Optimal Hankel norm model reduction by truncation of trajectories

Berend Roorda
Dept. of Economics
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
roorda@kub.nl

Siep Weiland
Dept. of Elec. Engineering
Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven
The Netherlands
s.weiland@ele.tue.nl

Keywords

Optimal Hankel norm approximation, Balancing, Linear systems, ℓ_2-systems, Faddeev sequences.

Abstract

We show how optimal Hankel-norm approximations of dynamical systems allow for a straightforward interpretation in terms of system trajectories. It is shown that for discrete time single-input systems optimal reductions are obtained by cutting 'balanced trajectories', i.e., by disconnecting the past and future in the input-output pairs relating to left- and right singular vectors of the system. A self-contained proof of optimality is given, and formulas are derived in terms of Faddeev sequences. Some parallels with the literature are briefly indicated.

1 Introduction

Optimal Hankel-norm approximation of dynamical systems is one of the exceptional reduction techniques in linear systems theory with a well-described optimality property. The importance of the Hankel-norm model reduction problem therefore goes beyond the question of complexity reduction alone, as the optimality properties of reduced order models also give substantial insight in the intricate structure of linear systems. This insight, however, may be easily obscured by the complexity of the equations and algorithms which lead to reduced order models, and also by the common abstract language of operator theory in which this problem is usually formulated.

Most expositions, including the original work by Adamjan, Arov and Krein (which we will abbreviate as AAK) [1, 2], are in terms of (infinite) dimensional operators on Hardy spaces (See, e.g. [9] and the references therein), or in terms of (finite dimensional) state space realizations (See [4], [12, Chapter 8] and the references therein), with emphasis on continuous time systems.

The main purpose of this paper is to show that the construction and derivation of optimal Hankel-norm approximants is quite transparant on the level of system trajectories of discrete time systems. Indeed, we show that by suitably disconnecting the past and future of 'balanced trajectories', optimal approximants (in the Hankel sense) are obtained by removing the anti-causal part. We restrict the analysis to single-input systems, mainly for expository reasons. Continuous time systems are not addressed, but the results carry over to this case by applying the well-known bilinear transformation, cf. [4]. Hankel-norm reductions of infinite dimensional systems will not be considered in this paper.

The mathematical essence of the method which we propose here is not new, and we indicate some points in the existing literature in which the same observations are made, albeit in quite a different setting and with a different language. Therefore, as far as the construction of reduced order models is concerned, the paper does not address new paradigms, but its interest lies in the approach we take. A completely self-contained derivation of results is given, in which no more mathematics is used than some basic geometric properties of square summable time series, together with a technical result that relates the causal degree of a system to its controllability gramian.

The work is related to [11], in which a similar construction is presented for reducing the complexity of a system in a behavioral framework, independent of input and output variables.

Notation

The space of square summable single-component trajectories on the time axis $\mathbb{Z}$ is denoted by ℓ_2, while ℓ_2^p denotes the set of vector valued time series whose p components belong to ℓ_2. The superscripts $+$ and $-$, e.g. in ℓ_2^- and ℓ_2^+, denote the restriction of the time axis to the *past* (i.e. $\mathbb{Z}_- := \{t \mid t < 0\}$) and to the *future* ($\mathbb{Z}_+ := \{t \mid t \geq 0\}$). The symbols $\|\cdot\|$, $\langle\cdot,\cdot\rangle$ and $\perp$ denote resp. the usual norm, inner product and orthogonality in any of these spaces. Trajectories are denoted in boldface, and zero time series (of any length) are denoted as $\mathbf{0}$. For $\tau \in \mathbb{Z}$, the τ-shift operator is denoted by σ^τ, with $(\sigma^\tau \mathbf{f})(t) := \mathbf{f}(t+\tau)$. Singular values are also denoted by σ, but never without a subscript. The symbol $\wedge$ denotes concatenation of time series at time $t = 0$. The k-th unit vector in $\mathbb{R}^{\cdot\times 1}$ is denoted by e_k, and I_n denotes the identity matrix in $\mathbb{R}^{n\times n}$.

2 Model Reduction

Consider a stable single-input system Σ_{full} with rational transfer function $G \in \mathbb{R}(z)^{p\times 1}$ of McMillan degree n,

$$G(z) = \frac{N(z)}{d(z)} = \frac{N_0 + N_1 z + \ldots + N_n z^n}{d_0 + d_1 z + \ldots + d_n z^n} \qquad (1)$$

We restrict attention to square summable signals, and define

$$\Sigma(G) := \{(\mathbf{u}, \mathbf{y}) \in \ell_2 \times \ell_2^p \mid \mathbf{y} = G(\sigma^{-1})(\mathbf{u})\} \qquad (2)$$

where the equation $\mathbf{y} = G(\sigma^{-1})\mathbf{u}$ is to be interpreted as the difference equation $d(\sigma^{-1})\mathbf{y} = N(\sigma^{-1})\mathbf{u}$. Recall that $(\sigma^{-1}f)(t) = f(t-1)$. By assumption, the polynomial d is Hurwitz, but (2) is equally well defined if d is not Hurwitz[1].

A particularly useful characterization of $\Sigma(G)$ is in terms of its *shortest lag* trajectory,

$$(\mathbf{d}, \mathbf{N}) \in \Sigma(G) \qquad (3)$$

with

$$\mathbf{d}(t) := \begin{cases} d_t & \text{if } t \in [0,n] \\ 0 & \text{otherwise} \end{cases}, \quad \mathbf{N}(t) := \begin{cases} N_t & \text{if } t \in [0,n] \\ 0 & \text{otherwise} \end{cases}.$$

It is well-known that no pair of different single-input systems of the form (2) share any non-zero trajectory (cf. e.g. [11]). Consequently, $(\mathbf{d}, \mathbf{N})$ uniquely specifies $\Sigma(G)$.

A (linear) state space system is a description of the form

$$\mathbf{x}(t+1) = A\mathbf{x}(t) + B\mathbf{u}(t) \qquad (4a)$$
$$\mathbf{y}(t) = C\mathbf{x}(t) + D\mathbf{u}(t). \qquad (4b)$$

We write $\Sigma(A,B,C,D)$ for the set of all square summable input-output pairs for which there exists a state sequence $\mathbf{x}$ such that (4) hold. Obviously,

$$\Sigma(A,B,C,D) = \Sigma(G)$$

if and only if $G(z) = C(Iz - A)^{-1}B + D$. State space representations can be constructed from the transfer function (1) by standard methods. (In Matlab this is performed by the command `[A,B,C,D]=tf2ss(N,d)` with `N` and `d` the (matrix) coefficients of N and d).

The degree $\delta(\Sigma)$ of the system $\Sigma = \Sigma(G)$ is defined as the McMillan degree of G, or, equivalently, as the dimension of state in a minimal state space representation of G. For single-input systems this is equal to the shortest lag of a non-zero input/output pair, cf. (3).

Model reduction involves the approximation of the system

$$\Sigma_{\text{full}} = \Sigma(G) = \Sigma(A,B,C,D)$$

by a system

$$\Sigma_{\text{red}} = \Sigma(\widehat{G}) = \Sigma(A',B',C',D')$$

with degree $\delta(\Sigma_{\text{red}}) = n' < n$. By the previous argument, this is equivalent to approximating the shortest lag trajectory $(\mathbf{d}, \mathbf{N}) \in \Sigma(G)$ by a non-zero trajectory $(\mathbf{d}', \mathbf{N}')$ of lag n'. Before we describe a method that optimizes an approximation criterion, we first briefly discuss a standard heuristic technique, based on balanced representations.

3 Balanced truncations

The first step in one of the best-known model reduction techniques amounts to bringing the state representation of Σ_{full} in balanced form, i.e., to apply a basis transformation of the state such that the observability and controllability gramian are both equal to a diagonal matrix. Precisely, $\Sigma(A,B,C,D)$ is balanced if

$$\begin{aligned} M &= A^\top M A + C^\top C \\ W &= A W A^\top + B B^\top \\ W = M &= \text{diag}(\sigma_1, \ldots \sigma_n) \end{aligned} \qquad (5)$$

with M the observability gramian, W the controllability gramian, and σ_j the (Hankel-) singular values in descending order for $j = 1, \ldots, n$. In Matlab this is implemented by `[A,B,C,sing,T]=dbalreal(A,B,C)`, where `sing` records the singular values σ_j and `T` the required state space transformation.

In a second step, the singular values are split into two disjoint sets of *large* and *small* values, resp. $\sigma_1, \ldots, \sigma_k$ and $\sigma_{k+1}, \ldots, \sigma_n$, and the matrices and state vector are partitioned accordingly:

$$\begin{aligned} \mathbf{x}_1(t+1) &= A_{11}\mathbf{x}_1(t) + A_{12}\mathbf{x}_2(t) &+ B_1\mathbf{u}(t) \\ \mathbf{x}_2(t+1) &= A_{21}\mathbf{x}_1(t) + A_{22}\mathbf{x}_2(t) &+ B_2\mathbf{u}(t) \\ \mathbf{y}(t) &= C_1\mathbf{x}_1(t) + C_2\mathbf{x}_2(t) &+ D\mathbf{u}(t). \end{aligned} \qquad (6)$$

The classical procedure of truncation by balancing ([8]) simply amounts to removing the 'small parts', and result in the reduced order model $\Sigma_{\text{red}} := \Sigma(A_{11}, B_1, C_1, D)$. A similar approach for the second step is implemented in a model reduction procedure in Matlab (`dmodred`), and amounts to replacing the 'lower state equation'

$$\mathbf{x}_2(t+1) = A_{21}\mathbf{x}_1(t) + A_{22}\mathbf{x}_2(t) + B_2\mathbf{u}(t)$$

by

$$\mathbf{x}_2(t) = A_{21}\mathbf{x}_1(t) + A_{22}\mathbf{x}_2(t) + B_2\mathbf{u}(t),$$

followed by elimination of $\mathbf{x}_2$. This results in the reduced model $\Sigma_{\text{red}} := \Sigma(A_{11} - A_{12}(A_{22}-I)^{-1}A_{21}, B_1 - A_{12}(A_{22} - I)^{-1}B_2, C_1 - C_2(A_{22}-I)^{-1}A_{21}, D - C_2(A_{22}-I)^{-1}B_2)$.

These techniques are attractively straightforward, and they have proved their practical value in a huge amount of applications. Yet, from a theoretical point of view, it is quite unsatisfactory that the resulting approximate models are not optimal in a well-defined sense.

[1] A complication arises when the polynomial d has roots on the unit circle. Then $\mathbf{u}$ is no longer an arbitrary time series in ℓ_2. This complication is not further addressed in this paper.

4 Hankel-norm approximation

The state space algorithm for optimal Hankel-norm approximation as described in e.g. [4] starts with the partitioning (6) and then constructs, in a more complicated second step, an approximant $\hat{G}$ of McMillan degree k that is optimal in the sense that the Hankel-distance

$$\delta_{\mathrm{H}}(G,\widehat{G}) := \sup_{\mathbf{0}\neq\mathbf{u}^-\in\ell_2^-} \frac{\|[(G-\hat{G})(\mathbf{u}^-\wedge\mathbf{0})]^+\|}{\|\mathbf{u}^-\|} \tag{7}$$

is minimized in the class of all stable rational matrices $\widehat{G}$ of McMillan degree at most k. As shown in [4], this minimum is attained and equal to σ_{k+1}, as we also prove later on.

Some remarks are in order here. The criterion (7) measures to what extend past inputs effect future outputs of the error system $G-\widehat{G}$. From an intuitive point of view, this may be considered as a measure on the *memory structure* of the error system, which is fed by past inputs and influences future outputs. The criterion discards, therefore, the possible difference in the *past* outputs of G and $\widehat{G}$, which makes it less appealing as an approximation criterion from a purely behavioral point of view. In particular, anti-causal modes are neglected in the criterion (7). That is, $\delta_{\mathrm{H}}(G,\widehat{G}_1)=\delta_{\mathrm{H}}(G,\widehat{G}_2)$ if the difference $\widehat{G}_1-\widehat{G}_2$ is anti-stable. Although the given system Σ_{full} is stable by assumption, it turns out that unstable systems arise naturally in Hankel-norm reduction. Therefore we now first discuss a refinement of the notion of degree of a system.

Let $\Sigma=\Sigma(G)$ and denote by Σ^+ and Σ^- its future and past behavior, respectively. Define the *causal* and *anti-causal* degree of a system, resp. by

$$\begin{aligned}\delta^+(\Sigma) &:= \dim\{\mathbf{y}^+\in(\ell_2^p)^+ \mid (\mathbf{0},\mathbf{y}^+)\in\Sigma^+\}\\ \delta^-(\Sigma) &:= \dim\{\mathbf{y}^-\in(\ell_2^p)^- \mid (\mathbf{0},\mathbf{y}^-)\in\Sigma^-\}.\end{aligned}$$

The causal degree is the dimension of the free future response of the system (that remains in $(\ell_2^p)^+$). The anti-causal degree has a similar interpretation for the time-reversed system. Obviously, these numbers are equal to the number of roots of d that are resp. inside and outside the unit circle. Assuming that there are no roots on the unit circle, it follows that

$$\delta(\Sigma)=\delta^+(\Sigma)+\delta^-(\Sigma)$$

and that a system $\Sigma(G)$ admits a decomposition into a stable (or better 'causal') and anti-stable (or 'anti-causal') part

$$\Sigma(G)=\Sigma(G^a)+\Sigma(G^c) \tag{8}$$

with G^a and G^c given by

$$G^a(z)=\frac{N^a(z)}{d^a(z)},\quad G^c(z)=\frac{N^c(z)}{d^c(z)} \tag{9}$$

with $d^c(z)$ and $d^a(z)$ the unique stable resp. anti-stable polynomials such that $d^a(z)d^c(z)=d(z)$, and the polynomials N^a and N^c such that $G^a+G^c=G$. This decomposition is unique modulo a static feedthrough term.

If $G_{\text{red}}=G^a_{\text{red}}+G^c_{\text{red}}$ is such a decomposition of a reduced order model, then $G_{\text{red}}-G^c_{\text{red}}$ is anti-stable and since the Hankel distance is invariant for anti-stable effects it follows that

$$\delta_{\mathrm{H}}(G,G_{\text{red}})=\delta_{\mathrm{H}}(G,G^c_{\text{red}}) \tag{10}$$

for every G and G_{red}. Consequently, from any candidate approximant the anti-stable part may be removed without affecting the Hankel distance. This also shows that, in principle, Hankel-norm reduction may be applied to unstable systems in a trivial way, namely by first removing the anti-stable part of a given unstable system.

In [4], the expressions of the optimal reduced order models are derived directly from (A,B,C,D). However, these expressions are hard to interpret on the level of system trajectories. The aim of this paper is to show that Hankel-norm model reduction allows an elegant interpretation on the level of system trajectories.

5 Hankel-norm approximation by cutting trajectories

In this section we sketch how optimal Hankel-norm approximants can be obtained by cutting system trajectories. In the next section some details are worked out, and a proof of correctness is given.

A system trajectory is an input-output pair that is compatible with the system laws. Of particular interest to us are those input-output pairs related to the singular values of the system. For this purpose, consider the input-output pairs

$$\Big((\mathbf{u}^-_{\langle k\rangle}\wedge 0)\ ,\ (\mathbf{y}^-_{\langle k\rangle}\wedge\mathbf{y}^+_{\langle k\rangle})\Big)$$

with

$$\mathbf{u}^-_{\langle k\rangle}(-j)=B^\top(A^\top)^{j-1}W^{-\frac{1}{2}}e_k \text{ for } j>0 \tag{11a}$$

$$\mathbf{y}^-_{\langle k\rangle}(-j)=F^\top(A^\top)^{j-1}W^{-\frac{1}{2}}e_k \text{ for } j>0 \tag{11b}$$

$$\mathbf{y}^+_{\langle k\rangle}(j)=C^\top(A^\top)^{j}W^{\frac{1}{2}}e_k \quad \text{ for } j\geq 0 \tag{11c}$$

with $F:=AWC^\top+BD^\top$. The past input $\mathbf{u}^-_{\langle k\rangle}$ and normalized future output $(1/\sigma_k)\mathbf{y}^+_{\langle k\rangle}$ are also called the k-th Schmidt pair of Σ_{full}.

A more basic, representation-free definition is as follows. The first singular vector $\mathbf{u}_{\langle 1\rangle}$ is characterized by the property that it maximizes the ratio of norms of future outputs and past inputs in the system, i.e.,

$$\mathbf{u}_{\langle 1\rangle}=\underset{\|\mathbf{u}^-\|=1}{\operatorname{argmax}}\ \frac{\|\mathbf{y}^+\|}{\|\mathbf{u}^-\|}, \tag{12}$$

where $\mathbf{y}^+$ is an abbreviation for $[G(\mathbf{u}^-\wedge\mathbf{0})]^+$. The corresponding maximum ratio is denoted as σ_1, which corresponds to σ_1 in (5).

For $k = 2, \ldots, n$ the past input $\mathbf{u}^-_{\langle k\rangle}$ maximizes the same criterion subject to some orthogonality requirements:

$$\mathbf{u}^-_{\langle k\rangle} = \underset{\|\mathbf{u}^-\|=1, \mathbf{u}^- \perp \mathbf{u}_{\langle 1\rangle}, \ldots, \mathbf{u}_{\langle k-1\rangle}}{\operatorname{argmax}} \frac{\|\mathbf{y}^+\|}{\|\mathbf{u}^-\|} \qquad (13)$$

The ratios of the norms define the numbers σ_k and correspond to the Hankel singular values σ_k in (5).

Notice that the aforementioned lower bound σ_{k+1} of the Hankel criterion (7) for approximations Σ_{red} of degree k follows immediately from this formulation. Indeed, a smaller value for this bound would imply that the space of free responses of Σ_{red} (i.e., all future outputs with zero future input) is not orthogonal to any element in the span of $\{\mathbf{y}^-_{\langle j\rangle}\}_{j=1,\ldots,k+1}$, which would contradict that Σ_{red} is of degree k.

Now define the rational function G_k as the (unique) linear time-invariant single-input system that maps the past input $\mathbf{u}^-_{\langle k\rangle}$ onto the truncated output,

$$G_k : (\mathbf{u}^-_{\langle k\rangle} \wedge \mathbf{0}) \to (\mathbf{y}^-_{\langle k\rangle} \wedge \mathbf{0}). \qquad (14)$$

That is, let G_k be defined by the system with the property that

$$\Big((\mathbf{u}^-_{\langle k\rangle} \wedge \mathbf{0})\ ,\ (\mathbf{y}^-_{\langle k\rangle} \wedge \mathbf{0}) \Big) \ \in\ \Sigma(G_k).$$

Consequently, the error system $\widetilde{G}_k := G - G_k$ is then determined by the property

$$\widetilde{G}_k : (\mathbf{u}^-_{\langle k\rangle} \wedge \mathbf{0}) \to (\mathbf{0} \wedge \mathbf{y}^+_{\langle k\rangle}). \qquad (15)$$

We remark that, by construction, $G_k = G - \widetilde{G_k}$ has finite ℓ_2-induced norm and therefore G_k has no poles on the unit circle. We can therefore decompose G_k into a causal and anti-causal part,

$$G_k = G^c_k + G^a_k \qquad (16)$$

cf. (9). Then

- $\Sigma(G^c_k)$ is an optimal $(k-1)$st order Hankel-norm approximant of $\Sigma(G)$, i.e., it has degree $k-1$ and, within the class of all rational stable functions of degree at most $k-1$, it has minimal Hankel-distance $\delta_{\text{H}}(G, G^c_k) = \sigma_k$.
- In particular, $\Sigma(G_n)$ is the optimal $(n-1)$st order Hankel-norm approximant of G, and
- $\Sigma(G_1)$ is the Nehari extension of G.
- G_k is the $(k-1)$st order Hankel approximation minus the Nehari extension of the $k-1$st order error system $\widetilde{G}_k$.

This implies that, at least for single-input systems, optimal Hankel-norm reductions are obtained by the *realization* of a system from a time series which is obtained by truncation of a specific output of the original system. This result therefore connects the problem of optimal Hankel-norm model reduction with realization theory.

The main observation in this section is not new, and we mention a few specific places in the literature were a similar observation is made. In fact, in the original AAK paper [1, p. 34, first formula], the error systems are defined as the quotient of Schmidt pairs, and approximations as the causal part of the difference between the original system and the error system, for siso systems. Some alternative formulations can be found in e.g. [7, Section 7], [5, formula (5.1)], [3, Section 10], and [12, Lemma 8.22, for $k = 1$].

Our formulation differs from this in that it emphasizes a direct characterization of the approximation system (G_k) in terms of $\mathbf{y}^-_{\langle k\rangle}$ (which is not a left- or right singular vector of Σ_{full}), rather than an indirect approach via the error system $(\widetilde{G}_k)$, which maps left- to right singular vectors.

6 Proof

We will proof the following statements

1. $\|\widetilde{G}_k(\mathbf{u})\| = \sigma_k \|\mathbf{u}\|$
2. G_k as defined in (14) is of (McMillan) degree at most $n-1$.
3. G^c_k is of McMillan degree at most $k-1$.

From the first item it follows that $\delta_{\text{H}}(G, G_k) = \sigma_k$, hence also $\delta_{\text{H}}(G, G^c_k) = \sigma_k$, as $G_k - G^c_k$ is anti-stable, cf. (10). The optimality of $\Sigma(G^c_k)$ then follows from the last item. The second item is an auxiliary result that is also used for deriving some explicit formula later on.

Proof of 1.
First observe that the statement is true for input $\mathbf{u} = \mathbf{u}^-_{\langle k\rangle}$, as the ratio of norms of $\mathbf{y}^+_{\langle k\rangle}$ and $\mathbf{u}^-_{\langle k\rangle}$ is equal to σ_k, cf. (13). Further, these two sequences have exactly the same correlations. Namely, for $k = 1$, observe that if $\sigma_1^2\ \langle \sigma^j \mathbf{u}_{\langle 1\rangle}, \mathbf{u}_{\langle 1\rangle}\rangle \neq \langle \sigma^j \mathbf{y}^+_{\langle 1\rangle}, \mathbf{y}^+_{\langle 1\rangle}\rangle$ for some j, implies $\mathbf{u}_{\langle 1\rangle} + \epsilon \sigma^j \mathbf{u}_{\langle 1\rangle}$ would correspond to a ratio of norms beyond σ_1 for sufficiently small (not necessarily positive) ϵ, which contradicts (12). For $k > 1$ it follows from a straightforward inductive argument (alternatively, it may be derived from the explicit formula (11) for the singular vectors). Equality of correlations implies that $\frac{\|\widetilde{G}_k(\mathbf{u})\|}{\|\mathbf{u}\|} = \sigma_k$, for $\mathbf{u}$ any linear combination of $\mathbf{u}^-_{\langle k\rangle}$, and hence for all $\mathbf{u} \in \ell^1_2$.

Proof of 2.
We first derive that for all $j > 0$, $\mathbf{u}^-_{\langle k\rangle}$ is orthogonal to all shifts of $\mathbf{d}$ that have zero future,

$$\mathbf{u}^-_{\langle k\rangle} \perp \sigma^j(\mathbf{d}) \text{ for all } j > n \qquad (17)$$

with $\mathbf{d}$ as defined below equation (3). Indeed, the output corresponding to these shifts is given by $\sigma^j(\mathbf{N})$, and also has zero future. Now if $\mathbf{u}^-_{\langle k\rangle}$ is not orthogonal to $\sigma^k(\mathbf{d})$ for some $k > n$, then substracting its projection onto $\sigma^k(\mathbf{d})$ would decrease the norm of past inputs, without changing the free response $\mathbf{y}^+$. This contradicts (12). The same argument also applies for $k > 1$, with a slight adaptation in order to take the extra orthogonality conditions into account.

Now (17) implies that

$$(d_0\mathbf{u}^-_{\langle k\rangle} + d_1\sigma\mathbf{u}^-_{\langle k\rangle} + \ldots + d_n\sigma^n\mathbf{u}^-_{\langle k\rangle}) \in \ell_2^-$$

is zero for $t < -n$. Hence also the corresponding output in Σ_{full} must be zero for $t < -n$. Consequently,

$$(d_0\mathbf{y}^-_{\langle k\rangle} + d_1\sigma\mathbf{y}^-_{\langle k\rangle} + \ldots + d_n\sigma^n\mathbf{y}^-_{\langle k\rangle}) \in (\ell_2^p)^-$$

is zero for $t < -n$, which implies that each individual output component in $\mathbf{y}^-_{\langle k\rangle}$ is orthogonal to the left-shifts of $\mathbf{d}$ with zero future,

$$e_k^\top \mathbf{y}^-_{\langle k\rangle} \perp \sigma^j(\mathbf{d}) \text{ for all } j > n, k = 1, \ldots, p. \tag{18}$$

We remark that these orthogonality properties can also be deduced on the basis of (11), from the observation that $d(z)$ is the characteristic polynomial of A and an application of the Cayley-Hamilton theorem.

Now define

$$\mathbf{d}_{\langle k\rangle} := (d_0\mathbf{u}^-_{\langle k\rangle} + d_1\sigma\mathbf{u}^-_{\langle k\rangle} + \ldots + d_n\sigma^n\mathbf{u}^-_{\langle k\rangle}) \tag{19}$$

$$\mathbf{N}_{\langle k\rangle} := (N_0\mathbf{y}^-_{\langle k\rangle} + N_1\sigma\mathbf{y}^-_{\langle k\rangle} + \ldots + N_n\sigma^n\mathbf{y}^-_{\langle k\rangle}) \tag{20}$$

Then $G_k(\mathbf{d}_{\langle k\rangle} \wedge \mathbf{0}) = (\mathbf{N}_{\langle k\rangle} \wedge \mathbf{0})$ and, as in (3),

$$\big((\mathbf{d}_{\langle k\rangle} \wedge \mathbf{0})\ ,\ (\mathbf{N}_{\langle k\rangle} \wedge \mathbf{0})\big) \ \in\ \Sigma(G_k).$$

We have shown that $\mathbf{d}_{\langle k\rangle}$ and $\mathbf{N}_{\langle k\rangle}$ are trajectories with support $[-n, -1]$. Hence G_k is of McMillan degree $n-1$ (or smaller if there are common zeros and poles). In fact, the elements of $\mathbf{d}_{\langle k\rangle}$ and $\mathbf{N}_{\langle k\rangle}$ restricted to $[-n, 0]$ consist of the coefficients of the denominator and numerator of G_k, cf. (1) and (3).

Proof of 3.
First notice that for $k = n$ the statement has already been proved in *2*. At the opposite side, for $k = 1$ it states that G_1 is in fact the Nehari extension of G. For values $1 < k < n$, G_k contains causal as well as anti-causal modes, and a decomposition of the stable and unstable part of G_k on a purely geometric level falls outside the scope of this paper. Instead, we give a proof that is based on the explicit formulas in (11).

A backward state representation for the input-output pairs $\mathbf{w} = \begin{bmatrix}\mathbf{u}\\ \mathbf{y}\end{bmatrix} \in \ell_2^{p+1}$ for G_k is given by

$$\begin{aligned}\mathbf{x}' &= A^\top\sigma\mathbf{x}' + A^\top e_k\mathbf{h}\\ \mathbf{w}' &= \begin{bmatrix}B^\top\\ F^\top\end{bmatrix}\sigma\mathbf{x}' + \begin{bmatrix}B^\top\\ F^\top\end{bmatrix}e_k\mathbf{h},\end{aligned}$$

which means that all input/output pairs of G_k can be generated by auxiliary input $\mathbf{h}$ through this state space system.

Eliminating $\mathbf{h}$ again gives

$$\begin{aligned}\mathbf{x}' &= A^\top T_k^\top\sigma\mathbf{x}' + \frac{A^\top e_k}{B^\top e_k}\mathbf{u}\\ \mathbf{y}' &= F^\top T_k^\top\sigma\mathbf{x}' + \frac{F^\top e_k}{B^\top e_k}\mathbf{u}\end{aligned}$$

with $T_k := I_n - \frac{Be_k^\top}{e_k^\top B}$.

The poles of this system (forward in time) are the eigenvalues of $(T_kA)^{-1}$, and we have to show that T_kA has at most $k-1$ anti-stable eigenvalues. This is proved in a seperate lemma.

Lemma Let be given an asymptotically stable $A \in \mathbb{R}^{n\times n}$, and $B \in \mathbb{R}^{n\times 1}$, such that they have diagonal controllability gramian $W = AWA^\top + BB^\top = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$ with nonincreasing diagonal entries. Then $(I_n - \frac{Be_k^\top}{e_k^\top B})A$ has at most $k-1$ anti-stable poles.

Proof. The latter matrix equals T_kA with T_k defined as above. If $\sigma_1 \geq 1$, first redefine B as μB with $0 < \mu < 1/\sigma_1$. As T_k is invariant under scalar multiplication of B, it now suffices to prove the lemma for $\sigma_1 < 1$.

Determine $\widetilde{B} \in \mathbb{R}^{n\times n}$ such that $AA^\top + BB^\top + \widetilde{B}\widetilde{B}^\top = I_n$. Then the (non-singular) controllabilility gramian of $(A, \widetilde{B})$ equals $I_n - W =: \widetilde{W}$, as $A(I_n - W)A^\top + \widetilde{B}\widetilde{B}^\top = I_n - BB^\top - AWA^\top = I_n - W$. Now the controllability gramian X of $(T_kA, T_k\widetilde{B})$ is given by $\widetilde{W} - \sigma_k^{-1}W$, which is derived as follows.

By definition, X is the solution of $X = T_kAXA^\top T_k^\top + T_k\widetilde{B}\widetilde{B}^\top T_k^\top$. As $T_kB = 0$ and $e_k^\top T_k = 0$, it follows that $T_k(AXA^\top + \widetilde{B}\widetilde{B}^\top)T_k^\top = T_k(AXA^\top + \widetilde{B}\widetilde{B}^\top - \sigma_k^{-1}BB^\top)T_k^\top = T_kXT_k^\top = X + (I_n - T_k)X(I_n - T_k)^\top = X$. Observe that X is diagonal, with $k-1$ negative entries. From a well-known result on Lyaponov solutions (which in fact admits a straightforward proof) this implies that T_kA has indeed $k-1$ anti-stable poles (cf. e.g. [4, Theorem 3.3] □

7 Reduction Formula

In the proof we derived some explicit formula for the functions G_k. For $k = n-1$ this is equal to the $n-1$-th order approximant, and for $k = 1$ this is the Nehari extension of G, but for k in between they contain an anti-stable part that still has to be removed in order to obtain the final approximation. Here we combine the results for G_k into one formula, and do not address the question of dissolving their stable part.

From (20) and (19) it follows that

$$G_k(z) = \frac{N^{\langle k\rangle}(z)}{d^{\langle k\rangle}(z)} \tag{21}$$

with

$$\begin{aligned}N^{\langle k\rangle}(z) &= N_0^{\langle k\rangle} + N_1^{\langle k\rangle}z + \ldots N_{n-1}^{\langle k\rangle}z^{n-1},\\ d^{\langle k\rangle}(z) &= d_0^{\langle k\rangle} + d_1^{\langle k\rangle}z + \ldots d_{n-1}^{\langle k\rangle}z^{n-1},\end{aligned} \tag{22}$$

and coefficients given by

$$\begin{aligned}N_j^{\langle k\rangle} &= F^\top A_{n-j-1}^\top e_k \text{ and}\\ d_j^{\langle k\rangle} &= B^\top A_{n-j-1}^\top e_k,\end{aligned} \tag{23}$$

where

$$A_j := d_0 A^j + \ldots + d_{j-1} A + d_j I_n.$$

The sequence $\{A_0, \ldots, A_{n-1}\}$ is known as the Faddeev sequence of A, and can be computed recursively by setting $A_0 = d_0 I_n$, and $A_k := AA_{k-1} + d_k I_n$. We remark that Feddeev sequences can be used for determining matrix inverses, and are applied for solving polynomial Lyaponov equations ([6, 10]). We have not yet investigated these connections in detail.

We conclude by some remarks on the use of these formulas. They are easily adapted for state representations that are not balanced, e.g. controller and observer canonical forms. As a state space basis transformation does not affect the characteristic polynomial of the A-matrix, and affect all terms in (22) in the same way, these formulas hold true for e_k replaced by some *unknown* $z_k \in \mathbb{R}^n$. Balancing then amounts to finding z_k such that the corresponding systems in (21) have the desired optimality properties.

Further notice that $d^{\langle k \rangle}$ has the stable poles of the $k-1$-th Hankel-norm approximant, together with (at most) $n-k$ anti-stable poles of the Nehari extension of the error system. The formula for $d^{\langle k \rangle}$ may be grouped into a square matrix

$$[A_{n-1}B, A_{n-2}B, \ldots, A_0 B] \in \mathbb{R}^{n \times n}$$

The k-th row then contains the coefficients of the denominator of G_k, and this compact formula may be used to further explore the proces of model reduction.

8 Conclusions

We gave a straighforward construction and a self-contained derivation of optimal Hankel-norm approximants for discrete time single-input systems. It has been shown that optimal Hankel-norm approximants can be obtained by the *realization* of a system from time series which are obtained by truncation of specific outputs of the original system. This result therefore connects the problem of model approximation with realization theory of discrete time systems. Indeed, we have shown that by suitably disconnecting the past and future of 'balanced trajectories' of the original system, optimal approximants (in the Hankel sense) are obtained by removing the anti-causal part. A completion of the proof on the level of the trajectories is under construction. Some explicit expressions have been obtained for Hankel approximants (with the Nehari extension of the error system not yet removed) in terms of Faddeev sequences.

Finally, the basic idea of reduction by cutting 'balanced' trajectories is more generally applicable. In ([11]) it is applied in a behavioral framework, and another variant involving a criterion with relative output errors will be described in the near future.

Acknowledgements

For the first author, this research is supported by the Netherlands Organization for Scientific Research (NWO - MaGW/ESR, project number 510-01-0025). The research was partially initiated during a post-doc position at the Mathematics Institute, University of Groningen (sponsored by the Dutch Institute of Systems and Control (DISC), 1996/1997).

References

[1] V.J. Adamjan, D.Z. Arov, and M.G. Krein, *Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem*, Mat. USSR Sbornik, Vol. 15, pp. 31-73, 1971.

[2] V.J. Adamjan, D.Z. Arov, and M.G. Krein, *Infinite block Hankel matrices and related extension problems*, AMS transl., Vol. 111, pp. 133–156, 1978.

[3] P.A. Fuhrmann, *A polynomial approach to Hankel norm and balanced approximations*, Linear Algebra and its Applications, vol. 146, pp. 133–220, 1991.

[4] K. Glover, *All optimal Hankel norm approximations of linear multivariable systems and their L_∞ Bound*, Int. Journal of Control, Vol. 39, (1984) 1115–1193.

[5] S.K. Kung and D.W. Lin, *A state-space formulation for optimal Hankel-Nnrm approximations*, IEEE Transactions on Automatic Control, Vol. 26:4, pp. 942–946, 1981.

[6] B. Hanzon and R.L.M. Peeters, *A Faddeev sequence method for solving Lyaponov and Sylvester equations*, Linear Algebra and its Applications, vol. 241–243, pp. 431–453, 1996.

[7] G. Michaletzky, *A note on Hankel approximation of all-pass functions*, Linear Algebra and its Applications, vol. 255, pp. 19–55, 1997.

[8] B.C. Moore, *Principal component analysis in linear systems: controllability, observability and model reduction*, IEEE Transactions on Automatic Control, Vol. 26:1, pp. 17–32, 1981.

[9] J.R. Partington, *An Introduction to Hankel Operators*, London Mathematical Society Student Texts, Vol. 12, Cambridge University Press, 1988.

[10] R.L.M. Peeters and P. Rapisarda, *A new algorithm to solve the polynomial Lyaponov equation*, Report M 99-003, Maastricht University, Mathematics.

[11] B. Roorda and S. Weiland, *Optimal angle reduction - A behavioral approach to linear system approximation*, Submitted to Linear Algebra and its Applications, 2000.

[12] K. Zhou, J.C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice Hall NJ, 1996.