}$ satisfying

$$[w,w] = \left[\begin{bmatrix}u\\ y\end{bmatrix},\begin{bmatrix}u\\ y\end{bmatrix}\right] = \left\langle \begin{bmatrix}0 & I\\ I & 0\end{bmatrix}\begin{bmatrix}u\\ y\end{bmatrix},\begin{bmatrix}u\\ y\end{bmatrix}\right\rangle = 2\mathrm{Re}\,\langle u,y\rangle_{\mathscr{U}}. \tag{5.3}$$

In the context of classical dissipative input–state–output systems, where $\mathscr{U}$ and $\mathscr{Y}$ denote input and output spaces, the right hand side of (5.2) and (5.3) are the scattering passive supply rate and the impedance passive supply rate respectively.

Definition 5.5. A subspace $\mathscr{S}$ of an indefinite inner–product space $(\mathscr{W},[\cdot,\cdot])$ is called nonnegative (respectively, neutral, nonpositive) if $[x,x]\geq 0$ ($[x,x]=0, [x,x]\leq 0$) for all $x\in\mathscr{S}$. A nonnegative subspace is called maximal if it is not a proper subset of another nonnegative subspace, with similar definitions for maximal neutral and nonpositive subspaces.

Definition 5.6. Given an indefinite inner–product space $(\mathscr{W},[\cdot,\cdot])$ we say that $\mathscr{W}_+,\mathscr{W}_-\subseteq\mathscr{W}$ is a fundamental decomposition of $\mathscr{W}$, denoted $\mathscr{W}=\mathscr{W}_+[+]-\mathscr{W}_-$ if,

(1) $\mathscr{W}_+$ equipped with $[\cdot,\cdot]_{|\mathscr{W}_+}$ and $\mathscr{W}_-$ equipped with $-[\cdot,\cdot]_{|\mathscr{W}_-}$ are Hilbert spaces.

(2) $\mathscr{W}$ is a direct sum of $\mathscr{W}_+$ and $\mathscr{W}_-$, orthogonal with respect to $[\cdot,\cdot]$.

Note that if $\mathscr{W}=\mathscr{W}_+[+]-\mathscr{W}_-$ then by (1) $\mathscr{W}_+$ is nonnegative and $\mathscr{W}_-$ is nonpositive. Fundamental decompositions are in general not unique.

Example 5.7. Let $(\mathscr{W},[\cdot,\cdot])$ denote an indefinite inner–product space and let E denote the signature operator from Lemma 5.2, which has eigenvalues ±1. Let $\mathscr{U}$ and $\mathscr{Y}$ denote the eigenspaces corresponding to $+1$ and -1 respectively, so that with respect to this decomposition

$$E = \begin{bmatrix} I_{\sigma_+(E)} & 0\\ 0 & -I_{\sigma_-(E)}\end{bmatrix}.$$

Then $\mathscr{U}$ and $\mathscr{Y}$ are clearly a direct sum decomposition of $\mathscr{W}$ and from (5.2) it follows that $\mathscr{U},\mathscr{Y}$ are in fact a fundamental decomposition. If $\sigma_+(E)=\sigma_-(E)$ then under the transformation

$$\begin{bmatrix}\mathscr{U}'\\ \mathscr{Y}'\end{bmatrix} := \frac{1}{\sqrt{2}}\begin{bmatrix}I & -I\\ I & I\end{bmatrix}\begin{bmatrix}\mathscr{U}\\ \mathscr{Y}\end{bmatrix}, \tag{5.4}$$

$\mathscr{W}$ is the direct sum of $\mathscr{U}'$ and $\mathscr{Y}'$ and the signature operator E with respect to this decomposition has the block form $\left[\begin{smallmatrix} 0 & I\\ I & 0\end{smallmatrix}\right]$. It now follows from (5.3) that $\mathscr{U}'$, $\mathscr{Y}'$ are *not* orthogonal with respect to $[\cdot,\cdot]$ and so are not a fundamental decomposition.

Definition 5.8. For a subspace $\mathscr{S}$ of an indefinite inner–product space $(\mathscr{W},[\cdot,\cdot])$ we denote by $\mathscr{S}^{[\perp]}$ the orthogonal companion with respect to $[\cdot,\cdot]$, which is defined as

$$\mathscr{S}^{[\perp]} = \{w\in\mathscr{W}\ :\ [w,v]=0,\ \forall\, v\in\mathscr{S}\}.$$

Note that $\mathscr{S}\cap\mathscr{S}^{[\perp]}\neq\{0\}$ in general.

Lemma 5.9. *If $\mathscr{W}_+$, $\mathscr{W}_-$ is a fundamental decomposition of an indefinite inner–product space $(\mathscr{W},[\cdot,\cdot])$ with signature operator E, then*

(i) $\mathscr{S}\subseteq\mathscr{W}$ is nonnegative if and only if $\mathscr{S}=\mathcal{G}(T)$, for $T:\mathscr{W}_+\supseteq D(T)\to\mathscr{W}_-$ a linear contraction, where $D(T)$ is the domain of T. Additionally $\mathscr{S}$ is maximal nonnegative if and only if $D(T)=\mathscr{W}_+$.

(ii) $\mathscr{S} \subseteq \mathscr{W}$ *is maximal nonnegative if and only if* $\mathscr{S}$ *is nonnegative and* $\mathscr{S}^{[\perp]}$ *is nonpositive.*

(iii) The dimension of any maximal nonnegative (nonpositive) subspace is equal to the multiplicity of 1 (-1) *as an eigenvalue of* E. *Hence any two maximal nonnegative (nonpositive) subspaces are isomorphic.*

(iv) The dimensions of the nonnegative parts of any two fundamental decompositions are the same.

Proof. For parts (*i*) and (*ii*) see [44], namely Theorem 4.2, Theorem 4.4 and Lemma 4.5 on pp. 105-106. For part (*iii*) see [45, Theorem 1.3, p. 15]. Part (*iv*) follows immediately from (*iii*). □

Corollary 5.10. *For a driving–variable system with node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, *indefinite inner–product signal space* $\mathscr{W}$ *and signature operator* E, *a fundamental decomposition* $\mathscr{W}_+$, $\mathscr{W}_-$ *of* $\mathscr{W}$ *is an admissible pair for* $\mathscr{W}$ *if and only if*

$$\dim \mathscr{W}_+ = \sigma_+(E) = \dim(\operatorname{im} D).$$

Proof. By Lemma 4.19, the direct sum decomposition $\mathscr{W}_+$, $\mathscr{W}_-$ of $\mathscr{W}$ is admissible if and only if $\dim \mathscr{W}_+ = \dim(\operatorname{im} D)$ as by Lemma 4.10 the pair $\operatorname{im} D$ and any complementary subspace is always admissible. That $\dim \mathscr{W}_+ = \sigma_+(E)$ follows from Lemma 5.9 (*iii*) above. □

6 Dissipative systems

Definition 6.1. Let $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})$ denote a driving–variable or output–nulling system with indefinite inner–product signal space $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$. We say that $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})$ is state–signal dissipative if there exists a positive, self–adjoint operator P on $\mathscr{X}$ such that for all $t \geq 0$

$$\int_0^t [w(s), w(s)]_{\mathscr{W}}\, ds \geq \langle Px(t), x(t)\rangle_{\mathscr{X}} - \langle Px(0), x(0)\rangle_{\mathscr{X}}, \quad \forall \left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \in \mathcal{T}. \tag{6.1}$$

We call $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})$ signal dissipative if for all $t \geq 0$

$$\int_0^t [w(s), w(s)]_{\mathscr{W}}\, ds \geq 0, \quad \forall\, w \in \mathcal{T}_{\text{ext}}. \tag{6.2}$$

Remark 6.2. (i) The right hand side of (6.1) represents the change in the internal energy of the system at time t, whilst the left hand side is the net energy that flows in to (or out of, depending on sign) the system up to time t.

(ii) The two concepts of dissipativity expressed by (6.1) and (6.2) are well–defined. Namely, by definition the state x is continuous and so the point evaluations on the right hand side of (6.1) make sense. Simple applications of the Hölder inequality using the facts that signals belong to L^2_{loc} and the expression (5.1) imply that for each $t \geq 0$ the left hand sides are also finite in absolute value.

From Definition 6.1 we see immediately that signal dissipativity is a necessary condition for state–signal dissipativity. In Theorem 8.5 the converse implication is addressed.

Definition 6.3. An input–state–output system with input and output spaces $\mathscr{U}$ and $\mathscr{Y}$ respectively, equipped with an indefinite inner-product on the corresponding signal space $\mathscr{W} := \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$ is called state–signal dissipative or signal dissipative if the corresponding driving variable system constructed in Lemma 4.18 is dissipative in the same sense.

The above definition describes the effect of equipping an input–state–output system with a quadratic supply rate in the language of [28], for instance, those described in Example 5.4. The next crucial result shows that, first, dissipativity is preserved under admissible input–output decompositions and, second, that bounded real and positive real input–state–output systems can be seen as particular decompositions of dissipative driving–variable systems.

Theorem 6.4. *Let Σ denote a driving–variable system with indefinite inner–product signal space $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$ and assume that the pair $\mathscr{U}, \mathscr{Y}$ is admissible. Let Σ_{D} denote the derived $(\mathscr{U}, \mathscr{Y})$ system.*

(i) If Σ is state–signal or signal dissipative then Σ_{D} is dissipative in the same sense as Σ.

(ii) If Σ is signal dissipative and $\mathscr{U}, \mathscr{Y}$ is a fundamental decomposition of $\mathscr{W}$ then Σ_{D} is bounded real.

(iii) If Σ is signal dissipative and $\mathscr{U}, \mathscr{Y}$ is a fundamental decomposition of $\mathscr{W}$ and additionally $\dim \mathscr{U} = \dim \mathscr{Y}$ then there exists an admissible pair $\mathscr{U}', \mathscr{Y}'$ such that the derived $(\mathscr{U}', \mathscr{Y}')$ system is positive real.

Proof. *(i):* This follows from the definitions and Theorem 4.15, which ensures that Σ and Σ_{D} have the same trajectories.

(ii): If $\mathscr{U}$, $\mathscr{Y}$ is a fundamental decomposition of $\mathscr{W}$ then for each $w \in \mathscr{W}$, there exist $u \in \mathscr{U}$ and $y \in \mathscr{Y}$ such that $w = u + y = \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ and so

$$\begin{aligned} 0 &\leq \int_0^t [w(s), w(s)]_{\mathscr{W}}\, ds = \int_0^t [u(s)+y(s), u(s)+y(s)]_{\mathscr{W}}\, ds = \int_0^t [u(s), u(s)]_{\mathscr{W}} + [y(s), y(s)]_{\mathscr{W}}\, ds \\ &= \int_0^t \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}}\, ds, \quad \forall\, t \geq 0, \end{aligned} \tag{6.3}$$

where we have used the orthogonality of the fundamental decomposition. From (6.3) the Bounded Real Lemma implies that Σ_{D} is bounded real.

(iii): By Lemma 5.9 the condition $\dim \mathscr{U} = \dim \mathscr{Y}$ is equivalent to $\sigma_+(E) = \sigma_-(E)$, where E is the signature operator of $\mathscr{W}$ and so the decomposition $\mathscr{U}'$, $\mathscr{Y}'$ of $\mathscr{W}$ from (5.4), Example 5.7, applies. This is an admissible decomposition of $\mathscr{W}$ by Corollary 4.20, and with respect to this decomposition, every $w \in \mathscr{W}$ satisfies $w = u' + y' = \begin{bmatrix} u' \\ y' \end{bmatrix}$ with $u' \in \mathscr{U}'$ and $y' \in \mathscr{Y}'$ such that

$$\begin{aligned} 0 &\leq \int_0^t [w(s), w(s)]_{\mathscr{W}}\, ds = \int_0^t [u'(s)+y'(s), u'(s)+y'(s)]_{\mathscr{W}}\, ds \\ &= \int_0^t 2\mathrm{Re}\, \langle u'(s), y'(s)\rangle_{\mathscr{U}'}\, ds, \quad \forall\, t \geq 0. \end{aligned} \tag{6.4}$$

From (6.4) the Positive Real Lemma implies that the derived $(\mathscr{U}', \mathscr{Y}')$ system is positive real. ☐

7 Dual systems

This section considers the duals of state space systems. We shall require duality (and in particular dissipative systems with dissipative duals) for a property called liveness (see Remark 8.6) that will ultimately be required for dissipative balanced truncation. We note that so far the signal and error spaces $\mathscr{V}$ have not required any geometry, and have simply been linear spaces. From hereon in we shall impose that $\mathscr{V}$ is a Hilbert space.

Definition 7.1. Given an input–state–output node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ with input, state and output spaces $\mathscr{U}, \mathscr{X}$ and $\mathscr{Y}$ respectively, we define the dual input–state–output node as the operator

$$\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{U} \end{bmatrix}. \tag{7.1}$$

We denote by $\mathcal{T}^*$ the corresponding set of trajectories (as in Definition 4.2) and call $(\left[\begin{smallmatrix} A^* & C^* \\ B^* & D^* \end{smallmatrix}\right], \mathcal{T}^*)_{\mathrm{iso}}$ the dual input–state–output system.

The dual of a driving–variable node is an output–nulling node and vice versa. In order to formulate these definitions we need some more notation.

Definition 7.2. Let $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$ denote an indefinite inner–product space. We define $(\mathscr{W}^*, [\cdot,\cdot]_{\mathscr{W}^*})$ as the linear space $\mathscr{W}$ equipped with the indefinite inner–product $-[\cdot,\cdot]_{\mathscr{W}}$, and call $\mathscr{W}^*$ the anti–space of $\mathscr{W}$. Moreover, given a Hilbert space $\mathscr{Z}$ and operators

$$S : \mathscr{Z} \to \mathscr{W}, \quad T : \mathscr{W} \to \mathscr{Z},$$

we define

$$S^\dagger : \mathscr{W}^* \to \mathscr{Z}, \quad T^\dagger : \mathscr{Z} \to \mathscr{W}^*,$$

as the adjoint maps, taken with respect to the Hilbert space inner–product on $\mathscr{Z}$ and the indefinite inner–product on $\mathscr{W}^*$. Thus $S^\dagger$ and $T^\dagger$ are such that

$$\left.\begin{aligned} [w, Sz]_{\mathscr{W}^*} &= \langle S^\dagger w, z\rangle_{\mathscr{Z}} \\ \langle z, Tw\rangle_{\mathscr{Z}} &= [T^\dagger z, w]_{\mathscr{W}^*} \end{aligned}\right\} \quad \forall\, z \in \mathscr{Z},\ \forall\, w \in \mathscr{W}. \tag{7.2}$$

Definition 7.3. Given a driving–variable node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, with indefinite inner–product signal space $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$, we define the dual output–nulling node by

$$\begin{bmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{W}^* \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{V} \end{bmatrix}, \tag{7.3}$$

which has error, state and signal spaces $\mathscr{V}$, $\mathscr{X}$ and $\mathscr{W}^*$ respectively. We denote by $\mathcal{T}^*$ the corresponding set of trajectories and call $(\left[\begin{smallmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{smallmatrix}\right], \mathcal{T}^*)_{\mathrm{on}}$ the dual output–nulling system.

Definition 7.4. Given an output–nulling node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, with indefinite inner–product signal space $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$, we define the dual driving–variable node by

$$\begin{bmatrix} A^* & C^\dagger \\ B^* & D^\dagger \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{V} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{W}^* \end{bmatrix}, \tag{7.4}$$

with driving–variable, state and signal spaces $\mathscr{V}, \mathscr{X}$ and $\mathscr{W}^*$ respectively. We denote by $\mathcal{T}^*$ the corresponding set of trajectories and call $(\left[\begin{smallmatrix} A^* & C^\dagger \\ B^* & D^\dagger \end{smallmatrix}\right], \mathcal{T}^*)_{\mathrm{dv}}$ the dual driving–variable system.

Proposition 7.5. *Given a driving–variable (output–nulling) system with set of trajectories $\mathcal{T}$, let $\mathcal{T}^*$ denote the set of trajectories of the dual output–nulling (driving–variable) system. Then*

$$\int_0^t [w(s), w_*(t-s)]_{\mathscr{W}}\, ds = 0, \quad \forall\, w \in \mathcal{T}_{\mathrm{ext}},\ \forall\, w_* \in \mathcal{T}^*_{\mathrm{ext}},\ \forall\, t \ge 0.$$

Proof. As both proofs are similar, we prove the case when $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})$ is a driving–variable system and $\mathcal{T}^*$ denotes the set of trajectories of the dual output–nulling system. For $w \in \mathcal{T}_{\mathrm{ext}}$, $w_* \in \mathcal{T}^*_{\mathrm{ext}}$ and $t \ge 0$ a calculation shows that

$$\begin{aligned} \int_0^t [w(s), w_*(t-s)]_{\mathscr{W}}\, ds &= \int_0^t [Cx(s) + Dv(s), w_*(t-s)]_{\mathscr{W}}\, ds \\ &= \int_0^t -\langle x(s), C^\dagger w_*(t-s)\rangle_{\mathscr{X}} - \langle v(s), D^\dagger w_*(t-s)\rangle_{\mathscr{V}}\, ds \\ &= \int_0^t \langle x(s), \dot{x}_*(t-s) - A^* x_*(t-s)\rangle_{\mathscr{X}} - \langle v(s), B^* x_*(t-s)\rangle_{\mathscr{V}}\, ds \\ &= \int_0^t \langle x(s), \dot{x}_*(t-s)\rangle_{\mathscr{X}} - \langle Ax(s) + Bv(s), x_*(t-s)\rangle_{\mathscr{X}}\, ds \\ &= -\int_0^t \frac{d}{ds}\langle x(s), x_*(t-s)\rangle_{\mathscr{X}}\, ds = \langle x(0), x_*(t)\rangle_{\mathscr{X}} - \langle x(t), x_*(0)\rangle_{\mathscr{X}} \\ &= 0. \end{aligned}$$

□

Remark 7.6. We comment that, loosely speaking, we have chosen throughout to specify 'systems' as nodes and the trajectories generated by these nodes. It is possible to define systems more abstractly as sets of trajectories (with certain properties) and then inferring the existence of such nodes that generate these trajectories. This latter approach is taken by Arov & Staffans in their development of state/signal systems. Although equivalent, we have chosen the former as we feel that it is more suitable to model reduction where the (usually physically motivated) systems considered are often specified by nodes. These differing approaches are also true for dual systems and for the latter approach Proposition 7.5 can be used as the *definition* of dual trajectories. This is in many ways more elegant than the somewhat peculiar definition that arises by considering the nodes first as we have done in Definitions 7.3 and 7.4.

Adjoint operators depend on the choice of inner–product the linear space is equipped with. The same is true of the adjoint operators with respect to indefinite inner–products. In Lemma 5.2 we have seen that indefinite inner–products can be characterised by definite inner–products and signature operators, usually denoted by E. The next lemma demonstrates how these three objects interact.

Lemma 7.7. *Let $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$ denote an indefinite inner–product space with signature operator E. Given a Hilbert space $\mathscr{X}$ and operators $S : \mathscr{X} \to \mathscr{W}$, $\quad T : \mathscr{W} \to \mathscr{X}$, the Hilbert space adjoints and indefinite inner–product space adjoints are related by*

$$S^\dagger = -S^* E, \qquad T^\dagger = -ET^*, \tag{7.5}$$

where $S^\dagger, T^\dagger$ are as in Definition 7.2. Furthermore, if $\mathscr{U}$ and $\mathscr{Y}$ are a direct sum decomposition for $\mathscr{W}$ then

$$\begin{aligned} S^*|_{\mathscr{U}} &= (S_{\mathscr{U}})^*, & S^*|_{\mathscr{Y}} &= (S_{\mathscr{Y}})^*, \\ (T|_{\mathscr{U}})^* &= (T^*)_{\mathscr{U}}, & (T|_{\mathscr{Y}})^* &= (T^*)_{\mathscr{Y}}, \end{aligned} \tag{7.6}$$

where recall that $S_{\mathscr{U}} = \pi^{\mathscr{Y}}_{\mathscr{U}} S$ and $S_{\mathscr{Y}} = \pi^{\mathscr{U}}_{\mathscr{Y}} S$.

Proof. We prove (7.5) and (7.6) for S only, as the proof is very similar for T. For $z \in \mathscr{X}$ and $w \in \mathscr{W}$ we see that

$$\langle S^\dagger w, z\rangle_{\mathscr{X}} = [w, Sz]_{\mathscr{W}^*} = -[w, Sz]_{\mathscr{W}} = -\langle Ew, Sz\rangle_{\mathscr{W}} = -\langle S^* Ew, z\rangle_{\mathscr{X}},$$

which gives (7.5) by the unicity of the adjoint. For $u \in \mathscr{U}$ and $z \in \mathscr{X}$ the equalities

$$\langle S_{\mathscr{U}} z, u\rangle_{\mathscr{U}} = \langle \pi^{\mathscr{Y}}_{\mathscr{U}} Sz, u\rangle_{\mathscr{U}} = \langle Sz, \left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right]\rangle_{\mathscr{W}} = \langle z, S^* \left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right]\rangle_{\mathscr{X}} = \langle z, S^*|_{\mathscr{U}} u\rangle_{\mathscr{X}},$$

establish (7.6) for $\mathscr{U}$. The case for $\mathscr{Y}$ is similar. □

Remark 7.8. We comment that a signature operator E of an indefinite inner–product space $\mathscr{W}$ is unitary and hence surjective. In particular, the operator E is suitable for considering E–admissible pairs for an output–nulling system with signal space $\mathscr{W}$, see Definition 4.8 (and hence the deliberate use of the same symbol).

Our main result of this section is the following proposition which demonstrates that with the above definitions of dual systems, the operations of duality and taking derived input–state–output systems using admissible pairs commute.

Proposition 7.9. *Let $\Sigma = (\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ denote a driving–variable system with indefinite inner–product signal space $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$ and corresponding signature operator E, and let $\Sigma^* = (\left[\begin{smallmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{smallmatrix}\right], \mathcal{T}^*)_{\mathrm{on}}$ denote the dual output–nulling system. If the pair $\mathscr{U}, \mathscr{Y}$ is an admissible pair for Σ then*

(i) $\mathscr{Y}, \mathscr{U}$ is an E–admissible pair for Σ^*,

(ii) $\left[\begin{smallmatrix} x_* \\ \hline y_* \\ u_* \end{smallmatrix}\right]$ is a trajectory in $\mathcal{T}^*$ if, and only if, $\left[\begin{smallmatrix} x_* \\ E\left[\begin{smallmatrix} -u_* \\ y_* \end{smallmatrix}\right] \end{smallmatrix}\right]$ is a trajectory in $\tilde{\mathcal{T}}$, the set of trajectories of the E-derived $(\mathscr{Y}, \mathscr{U})$ system of Σ^*.

(iii) for $\left(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T}\right)_{\mathrm{iso}}$ *denoting the derived* $(\mathscr{U}, \mathscr{Y})$ *system of* Σ*, the following diagram commutes:*

$$\begin{tikzcd}
\left(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T}\right)_{\mathrm{dv}} \arrow[r, "\text{dual}"] \arrow[d, "\text{derived}"'] & \left(\left[\begin{smallmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{smallmatrix}\right], \mathcal{T}^*\right)_{\mathrm{on}} \arrow[d, "\text{derived}"] \\
\left(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T}\right)_{\mathrm{iso}} \arrow[r, "\text{dual}"] & \left(\left[\begin{smallmatrix} A_{\mathrm{D}}^* & C_{\mathrm{D}}^* \\ B_{\mathrm{D}}^* & D_{\mathrm{D}}^* \end{smallmatrix}\right], \mathcal{T}^*\right)_{\mathrm{iso}}.
\end{tikzcd}$$

Proof. (i): From Definition 4.8 we are required to prove that $(D^\dagger E)|_{\mathscr{U}} : \mathscr{U} \to \mathscr{V}$ is invertible. By (7.5) and the fact that E is self–adjoint and unitary we have

$$(D^\dagger E)|_{\mathscr{U}} = -(D^* E^2)|_{\mathscr{U}} = -(D^*)|_{\mathscr{U}} = -D_{\mathscr{U}}^*, \tag{7.7}$$

where the last equality is from (7.6). By Definition 4.8, $\mathscr{U}, \mathscr{Y}$ an admissible pair for Σ implies that $D_{\mathscr{U}}$ is invertible. Thus from (7.7) we see that $(D^\dagger E)|_{\mathscr{U}}$ is also invertible.

(ii) : The proof is very similar to that of Theorem 4.15. As the output–nulling case was not treated there, we do provide the proof here. The direct sum decomposition $\mathscr{W} = \mathscr{U} \oplus \mathscr{Y}$ and the surjectivity of E imply that any $w_* \in L^2_{\mathrm{loc}}(\mathbb{R}^+; \mathscr{W}^*)$ can be written as

$$w_* = -Eu_* + Ey_* = E \begin{bmatrix} -u_* \\ y_* \end{bmatrix}, \tag{7.8}$$

for $u_* \in L^2_{\mathrm{loc}}(\mathbb{R}^+; \mathscr{U})$ and $y_* \in L^2_{\mathrm{loc}}(\mathbb{R}^+; \mathscr{Y})$. We remark that when deriving input/output pairs u, y from the signal w of an output–nulling trajectory we have chosen to put a minus sign with the component that is the output (which is usually y, compare with (4.24)). In the dual case the input and output spaces interchange and so in the statement (ii) and (7.8) we have put the minus sign on u_*.

Let x_* and $w_* = E\left[\begin{smallmatrix} -u_* \\ y_* \end{smallmatrix}\right]$ denote the components of a trajectory in $\mathcal{T}^*$, so that

$$\begin{aligned}
\dot{x}_* &= A^* x - C^\dagger E \begin{bmatrix} -u_* \\ y_* \end{bmatrix} = A^* x_* + (C^\dagger E)|_{\mathscr{U}} u_* - (C^\dagger E)|_{\mathscr{Y}} y_*, \\
0 &= -B^* x_* + D^\dagger E \begin{bmatrix} -u_* \\ y_* \end{bmatrix} = -B^* x_* - (D^\dagger E)|_{\mathscr{U}} u_* + (D^\dagger E)|_{\mathscr{Y}} y_*
\end{aligned} \tag{7.9}$$

We have already seen that $(D^\dagger E)|_{\mathscr{U}}$ is invertible and hence from (7.9) we can obtain an input–state–output relation between y_* and u_*, namely

$$\begin{aligned}
\dot{x}_* &= (A^* - (C^\dagger E)|_{\mathscr{U}} (D^\dagger E)|_{\mathscr{U}}^{-1} B^*) x_* + (-(C^\dagger E)|_{\mathscr{Y}} + (C^\dagger E)|_{\mathscr{U}} (D^\dagger E)|_{\mathscr{U}}^{-1} (D^\dagger E)|_{\mathscr{Y}}) y_*, \\
u_* &= -(D^\dagger E)|_{\mathscr{U}}^{-1} B^* x_* + (D^\dagger E)|_{\mathscr{U}}^{-1} (D^\dagger E)|_{\mathscr{Y}} y_*,
\end{aligned} \tag{7.10}$$

so that x_* and $\left[\begin{smallmatrix} y_* \\ u_* \end{smallmatrix}\right]$ are the components of a trajectory in $\mathcal{T}^*_{\mathrm{iso}}$. To see the converse we reverse the above steps.

(iii): The bottom route through the diagram gives

$$\begin{aligned}
\begin{bmatrix} A & B \\ C & D \end{bmatrix} &\xrightarrow{\text{derived}} \begin{bmatrix} A - BD_{\mathscr{U}}^{-1} C_{\mathscr{U}} & BD_{\mathscr{U}}^{-1} \\ C_{\mathscr{Y}} - D_{\mathscr{Y}} D_{\mathscr{U}}^{-1} C_{\mathscr{U}} & D_{\mathscr{Y}} D_{\mathscr{U}}^{-1} \end{bmatrix} \\
&\xrightarrow{\text{dual}} \begin{bmatrix} A^* - (C_{\mathscr{U}})^* D_{\mathscr{U}}^{-*} B^* & (C_{\mathscr{Y}})^* - (C_{\mathscr{U}})^* D_{\mathscr{U}}^{-*} (D_{\mathscr{Y}})^* \\ D_{\mathscr{U}}^{-*} B^* & D_{\mathscr{U}}^{-*} (D_{\mathscr{Y}})^* \end{bmatrix}.
\end{aligned} \tag{7.11}$$

The top route through the diagram has effectively already been considered in (7.9) and (7.10). All that remains to verify is that the nodes in (7.10) and (7.11) are the same, but this follows by inspection using (7.5) and (7.6). That the trajectories coincide follows from the definitions of dual trajectories, Theorem 4.15 and the equality established in (ii). $\square$

8 Jointly dissipative systems

Here we consider dissipative state space systems which have dissipative duals (in the same sense), a notion we call jointly dissipative. Theorem 8.5 generalises the classical KYP Lemma and demonstrates that for driving–variable systems joint dissipativity is equivalent to a set of Lur'e equations having a positive self–adjoint solution. We use the extremal solutions of these equations to construct dissipative balanced truncations and thus perform model reduction by dissipative balanced truncation.

Definition 8.1. We say that a driving–variable or output–nulling system is jointly signal dissipative if it is signal dissipative and its dual is signal dissipative. We say that a driving–variable or output–nulling system is jointly state–signal dissipative if it is state–signal dissipative with respect to a positive, self–adjoint operator P and its dual is state–signal dissipative with respect to P^{-1}.

It is well known that the duals of bounded real and positive real input–state–output systems are again respectively bounded real and positive real. The following example, which is the continuous time version of [32, Example 5.5], demonstrates that the same is not true for driving–variable and output–nulling systems.

Example 8.2. Consider the following static driving–variable system

$$w(t) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} v(t) =: Dv(t), \quad t \geq 0, \tag{8.1}$$

where $\mathscr{V} = \mathbb{C}$, $\mathscr{X} = \{0\}$ and $\mathscr{W} = \mathbb{C}^3$. Introduce the indefinite inner–product

$$[\cdot,\cdot]_{\mathscr{W}} : \mathbb{C}^3 \times \mathbb{C}^3 \to \mathbb{C}, \quad [w, z]_{\mathscr{W}} := \left(\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \alpha \end{bmatrix}}_{=:Q} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}, \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \right)_{\mathbb{C}^3} = w_1\bar{z}_1 - w_2\bar{z}_2 + \alpha w_3 \bar{z}_3,$$

with $\alpha > 0$. The driving–variable system is signal dissipative with respect to $[\cdot,\cdot]_{\mathscr{W}}$ as for $t \geq 0$

$$\int_0^t [w(s), w(s)]_{\mathscr{W}}\, ds = \int_0^t |w_1(s)|^2 - |w_2(s)|^2 + \alpha |w_3(s)|^2\, ds = \int_0^t |v(s)|^2 - |v(s)|^2\, ds = 0,$$

by (8.1). However, the dual is *not* signal dissipative. The dual output–nulling node is $\left[\begin{smallmatrix} 0 & 0 \\ 0 & D^\dagger \end{smallmatrix}\right]$ so that the trajectories w_* satisfy

$$0 = D^\dagger w_*(t) = \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} w_*^{(1)}(t) \\ w_*^{(2)}(t) \\ w_*^{(3)}(t) \end{bmatrix}, \quad t \geq 0 \quad \Rightarrow \quad w_*^{(1)}(t) = w_*^{(2)}(t), \quad t \geq 0, \tag{8.2}$$

where we have used the fact that $D^\dagger = -D^T Q$, for D^T the complex conjugate transpose of D. Therefore

$$\begin{aligned} \int_0^t [w_*(s), w_*(s)]_{\mathscr{W}^*}\, ds &= \int_0^t - \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \alpha \end{bmatrix} \begin{bmatrix} w_*^{(1)}(s) \\ w_*^{(2)}(s) \\ w_*^{(3)}(s) \end{bmatrix}, \begin{bmatrix} w_*^{(1)}(s) \\ w_*^{(2)}(s) \\ w_*^{(3)}(s) \end{bmatrix} \right) ds \\ &= \int_0^t -|w_*^{(1)}(s)|^2 + |w_*^{(2)}(s)|^2 - \alpha |w_*^{(1)}(s)|^2\, ds = -\int_0^t \alpha |w_*^{(3)}(s)|^2\, ds \leq 0. \end{aligned} \tag{8.3}$$

From (8.2) we see that the component $w_*^{(3)}$ of a dual trajectory w_* is arbitrary. Therefore, from (8.3) it follows that the dual system is *not* signal–dissipative.

The main results of this section address joint admissibility for driving–variable systems.

Theorem 8.3. *Let $\mathscr{W}$ denote the signal space of a jointly signal dissipative driving–variable system. Then every fundamental decomposition of $\mathscr{W}$ is admissible.*

The proof of Theorem 8.3 makes use of the following elementary lemma; for a proof see [40, Lemma 3.5.3].

Lemma 8.4. *Let $\Sigma = ([\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}], \mathcal{T})_{\mathrm{dv}}$ denote a driving–variable system with indefinite inner–product signal space $\mathscr{W}$ and let $\Sigma^* = (\begin{bmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{bmatrix}, \mathcal{T}^*)_{\mathrm{on}}$ denote the dual system.*

(i) If Σ is signal dissipative then $\operatorname{im} D$ is nonnegative in $\mathscr{W}$.

(ii) If Σ^* is signal dissipative then $\ker D^\dagger$ is nonnegative in $\mathscr{W}^*$.

(iii) For $\mathscr{S} \subset \mathscr{W}$, $\mathscr{S}$ is nonpositive in $\mathscr{W}$ if, and only if, $\mathscr{S}$ is nonnegative in $\mathscr{W}^*$.

(iv) $(\operatorname{im} D)^{[\perp]} = \ker D^\dagger$.

Proof of Theorem 8.3. Let $\mathscr{W} = \mathscr{W}_+[+] - \mathscr{W}_-$ denote a fundamental decomposition of $\mathscr{W}$. From Lemma 8.4 (i), $\operatorname{im} D$ is nonnegative in $\mathscr{W}$ and combining parts $(ii) - (iv)$ implies that $(\operatorname{im} D)^{[\perp]}$ is nonpositive in $\mathscr{W}$. From Lemma 5.9 (ii), we see that $\operatorname{im} D$ is maximal nonnegative. Thus by Lemma 5.9 (i), there is a linear contraction $T : \mathscr{W}_+ \to \mathscr{W}_-$ such that

$$\operatorname{im} D = \mathcal{G}(T),$$

and now it follows from Lemma 4.19 (with $T = \tilde{D}$) that the pair $\mathscr{W}_+, \mathscr{W}_-$ is admissible. □

The next result can be seen as a generalisation of the Bounded Real and Positive Real Lemmas and characterised jointly dissipative driving–variable systems.

Theorem 8.5. *Given a minimal driving–variable system $\Sigma = ([\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}], \mathcal{T})_{\mathrm{dv}}$, with indefinite inner–product signal space $\mathscr{W}$, the following are equivalent:*

(i) Σ is jointly signal dissipative.

(ii) Σ is jointly state–signal dissipative.

(iii) The signature operator E of $\mathscr{W}$ satisfies $\sigma_+(E) = \dim(\operatorname{im} D)$ and there exists a positive, self–adjoint operator P on $\mathscr{X}$ and operators $M : \mathscr{X} \to \mathscr{W}$, $N : \mathscr{V} \to \mathscr{W}$ satisfying the indefinite KYP Lur'e equations

$$A^*P + PA - C^*EC = -M^*M, \tag{8.4a}$$

$$PB - C^*ED = -M^*N, \tag{8.4b}$$

$$D^*ED = N^*N. \tag{8.4c}$$

In addition, if any of the above hold then there exist positive, self–adjoint solutions P_m, P_M of (8.4) such that every self–adjoint solution P of (8.4) satisfies $0 < P_m \le P \le P_M$. The extremal operators P_m, P_M are the optimal cost operators of the indefinite optimal control problems, namely:

$$\langle P_M x_0, x_0\rangle_{\mathscr{X}} = \inf_{\substack{w \in L^2(\mathbb{R}^-;\mathscr{W}) \\ x(0)=x_0}} \int_{\mathbb{R}^-} [w(s), w(s)]_{\mathscr{W}}\, ds, \tag{8.5a}$$

$$-\langle P_m x_0, x_0\rangle_{\mathscr{X}} = \inf_{\substack{\mathcal{T}(x_0) \\ w \in L^2(\mathbb{R}^+;\mathscr{W})}} \int_{\mathbb{R}^+} [w(s), w(s)]_{\mathscr{W}}\, ds. \tag{8.5b}$$

The minimisation problems (8.5a) and (8.5b) are subject to the driving–variable node $[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}]$ over $\mathbb{R}^-$ and $\mathbb{R}^+$ respectively.

We provide some comments on the above result.

Remark 8.6. (i) By Theorem 8.5 the two notions of dissipativity considered are equivalent for minimal jointly dissipative driving–variable systems. There is therefore no ambiguity in calling such systems simply jointly dissipative.

(ii) The assumption that $\sigma_+(E) = \dim(\operatorname{im} D)$ is essential in proving the implication $(iii) \Rightarrow (ii)$. Although existence of a positive, self–adjoint solution P to the indefinite KYP Lur'e equations (8.4) does imply state–signal dissipativity of the system, it does not necessarily imply state–signal dissipativity of the dual system. The assumption that $\sigma_+(E) = \dim(\operatorname{im} D)$ also features in behavorial theory, and is a property referred to as *liveness* (see [41, p.56]). As we have already seen, for our purposes it ensures that the dual of a system is dissipative when the system itself is dissipative.

(iii) We use the already established Bounded Real Lemma to prove Theorem 8.5 instead of deriving it from first principles. However, that said, the Bounded Real and Positive Lemmas are special cases of Theorem 8.5.

(iv) A consequence of the proof of Theorem 8.5 is that the operators M, N in (8.4) map into $\mathscr{U} \subseteq \mathscr{W}$, the nonnegative component of some fundamental decomposition $\mathscr{W}$. In other words we can find a rank minimising (see for example [47]) solution (P, M, N) of (8.4).

Proof of Theorem 8.5: (i) $\Rightarrow$ *(iii)* : Let $\mathscr{U}, \mathscr{Y}$ denote a fundamental decomposition of $\mathscr{W}$ which by Theorem 8.3 is an admissible pair. From Corollary 5.10 it follows that

$$\dim(\operatorname{im} D) = \dim \mathscr{U} = \sigma_+(E).$$

We recall that with respect to the fundamental decomposition $\mathscr{U}, \mathscr{Y}$ the signature operator E has the block diagonal form

$$\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} : \begin{bmatrix} \mathscr{U} \\ \mathscr{Y} \end{bmatrix} \to \begin{bmatrix} \mathscr{U} \\ \mathscr{Y} \end{bmatrix}.$$

Let $\Sigma_{\mathrm{D}} = (\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ denote the derived $(\mathscr{U}, \mathscr{Y})$ system, which by Lemma 4.24 is minimal and bounded real by Theorem 6.4. Hence, by the Bounded Real Lemma, there are linear operators $K : \mathscr{X} \to \mathscr{U}$, $W : \mathscr{U} \to \mathscr{U}$ and self–adjoint positive P on $\mathscr{X}$ satisfying the bounded real Lur'e equations (3.2) (with realisation $\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right]$). We prove that the triple (P, K, W) solving (3.2) gives a solution (P, M, N) of (8.4) where

$$M := K + WC_{\mathscr{U}} : \mathscr{X} \to \mathscr{U} \subseteq \mathscr{W}, \quad N := WD_{\mathscr{U}} : \mathscr{V} \to \mathscr{U} \subseteq \mathscr{W}. \tag{8.6}$$

Expanding first equation (3.2c) we obtain

$$I - D_{\mathrm{D}}^* D_{\mathrm{D}} = I - D_{\mathscr{U}}^{-*} D_{\mathscr{Y}}^* D_{\mathscr{Y}} D_{\mathscr{U}}^{-1} = W^* W \tag{8.7}$$

$$\Rightarrow \quad D^* E D = D_{\mathscr{U}}^* D_{\mathscr{U}} - D_{\mathscr{Y}}^* D_{\mathscr{Y}} = (WD_{\mathscr{U}})^* (WD_{\mathscr{U}}) = N^* N, \tag{8.8}$$

where N is as in (8.6). From equations (3.2b) and (8.7) we obtain

$$\begin{aligned} PB - C^* ED &= PB_{\mathrm{D}} D_{\mathscr{U}} - C_{\mathscr{U}}^* D_{\mathscr{U}} + C_{\mathscr{Y}}^* D_{\mathscr{Y}} = (PB_{\mathrm{D}} + C_{\mathrm{D}}^* D_{\mathrm{D}}) D_{\mathscr{U}} - C_{\mathscr{U}}^* W^* W D_{\mathscr{U}} \\ &= -(K + WC_{\mathscr{U}})^* W D_{\mathscr{U}} = -M^* N. \end{aligned} \tag{8.9}$$

Finally, from (3.2a) we infer that

$$\begin{aligned} A^* P + PA - C^* EC &= - K^* K + C_{\mathscr{U}}^* D_{\mathscr{U}}^{-*} B^* P + PBD_{\mathscr{U}}^{-1} C_{\mathscr{U}} - C_{\mathscr{U}}^* D_{\mathscr{U}}^{-*} D_{\mathscr{Y}}^* D_{\mathscr{Y}} D_{\mathscr{U}}^{-1} C_{\mathscr{U}} - C_{\mathscr{U}}^* C_{\mathscr{U}} \\ &= - (K + WC_{\mathscr{U}})^* (K + WC_{\mathscr{U}}) = -M^* M. \end{aligned} \tag{8.10}$$

Equations (8.8), (8.9) and (8.10) are (8.4a), (8.4b) and (8.4c) respectively.

(iii) $\Rightarrow$ *(ii)* : We can rewrite the equations (8.4) as

$$\mathcal{M} := \begin{bmatrix} A^* P + PA - C^* EC & PB - C^* ED \\ B^* P - D^* EC & -D^* ED \end{bmatrix} = - \begin{bmatrix} M^* M & M^* N \\ N^* M & N^* N \end{bmatrix} = - \begin{bmatrix} M^* \\ N^* \end{bmatrix} \begin{bmatrix} M & N \end{bmatrix} \leq 0. \tag{8.11}$$

Given a trajectory $\left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \in \mathcal{T}$, let $v \in L^2_{\text{loc}}(\mathbb{R}^+;\mathscr{V})$ denote a corresponding driving–variable. For $t \geq 0$ from (8.11) we have

$$\int_0^t \left\langle \mathcal{M}\begin{bmatrix} x(s) \\ v(s) \end{bmatrix}, \begin{bmatrix} x(s) \\ v(s) \end{bmatrix} \right\rangle ds \leq 0. \tag{8.12}$$

Unravelling inequality (8.12) and using (4.10) gives

$$\int_0^t \frac{d}{ds}\langle Px(s), x(s)\rangle_{\mathscr{X}}\, ds \leq \int_0^t \langle Ew(s), w(s)\rangle_{\mathscr{W}}\, ds = \int_0^t [w(s), w(s)]_{\mathscr{W}}\, ds.$$

establishing that Σ is state–signal dissipative.

It remains to prove that the dual output–nulling system Σ^* is state–signal dissipative with respect to P^{-1}. By Corollary 5.10, $\sigma_+(E) = \dim(\operatorname{im} D)$ implies that any fundamental decomposition $\mathscr{U}$, $\mathscr{Y}$ of $\mathscr{W}$ is admissible. The equations (8.4) collapse to the bounded real Lur'e equations for the derived $(\mathscr{U},\mathscr{Y})$ system, which thus have a positive, self–adjoint solution P. We infer that Σ_{D} is bounded real, and hence so is the dual input–state–output system Σ^*_{D}. As discussed in Section 3, it follows that P^{-1} solves the dual bounded real Lur'e equations (3.5). Hence part (*iii*) of the Bounded Real Lemma applied to Σ^*_{D} implies that for state $x_* \in C(\mathbb{R}^+;\mathscr{X})$, input $y_* \in L^2_{\text{loc}}(\mathbb{R}^+;\mathscr{Y})$ and output $u_* \in L^2_{\text{loc}}(\mathbb{R}^+;\mathscr{U})$ of Σ^*_{D}

$$\langle P^{-1}x_*(t), x_*(t)\rangle_{\mathscr{X}} - \langle P^{-1}x_*(0), x_*(0)\rangle_{\mathscr{X}} \leq \int_0^t \|y_*(s)\|^2_{\mathscr{Y}} - \|u_*(s)\|^2_{\mathscr{U}}\, ds, \quad t \geq 0. \tag{8.13}$$

Proposition 7.9 (*ii*) gives the relation between trajectories of Σ^* and Σ^*_D. In particular, every signal w_* of Σ^* (with state x_*) can be decomposed as

$$E\begin{bmatrix} -u_* \\ y_* \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}\begin{bmatrix} -u_* \\ y_* \end{bmatrix} = \begin{bmatrix} -u_* \\ -y_* \end{bmatrix},$$

where y_*, u_* is an input/output pair for Σ^*_D also with state x_*. Therefore,

$$\begin{aligned} \int_0^t [w_*(s), w_*(s)]_{\mathscr{W}^*}\, ds &= \int_0^t -\langle w_*(s), Ew_*(s)\rangle_{\mathscr{W}}\, ds \\ &= \int_0^t -\left\langle \begin{bmatrix} -u_*(s) \\ -y_*(s) \end{bmatrix}, \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}\begin{bmatrix} -u_*(s) \\ -y_*(s) \end{bmatrix} \right\rangle_{\mathscr{U}\times\mathscr{Y}} ds \\ &= \int_0^t \|y_*(s)\|^2_{\mathscr{Y}} - \|u_*(s)\|^2_{\mathscr{Y}}\, ds. \end{aligned} \tag{8.14}$$

Combining (8.13) and (8.14) we conclude that Σ^* is state–signal dissipative with respect to P^{-1}, as required.

(*ii*) $\Rightarrow$ (*i*) : This implication is trivial.

Finally, suppose any (hence all) of (*i*)-(*iii*) above hold, any one of which imply that we can choose an admissible fundamental decomposition of the signal space where the derived input–state–output system Σ_{D} is bounded real. The proof of Theorem 8.5 so far has established a one–to–one correspondence between triples (P, K, W) solving (3.2) for Σ_{D} and triples (P, M, N) solving (8.4) for Σ.

The Bounded Real Lemma for Σ_{D} implies that there exist extremal positive self–adjoint operators P_m, P_M to (3.2), which by the above correspondence are thus extremal positive self–adjoint solutions of (8.4). The operators P_M and P_m are the optimal cost operators in (3.4a) and (3.4b) respectively for Σ_{D}. That P_M and P_m are also the optimal cost operators in (8.5a) and (8.5b) respectively now follows from (3.4a) and (3.4b), the equivalence of trajectories Theorem 4.15 and the equality of the scattering passive supply rate and indefinite inner product by Theorem 6.4. $\square$

9 Dissipative balanced approximations

In this section we define the dissipative balanced truncation of a minimal jointly dissipative driving–variable system and thus carry out model reduction by balanced truncation in a framework free from inputs and outputs. The established connections between driving–variable and input–state–output systems enable us to see existing bounded real and positive real balanced truncation as special cases.

9.1 Dissipative balanced truncation

Here the quantities to be balanced are the extremal solutions from the Indefinite KYP Lemma.

Definition 9.1. A minimal, jointly dissipative driving–variable system is called dissipative balanced or in dissipative balanced co–ordinates if

$$P_m = P_M^{-1} =: \Pi, \tag{9.1}$$

where P_m and P_M are the extremal solutions of the KYP Lur'e equations (8.4). The dissipative singular values, which we denote by $(\sigma_k)_{k=1}^m$, are the nonnegative square roots of the eigenvalues of the product $P_m P_M^{-1}$. The dissipative singular values are ordered such that $\sigma_k > \sigma_{k+1} > 0$ for each k and we let r_k denote the (geometric) multiplicity of σ_k.

Proposition 9.2. *Let $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ denote a minimal jointly dissipative driving–variable system. Then there exists an invertible operator $T : \mathscr{X} \to \mathscr{X}$, which we call a dissipative balancing transformation, such that the similarity transformed system $(\left[\begin{smallmatrix} T^{-1}AT & T^{-1}B \\ CT & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ is dissipative balanced.*

Proof. This is an application of [21, Lemma 7.3]. □

Definition 9.3. Let $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ denote a minimal jointly dissipative driving–variable system in dissipative balanced co–ordinates. Let $(\sigma_i)_{i=1}^m$ denote the dissipative singular values, with multiplicities r_i. For $r < m$ let $\mathscr{X}_r$ denote the sum of the first r eigenspaces of Π, with corresponding orthogonal projection $P_{\mathscr{X}_r}$. Define the operators

$$\Pi_1 = P_{\mathscr{X}_r}\Pi|_{\mathscr{X}_r}, \qquad A_{11} = P_{\mathscr{X}_r}A|_{\mathscr{X}_r}, \qquad B_1 = P_{\mathscr{X}_r}B, \qquad C_1 = C|_{\mathscr{X}_r}. \tag{9.2}$$

Let $\mathcal{T}_r$ denote the trajectories corresponding to the driving–variable node $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$. We call the driving–variable system $(\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right], \mathcal{T}_r)_{\mathrm{dv}}$ the dissipative balanced truncation (of order $\sum_{i=1}^r r_i$) of $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$.

Definition 9.4. Let $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ denote a minimal input–state–output system that is signal dissipative with respect to an indefinite inner–product with signature operator E that satisfies $\sigma_+(E) = \dim \mathscr{U}$. We say that the input–state–output system $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ is dissipative balanced if the jointly dissipative driving–variable system

$$\left(\left[\begin{array}{c|c} A & B \\ \hline 0 & I \\ C & D \end{array}\right], \mathcal{T}\right)_{\mathrm{dv}},$$

is dissipative balanced in the sense of Definition 9.1. We call the input–state–output system corresponding to the node $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$, where A_{11}, B_1 and C_1 are as in (9.2) the dissipative balanced truncation (of order $\sum_{i=1}^r r_i$) of $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$.

Lemma 9.5. *Let $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ denote a driving–variable system with admissible pair $\mathscr{U}, \mathscr{Y}$ and let $T \in B(\mathscr{X})$ be invertible. Letting $(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ denote the derived $(\mathscr{U}, \mathscr{Y})$ system the following diagram commutes*

$$\begin{array}{ccc} (\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}} & \xrightarrow{\text{transform}} & (\left[\begin{smallmatrix} T^{-1}AT & T^{-1}B \\ CT & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}} \\ {\scriptstyle\text{derived}}\big\downarrow & & \big\downarrow{\scriptstyle\text{derived}} \\ (\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}} & \xrightarrow{\text{transform}} & (\left[\begin{smallmatrix} T^{-1}A_{\mathrm{D}}T & T^{-1}B_{\mathrm{D}} \\ C_{\mathrm{D}}T & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}. \end{array}$$

If additionally $([\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}], \mathcal{T})_{\mathrm{dv}}$ *is minimal, jointly dissipative and dissipative balanced then the following diagram commutes*

$$\begin{tikzcd} ([\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}], \mathcal{T})_{\mathrm{dv}} \arrow[r, "\text{truncate}"] \arrow[d, "\text{derived}"'] & ([\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}], \mathcal{T}_r)_{\mathrm{dv}} \arrow[d, "\text{derived}"] \\ ([\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}], \mathcal{T})_{\mathrm{iso}} \arrow[r, "\text{truncate}"] & ([\begin{smallmatrix} (A_{\mathrm{D}})_{11} & (B_{\mathrm{D}})_1 \\ (C_{\mathrm{D}})_1 & D_{\mathrm{D}} \end{smallmatrix}], \mathcal{T}_r)_{\mathrm{iso}}. \end{tikzcd}$$

As such the derived $(\mathscr{U}, \mathscr{Y})$ *system of a dissipative balanced truncation driving–variable system is the same as taking the derived* $(\mathscr{U}, \mathscr{Y})$ *system of the original driving–variable system and then dissipative balancing and truncating the resulting input–state–output system.*

Proof. Consider the first diagram. The bottom route through the diagram states that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \xrightarrow{\text{derived}} \begin{bmatrix} A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}} & BD_{\mathscr{U}}^{-1} \\ C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix} \xrightarrow{\text{transform}} \begin{bmatrix} T^{-1}(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})T & T^{-1}BD_{\mathscr{U}}^{-1} \\ (C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}})T & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix}. \tag{9.3}$$

The top route through the diagram states that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \xrightarrow{\text{transform}} \begin{bmatrix} T^{-1}AT & T^{-1}B \\ CT & D \end{bmatrix} \xrightarrow{\text{derived}} \begin{bmatrix} T^{-1}AT - T^{-1}BD_{\mathscr{U}}^{-1}(CT)_{\mathscr{U}} & T^{-1}BD_{\mathscr{U}}^{-1} \\ (CT)_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}(C_{\mathscr{U}})T & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix}. \tag{9.4}$$

That (9.3) and (9.4) are the same follows by inspection and the fact that $(CT)_{\mathscr{U}} = \pi_{\mathscr{U}}^{\mathscr{Y}} CT = C_{\mathscr{U}}T$ (and similarly for $\mathscr{Y}$ instead of $\mathscr{U}$). We now prove that the second diagram commutes. The bottom route gives

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \xrightarrow{\text{derived}} \begin{bmatrix} A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}} & BD_{\mathscr{U}}^{-1} \\ C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix} \xrightarrow{\text{truncate}} \begin{bmatrix} P_{\mathscr{X}_r}(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})|_{\mathscr{X}_r} & P_{\mathscr{X}_r}BD_{\mathscr{U}}^{-1} \\ (C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}})|_{\mathscr{X}_r} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix}. \tag{9.5}$$

The top route through the diagram states that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \xrightarrow{\text{truncate}} \begin{bmatrix} P_{\mathscr{X}_r}A|_{\mathscr{X}_r} & P_{\mathscr{X}_r}B \\ C|_{\mathscr{X}_r} & D \end{bmatrix} \xrightarrow{\text{derived}} \begin{bmatrix} P_{\mathscr{X}_r}A|_{\mathscr{X}_r} - P_{\mathscr{X}_r}BD_{\mathscr{U}}^{-1}(C|_{\mathscr{X}_r})_{\mathscr{U}} & P_{\mathscr{X}_r}BD_{\mathscr{U}}^{-1} \\ (C|_{\mathscr{X}_r})_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}(C|_{\mathscr{X}_r})_{\mathscr{U}} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix}. \tag{9.6}$$

That (9.5) and (9.6) are the same again follows by inspection, the facts that restriction and projection are linear and that $(C|_{\mathscr{X}_r})_{\mathscr{U}} = \pi_{\mathscr{U}}^{\mathscr{Y}} C|_{\mathscr{X}_r} = (C_{\mathscr{U}})|_{\mathscr{X}_r}$ (and similarly for $\mathscr{Y}$ instead of $\mathscr{U}$). ☐

Corollary 9.6. *Let* Σ *denote a minimal jointly dissipative driving–variable system, let* $\mathscr{U}$, $\mathscr{Y}$ *denote an admissible pair and let* Σ_{D} *denote the derived* $(\mathscr{U}, \mathscr{Y})$ *system.*

(i) If Σ_{D} *is bounded real (positive real) then the dissipative singular values of* Σ *are precisely the bounded real singular values (positive real singular values) of* Σ_{D}.

(ii) If Σ *is dissipative balanced and* Σ_{D} *is bounded real (positive real) then* Σ_{D} *is bounded real balanced (positive real balanced).*

(iii) Let Σ_r *denote the dissipative balanced truncation of order* $\sum_{i=1}^{r} r_i$ *(using the notation of Definition 9.3). If* Σ_{D} *is bounded real (positive real) then the* $(\mathscr{U}, \mathscr{Y})$ *derived system of* Σ_r *is the bounded real balanced truncation (positive real balanced truncation) of order* $\sum_{i=1}^{r} r_i$ *of* Σ_{D}.

Proof. Claims (*i*) and (*ii*) follow from the indefinite KYP Lemma and the Bounded Real Lemma (Positive Real Lemma). Specifically the extremal solutions P_m and P_M of the KYP Lur'e equations (8.4) of Σ *are* the extremal solutions of the bounded real Lur'e equations (3.2) (positive real Lur'e equations (3.11)) of the $(\mathscr{U}, \mathscr{Y})$ derived system. Claim (*iii*) then follows from claims (*i*) and (*ii*) and the commuting diagrams in Lemma 9.5. ☐

Corollary 9.7. *Let* Σ *denote a minimal jointly dissipative driving-variable system. Then for every* r *the dissipative balanced truncation* Σ_r *is jointly dissipative.*

Proof. This follows from the indefinite KYP Lemma and Corollary 9.6. □

We are now in position to prove our main result, stated in Section 2, which is an error bound in the gap metric for dissipative balanced truncation.

Proof of Theorem 2.1: We make use of Theorem 3.7, an error bound for bounded real input–state–output systems. Choose $r < m$ and a fundamental decomposition $\mathscr{U}, \mathscr{Y}$ of the signal space, which is admissible by Theorem 8.3. The derived $(\mathscr{U}, \mathscr{Y})$ system is bounded real by Theorem 6.4. Denote the transfer function and input–output map of this system by G and $\mathfrak{D}_G$ respectively. By Corollary 9.6 (i) the bounded real singular values of G are precisely the dissipative singular values of Σ. By Corollary 9.6 (iii) the bounded real balanced truncation, with transfer function G_r and input–output map $\mathfrak{D}_{G_r}$, is the derived $(\mathscr{U}, \mathscr{Y})$ system of Σ_r. The error bound (3.8)

$$\|G - G_r\|_{H^\infty} \leq 2 \sum_{i=r+1}^{m} \sigma_i,$$

from Theorem 3.7 then applies. Let $\mathcal{S}$ and $\mathcal{S}_r$ denote the sets of stable externally generated trajectories of Σ and Σ_r respectively. Since G is bounded real, it belongs to H^∞ and hence the input–output map $\mathfrak{D}_G$ is bounded $L^2(\mathbb{R}^+; \mathscr{U}) \to L^2(\mathbb{R}^+; \mathscr{Y})$. In this instance the conclusion of Corollary 4.17 can be strengthened to

$$\mathcal{S} = \mathcal{G}(\mathfrak{D}_G).$$

The same relation holds for the dissipative balanced truncation and the bounded real balanced truncation of the derived system (for the same reasons), namely

$$\mathcal{S}_r = \mathcal{G}(\mathfrak{D}_{G_r}).$$

Therefore

$$\hat{\delta}(\Sigma, \Sigma_r) := \hat{\delta}(\mathcal{S}, \mathcal{S}_r) = \hat{\delta}(\mathcal{G}(\mathfrak{D}_G), \mathcal{G}(\mathfrak{D}_{G_r})) = \hat{\delta}(\mathfrak{D}_G, \mathfrak{D}_{G_r}). \tag{9.7}$$

We also require the well known equality (see for example [48])

$$\|\mathfrak{D}_G - \mathfrak{D}_{G_r}\| = \|G - G_r\|_\infty. \tag{9.8}$$

Combining the bounds (3.8) and (4.31) with the equalities (9.7) and (9.8) yields the desired bound (2.9). □

The above bound holds for any admissible decomposition into inputs and outputs (since the gap metric is independent of such a splitting), and in particular for the positive real case. Therefore we obtain a gap metric error bound for stable positive real balanced truncation and, by using the equivalence between the H^∞ norm and the gap metric for stable systems, we get an H^∞ error bound for positive real balanced truncation. These results are formulated as Corollary 9.8 below, which extends Corollary 2.2. The gap metric error bound (2.10) (equivalently (9.9)) has been independently established by Timo Reis [8].

Corollary 9.8. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a positive real rational transfer function with positive real singular values $(\sigma_i)_{i=1}^m$ and for $r < m$ let J_r denote the positive real balanced truncation. Then the following bounds hold,*

$$\hat{\delta}(J, J_r) \leq 2 \sum_{i=r+1}^{m} \sigma_i, \tag{9.9}$$

and

$$\|J - J_r\|_{H^\infty} \leq 2 \min\left\{(1 + \|J\|_{H^\infty}^2)(1 + \|J_r\|_{H^\infty}), (1 + \|J\|_{H^\infty})(1 + \|J_r\|_{H^\infty}^2)\right\} \sum_{i=r+1}^{m} \sigma_i. \tag{9.10}$$

In inequality (9.9) we are abusing notation by writing $\hat{\delta}(J, J_r) = \hat{\delta}(\mathfrak{D}_J, \mathfrak{D}_{J_r})$, where $\mathfrak{D}_J$ and $\mathfrak{D}_{J_r}$ are the input–output maps corresponding to J and J_r respectively.

Remark 9.9. The H^∞ error bound (9.10) is not an *a priori* error bound, as it requires $\|J_r\|_\infty$, thus limiting its usefulness in practise.

Proof of Corollary 9.8: Let $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ denote a minimal realisation of J, from which we build a driving–variable system Σ as in Lemma 4.18, which is minimal and dissipative with respect to the indefinite inner–product induced by $E := \left[\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right]$ by Definition 6.3. Since the dual input–state–output system with node $\left[\begin{smallmatrix} A^* & C^* \\ B^* & D^* \end{smallmatrix}\right]$ is positive real, by Proposition 7.9 we see that Σ is jointly dissipative. By Corollary 9.6 (i) the dissipative singular values of Σ are precisely the positive real singular values of J and by part (iii) of that result the $(\mathscr{U}, \mathscr{U})$ derived dissipative balanced truncation of Σ is the positive real balanced truncation of J. The gap metric error bound for $\hat{\delta}(J, J_r)$ now follows from Theorem 2.1 and Theorem 4.15.

To prove the H^∞ bound we use the equivalence of the gap metric restricted to bounded, linear operators and the operator norm, and we refer the reader to [40, Corollary 3.6.9.] for the details. □

9.2 Singular perturbation approximation

So far we have considered model reduction by direct truncation. In this section we demonstrate that singular perturbation approximation is often another suitable method for model reduction of a driving–variable system.

Definition 9.10. Let $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ denote a minimal jointly dissipative driving–variable system in dissipative balanced co–ordinates. Let $(\sigma_i)_{i=1}^m$ denote the dissipative singular values, each with multiplicity r_i. For $s < m$ let $\mathscr{X}_s$ and $\mathscr{Z}_s$ denote the sum of the first s and last $m-s$ eigenspaces of Π respectively, with respective orthogonal projections $P_{\mathscr{X}_s}$ and $P_{\mathscr{Z}_s}$. Then with respect to the orthogonal decomposition $\mathscr{X} = \mathscr{X}_s \oplus \mathscr{Z}_s$, the operators A, B, C and Π split as

$$\Pi = \begin{bmatrix} P_{\mathscr{X}_s}\Pi|_{\mathscr{X}_s} & 0 \\ 0 & P_{\mathscr{Z}_s}\Pi|_{\mathscr{Z}_s} \end{bmatrix} = \begin{bmatrix} \Pi_1 & 0 \\ 0 & \Pi_2 \end{bmatrix}, \; A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \; B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \; C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}.$$

Assuming that A_{22}^{-1} exists define

$$\begin{aligned} A_{\mathrm{s}} &:= A_{11} - A_{12}A_{22}^{-1}A_{21}, \quad & B_{\mathrm{s}} &:= B_1 - A_{12}A_{22}^{-1}B_2, \\ C_{\mathrm{s}} &:= C_1 - C_2A_{22}^{-1}A_{21}, \quad & D_{\mathrm{s}} &:= D - C_2A_{22}^{-1}B_2. \end{aligned} \tag{9.11}$$

Let $\mathcal{T}_s$ denote the trajectories corresponding to the driving–variable node $\left[\begin{smallmatrix} A_{\mathrm{s}} & B_{\mathrm{s}} \\ C_{\mathrm{s}} & D_{\mathrm{s}} \end{smallmatrix}\right]$. We call the driving–variable system $(\left[\begin{smallmatrix} A_{\mathrm{s}} & B_{\mathrm{s}} \\ C_{\mathrm{s}} & D_{\mathrm{s}} \end{smallmatrix}\right], \mathcal{T}_s)_{\mathrm{dv}}$ the (dissipative) singular perturbation approximation (of order $\sum_{i=1}^s r_i$) of $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$.

Theorem 9.11. *Given a minimal jointly dissipative driving–variable system Σ let $(\sigma_i)_{i=1}^m$ denote the dissipative singular values with multiplicities r_i and for $s < m$ assume that the dissipative singular perturbation approximation Σ_s (of order $\sum_{i=1}^s r_i$) exists. Then*

$$\hat{\delta}(\Sigma, \Sigma_s) \leq 2\sum_{i=s+1}^m \sigma_i. \tag{9.12}$$

Proof. Let $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ denote a dissipative balanced driving–variable node of Σ. Choose a fundamental decomposition $\mathscr{U}, \mathscr{Y}$ of the signal space, which is admissible by Theorem 8.3. The derived $(\mathscr{U}, \mathscr{Y})$ system, denoted by Σ_{D}, and with transfer function G, is bounded real by Theorem 6.4. We let $\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right]$ denote the (minimal) derived input–state–output node of Σ_{D}. By Theorem 3.7 the components $(A_{\mathrm{D}})_{11}$ and $(A_{\mathrm{D}})_{22}$ of the bounded real balanced truncation are Hurwitz, and so

$$(A_{\mathrm{D}})_{22} = P_{\mathscr{Z}_s}A_{\mathrm{D}}|_{\mathscr{Z}_s} = P_{\mathscr{Z}_s}(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})|_{\mathscr{Z}_s} = A_{22} - B_2D_{\mathscr{U}}^{-1}C_{\mathscr{U},2}, \quad \text{is invertible,} \tag{9.13}$$

where $C_{\mathscr{U},2} = \pi^{\mathscr{Y}}_{\mathscr{U}} C_2$. Furthermore, the assumption that the singular perturbation approximation Σ_s exists implies that A_{22} is invertible and hence $D_{\mathscr{U}} - C_{\mathscr{U},2}A_{22}^{-1}B_2$ is well–defined. The combination of $D_{\mathscr{U}}$ invertible and (9.13) implies that

$$D_{\mathscr{U}} - C_{\mathscr{U},2}A_{22}^{-1}B_2, \quad \text{is invertible}; \tag{9.14}$$

the inverse can be written down using the well known block matrix inversion identity:

$$(D_{\mathscr{U}} - C_{\mathscr{U},2}A_{22}^{-1}B_2)^{-1} = D_{\mathscr{U}}^{-1} + D_{\mathscr{U}}^{-1}C_{\mathscr{U},2}(A_{22} - B_2D_{\mathscr{U}}^{-1}C_{\mathscr{U},2})^{-1}B_2D_{\mathscr{U}}^{-1}.$$

The combination of (9.13) and (9.14) implies that the following diagram commutes

$$\begin{array}{ccc} \left(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T}\right)_{\mathrm{dv}} & \xrightarrow{\text{spa}} & \left(\left[\begin{smallmatrix} A_{\mathrm{s}} & B_{\mathrm{s}} \\ C_{\mathrm{s}} & D_{\mathrm{s}} \end{smallmatrix}\right], \mathcal{T}_s\right)_{\mathrm{dv}} \\ \text{derived}\Big\downarrow & & \Big\downarrow\text{derived} \\ \left(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T}\right)_{\mathrm{iso}} & \xrightarrow{\text{spa}} & \left(\left[\begin{smallmatrix} (A_{\mathrm{D}})_{\mathrm{s}} & (B_{\mathrm{D}})_{\mathrm{s}} \\ (C_{\mathrm{D}})_{\mathrm{s}} & (D_{\mathrm{D}})_{\mathrm{s}} \end{smallmatrix}\right], \mathcal{T}_s\right)_{\mathrm{iso}} \end{array}. \tag{9.15}$$

Here spa denotes taking the singular perturbation approximation. The proof of (9.15) is a rather long, but elementary, series of calculations which we do not give here but can be found in [40, Appendix A].

The proof of the error bound now mirrors that of Theorem 2.1, only instead appealing to singular perturbation results instead of those for balanced truncation. The commuting diagram (9.15) implies that Corollary 9.6 (*iii*) applies with balanced truncation replaced by singular perturbation approximation. Specifically, the bounded real singular perturbation approximation of Σ_{D}, with transfer function G_s and input–output map $\mathfrak{D}_{G_s}$, is equal to the derived $(\mathscr{U}, \mathscr{Y})$ system of the singular perturbation approximation Σ_s of Σ. Therefore

$$\hat{\delta}(\Sigma, \Sigma_s) := \hat{\delta}(\mathcal{S}, \mathcal{S}_s) = \hat{\delta}(\mathcal{G}(\mathfrak{D}_G), \mathcal{G}(\mathfrak{D}_{G_s})) = \hat{\delta}(\mathfrak{D}_G, \mathfrak{D}_{G_s}) \leq \|G - G_s\|_{H^\infty} \leq 2\sum_{i=r+1}^{m} \sigma_i,$$

where the last inequality is proven in [16, Theorem 3]. □

Similar adaptations can be made to Corollary 9.8 to give the following error bounds for singular perturbation approximation of stable positive real input–state–output systems.

Corollary 9.12. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a positive real rational transfer function with positive real singular values $(\sigma_i)_{i=1}^m$ and for $s < m$ let J_s denote the singular perturbation approximation. Then the following bounds hold,*

$$\hat{\delta}(J, J_s) \leq 2\sum_{i=s+1}^{m} \sigma_i, \tag{9.16}$$

and

$$\|J - J_s\|_{H^\infty} \leq 2\min\left\{(1 + \|J\|_{H^\infty}^2)(1 + \|J_s\|_{H^\infty}), (1 + \|J\|_{H^\infty})(1 + \|J_s\|_{H^\infty}^2)\right\}\sum_{i=s+1}^{m} \sigma_i. \tag{9.17}$$

Proof. For $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ a minimal realisation of $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ we have that A_{22} is Hurwitz and hence invertible. This follows from Theorem 3.11 where there it is stated that A_{11} is Hurwitz, but the same arguments apply to A_{22}. Therefore the singular perturbation approximation of J exists. The proof proceeds identically to that of Corollary 9.8, only appealing to Theorem 9.11 instead of Theorem 2.1. Note that the input–state–output node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ and the driving–variable node of the the driving–variable system Σ 'built' from $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ as in Lemma 4.18 have the same A operator. Therefore the singular perturbation approximation of Σ exists and Theorem 9.11 applies. The bound (9.16) now follows from (9.12) and Theorem 4.15. The H^∞ bound (9.17) follows from (9.16) by arguing as in Corollary 9.8; that is, by using the equivalence of the gap metric restricted to bounded, linear operators and the operator norm. □

References

[1] B. C. Moore, “Principal component analysis in linear systems: controllability, observability, and model reduction,” *IEEE Trans. Automat. Control*, vol. 26, no. 1, pp. 17–32, 1981.

[2] D. F. Enns, “Model reduction with balanced realizations: An error bound and a frequency weighted generalization,” in *Proc. CDC*, pp. 127–132, 1984. doi: 10.1109/CDC.1984.272286.

[3] K. Glover, “All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds,” *Internat. J. Control*, vol. 39, no. 6, pp. 1115–1193, 1984.

[4] U. B. Desai and D. Pal, “A transformation approach to stochastic model reduction,” *IEEE Trans. Automat. Control*, vol. 29, no. 12, pp. 1097–1100, 1984.

[5] P. C. Opdenacker and E. A. Jonckheere, “A contraction mapping preserving balanced reduction scheme and its infinity norm error bounds,” *IEEE Trans. Circuits and Systems*, vol. 35, no. 2, pp. 184–189, 1988.

[6] S. Gugercin and A. C. Antoulas, “A survey of model reduction by balanced truncation and some new results,” *Internat. J. Control*, vol. 77, no. 8, pp. 748–766, 2004.

[7] X. Chen and J. T. Wen, “Positive realness preserving model reduction with $\mathcal{H}^\infty$ norm error bounds,” *IEEE Trans. Circuits Systems I Fund. Theory Appl.*, vol. 42, no. 1, pp. 23–29, 1995.

[8] T. Reis, “Private correspondence,” 2010.

[9] K. Zhou and J. C. Doyle, *Essentials of robust control*, vol. 104. Prentice Hall Upper Saddle River, NJ, 1998.

[10] L. Qiu and E. J. Davison, “Feedback stability under simultaneous gap metric uncertainties in plant and controller,” *Systems Control Lett.*, vol. 18, no. 1, pp. 9–22, 1992.

[11] D. Z. Arov and O. J. Staffans, “State/signal linear time-invariant systems theory. I. Discrete time systems,” in *The state space method generalizations and applications*, vol. 161 of *Oper. Theory Adv. Appl.*, pp. 115–177, Basel: Birkhäuser, 2006.

[12] M. Kurula and O. J. Staffans, “Well-posed state/signal systems in continuous time,” *Complex Anal. Oper. Theory*, vol. 4, no. 2, pp. 319–390, 2010.

[13] M. B. Ha, *Model Reduction in a Behavioral Framework*. PhD thesis, Groningen, 2009.

[14] H. L. Trentelman, “Bounded real and positive real balanced truncation using Sigma-normalised coprime factors,” *Systems Control Lett.*, vol. 58, no. 12, pp. 871 – 879, 2009.

[15] H. L. Trentelman, “A behavioral characterization of the positive real and bounded real characteristic values in balancing,” *Systems Control Lett.*, vol. 60, no. 1, pp. 7–14, 2011.

[16] G. Muscato, G. Nunnari, and L. Fortuna, “Singular perturbation approximation of bounded real balanced and stochastically balanced transfer matrices,” *Internat. J. Control*, vol. 66, no. 2, pp. 253–269, 1997.

[17] G. Li, G. Weiss, M. Mueller, S. Townley, and M. R. Belmont, “Wave energy converter control by wave prediction and dynamic programming,” *Renewable Energy*, vol. 48, pp. 392–403, 2012.

[18] J. C. Willems, “The behavioral approach to open and interconnected systems: modeling by tearing, zooming, and linking,” *IEEE Control Syst. Mag.*, vol. 27, no. 6, pp. 46–99, 2007.

[19] J. W. Polderman and J. C. Willems, *Introduction to mathematical systems theory: A behavioral approach*, vol. 26 of *Texts in Applied Mathematics*. New York: Springer-Verlag, 1998.

[20] K. Morris and C. Navasca, “Approximation of low rank solutions for linear quadratic control of partial differential equations,” *Comput. Optim. Appl.*, vol. 46, no. 1, pp. 93–111, 2010.

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

[22] B. D. Anderson and S. Vongpanitlerd, *Network analysis and synthesis: a modern systems theory approach*. Prentice Hall, 1973.

[23] T. Reis and T. Stykel, "Positive real and bounded real balancing for model reduction of descriptor systems," *International Journal of Control*, vol. 83, no. 1, pp. 74–88, 2010.

[24] C. Guiver and M. R. Opmeer, "Bounded real and positive real balanced truncation for infinite-dimensional systems," *Math. Control Relat. Fields*, vol. 3, no. 1, pp. 83–119, 2013.

[25] J. T. Wen, "Time domain and frequency domain conditions for strict positive realness," *IEEE Trans. Automat. Control*, vol. 33, no. 10, pp. 988–992, 1988.

[26] W. Sun, P. P. Khargonekar, and D. Shim, "Solution to the positive real control problem for linear time-invariant systems," *IEEE Trans. Automat. Control*, vol. 39, no. 10, pp. 2034–2046, 1994.

[27] O. J. Staffans, "Passive and conservative continuous-time impedance and scattering systems. I. Well-posed systems," *Math. Control Signals Systems*, vol. 15, no. 4, pp. 291–315, 2002.

[28] J. C. Willems, "Dissipative dynamical systems. II. Linear systems with quadratic supply rates," *Arch. Rational Mech. Anal.*, vol. 45, pp. 352–393, 1972.

[29] P. Harshavardhana, E. A. Jonckheere, and L. M. Silverman, "Stochastic balancing and approximation-stability and minimality," *IEEE Trans. Automat. Control*, vol. 29, no. 8, pp. 744–746, 1984.

[30] C. Guiver and M. R. Opmeer, "A counterexample to "Positive realness preserving model reduction with $\mathscr{H}_\infty$ norm error bounds"," *IEEE Trans. Circuits Syst. I. Regul. Pap.*, vol. 58, no. 6, pp. 1410–1411, 2011.

[31] J. C. Willems, "Input-output and state-space representations of finite-dimensional linear time-invariant systems," *Linear Algebra Appl.*, vol. 50, pp. 581–608, 1983.

[32] D. Z. Arov and O. J. Staffans, "State/signal linear time-invariant systems theory. II. Passive discrete time systems," *Internat. J. Robust Nonlinear Control*, vol. 17, no. 5-6, pp. 497–548, 2007.

[33] D. Z. Arov and O. J. Staffans, "State/signal linear time-invariant systems theory. III. Transmission and impedance representations of discrete time systems," in *Operator theory, structured matrices, and dilations*, vol. 7 of *Theta Ser. Adv. Math.*, pp. 29–68, Theta, Bucharest, 2007.

[34] D. Z. Arov and O. J. Staffans, "State/signal linear time-invariant systems theory. IV. Affine and generalized transfer functions of discrete time systems," *Complex Anal. Oper. Theory*, vol. 1, no. 4, pp. 457–521, 2007.

[35] D. Z. Arov, M. Kurula, and O. J. Staffans, "Canonical state/signal shift realizations of passive continuous time behaviors," *Complex Anal. Oper. Theory*, vol. 5, no. 2, pp. 331–402, 2011.

[36] M. Kurula, *Towards input/output-free modelling of linear infinite-dimensional systems in continuous time*. PhD thesis, Åbo Akademi University, 2010. Available from: http://users.abo.fi/mkurula/pdf/phd-kurula.pdf.

[37] M. Kurula, "On passive and conservative state/signal systems," *Integral Equations Operator Theory*, vol. 67, no. 3, pp. 377–424,449, 2010.

[38] M. Kurula and O. J. Staffans, "Connections between classical and generalised trajectories of a state/signal system," *Complex Anal. Oper. Theory*, vol. 5, no. 2, pp. 403–422, 2011.

[39] R. F. Curtain and H. Zwart, *An introduction to infinite-dimensional linear systems theory*, vol. 21 of *Texts in Applied Mathematics*. New York: Springer-Verlag, 1995.

[40] C. Guiver, *Model Reduction by Balanced Truncation.* PhD thesis, University of Bath, 2012. Available from: http://opus.bath.ac.uk/32863/.

[41] J. C. Willems and H. L. Trentelman, "Synthesis of dissipative systems using quadratic differential forms. I," *IEEE Trans. Automat. Control*, vol. 47, no. 1, pp. 53–69, 2002.

[42] T. Kato, *Perturbation theory for linear operators.* Classics in Mathematics, Berlin: Springer-Verlag, 1995. Reprint of the 1980 edition.

[43] L. Dai, *Singular control systems*, vol. 118 of *Lecture Notes in Control and Information Sciences.* Berlin: Springer-Verlag, 1989.

[44] J. Bognár, *Indefinite inner product spaces.* New York: Springer-Verlag, 1974. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 78.

[45] I. Gohberg, P. Lancaster, and L. Rodman, *Matrices and indefinite scalar products*, vol. 8 of *Operator Theory: Advances and Applications.* Basel: Birkhäuser Verlag, 1983.

[46] I. Gohberg, P. Lancaster, and L. Rodman, *Indefinite linear algebra and applications.* Basel: Birkhäuser Verlag, 2005.

[47] D. J. Clements, B. D. O. Anderson, A. J. Laub, and J. B. Matson, "Spectral factorization with imaginary-axis zeros," *Linear Algebra Appl.*, vol. 250, pp. 225–252, 1997.

[48] G. Weiss, "Representation of shift-invariant operators on L^2 by H^∞ transfer functions: an elementary proof, a generalization to L^p, and a counterexample for L^∞," *Math. Control Signals Systems*, vol. 4, no. 2, pp. 193–203, 1991.