Model reduction for controller design for infinite-dimensional systems

Opmeer, Mark Robertus

Chapter 10

Balanced realizations

10.1 Lyapunov-balanced realizations

In this section we collect some results on Lyapunov-balanced realizations which are available in the literature.

Definition 10.1. A discrete-time system is called **Lyapunov-balanced** if it is input and output stable and its observability and controllability gramian are equal.

The following result shows the existence and uniqueness of Lyapunov-balanced realizations.

Proposition 10.2. *Any $H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$ function has a minimal Lyapunov-balanced realization. Minimal Lyapunov-balanced realizations are unique up to a unitary similarity transformation in the state space. Their state operator is a contraction. Both the minimal Lyapunov-balanced realization and its dual system are strongly stable.*

Proof. See Young [99] or Theorems 11.2.5 and 11.2.9 in Peller [75] for all the above statements except the ones about strong stability. The statements about strong stability can be found in Ober and Wu [63]. The complete theorem as stated above is contained in Theorem 9.5.6 of Staffans [89] (in the continuous-time version). □

Definition 10.3. A discrete-time system is called **compact Lyapunov-balanced** if it is Lyapunov-balanced and its gramian is compact.

We note that any $H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$ function has a bounded Hankel operator (see Definition A.24). It follows from Lemma A.26 that this Hankel operator is similar, through the Z-transform, to the Hankel map of any realization of the given function. We will call this the Hankel map of the given $H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$ function.

Proposition 10.4. *Any $H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$ function with a compact Hankel map has a minimal compact Lyapunov-balanced realization.*

Proof. By Proposition 10.2 the given function has a minimal Lyapunov-balanced realization. Since the Hankel map is independent of the realization this Lyapunov-balanced realization has a compact Hankel map. Denote the gramian by L. It follows, using Lemma 2.4 that $L^2 = \mathcal{C}^*\mathcal{C}\mathcal{B}\mathcal{B}^* = \mathcal{C}^*\mathcal{H}\mathcal{B}^*$ is compact. From this we conclude that L is compact. □

Remark 10.5. From the proof of Propositions 10.2 and 10.4 one can obtain the following explicit form of a compact Lyapunov-balanced realization.

Assume that we are given a $H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$ function with a compact Hankel map $\mathcal{H}$. Recall the backward shift realization Σ^{rs} from Remark 2.13. Let $\mathscr{X}$ be the closure of the range of the Hankel map. Since $\mathcal{H}$ is compact there exist a nonincreasing positive sequence (σ_i) and orthonormal bases (v_i) of the closure of the range of $\mathcal{H}^*$ and (w_i) of $\mathscr{X}$ such that

$$\mathcal{H}v_i = \sigma_i w_i, \quad \mathcal{H}^* w_i = \sigma_i v_i$$

(this is known as the structure theorem for compact operators, note that the σ_i are positive since we only want the (w_i) to be a basis for $\mathscr{X}$, not for the whole of $l^2(\mathbb{Z}^+, \mathscr{Y})$). The σ_i are called the **Hankel singular values** of the system and the (v_i, w_i) the **Schmidt pairs**. We have

$$\langle A^{\text{bal}} w_j, w_i \rangle_{l^2(\mathbb{Z}^+, \mathscr{Y})} = \sqrt{\frac{\sigma_j}{\sigma_i}} \langle A^{\text{bs}} w_j, w_i \rangle_{l^2(\mathbb{Z}^+, \mathscr{Y})}.$$

Pick an orthonormal basis (u_i) in $\mathscr{U}$, then

$$\langle B^{\text{bal}} u_j, w_i \rangle_{l^2(\mathbb{Z}^+, \mathscr{Y})} = \frac{1}{\sqrt{\sigma_i}} \langle B^{\text{bs}} u_j, w_i \rangle_{l^2(\mathbb{Z}^+, \mathscr{Y})}.$$

Note that since $B^{\text{bs}} u = \mathcal{H}\underline{u}$, where $\underline{u} : \mathbb{Z}^- \to \mathscr{U}$ is defined by $\underline{u}_{-1} = u$ and $\underline{u}_{-i} = 0$ if $i > 1$, we have

$$\langle B^{\text{bal}} u_j, w_i \rangle_{l^2(\mathbb{Z}^+, \mathscr{Y})} = \sqrt{\sigma_i} \langle \underline{u_j}, v_i \rangle_{l^2(\mathbb{Z}^-, \mathscr{U})}.$$

Choose an orthonormal basis (y_i) in $\mathscr{Y}$, then

$$\langle C^{\text{bal}} w_j, y_i \rangle_{\mathscr{Y}} = \sqrt{\sigma_j} \langle C^{\text{bs}} w_j, y_i \rangle_{\mathscr{Y}}.$$

Of course D^{bal} equals the value of the given $H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$ function at zero, as is the case with any realization. The gramians are both equal to $L^{\text{bal}} := \sqrt{\mathcal{H}\mathcal{H}^*}$. So with respect to the orthonormal basis (w_i) the gramian is diagonal, we have

$$\langle L^{\text{bal}} w_j, w_i \rangle = \sigma_i \delta_{ij},$$

with δ_{ij} the Kronecker delta.

Definition 10.6. Let $\mathsf{G} \in H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$ have a compact Hankel map $\mathcal{H}$. The realization from Remark 10.5 is called the compact Lyapunov-balanced realization of G with respect to the sequence of eigenvectors (w_i) and is denoted by $\Sigma^{\text{bal}}_{(w_i)}$.

Note that $\Sigma^{\text{bal}}_{(w_i)}$ is always approximately controllable and observable since the gramian is positive.

Definition 10.7. Let $\mathsf{G} \in H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$ have a compact Hankel map $\mathcal{H}$. Let (w_i) be an ordered sequence of eigenvectors of $\mathcal{H}\mathcal{H}^*$ (the ordering is such that the corresponding eigenvalues σ_i^2 form a nonincreasing sequence). Let $n \in \mathbb{Z}^+$ be such that $\sigma_n > \sigma_{n+1}$. The **truncated Lyapunov-balanced realization** of dimension n of G with respect to the sequence of eigenvectors (w_i) is defined as the restriction/projection of $\Sigma^{\text{bal}}_{(w_i)}$ onto $\mathscr{X}_n := \{w_i : i = 1, \ldots, n\}$.

Remark 10.8. Note that we used the term 'an ordered sequence of eigenvectors of $\mathcal{H}\mathcal{H}^*$' since such a sequence is not unique if $\mathcal{H}\mathcal{H}^*$ has repeated eigenvalues. As we will show in the next lemma the condition $\sigma_n > \sigma_{n+1}$ ensures that the choice of ordered sequence of eigenvectors is to a large extent unimportant. Also note that in the case that $\mathcal{H}\mathcal{H}^*$ has repeated eigenvalues the truncated Lyapunov-balanced realization of dimension n is not defined for every $n \in \mathbb{Z}^+$.

Proposition 10.9. *Let $\mathsf{G} \in H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$ have a compact Hankel map $\mathcal{H}$. Let (w_i) and $(\tilde{w}_i)$ be ordered bases of eigenvectors of $\mathcal{H}\mathcal{H}^*$. The truncated Lyapunov-balanced realization of dimension n of G with respect to the sequence of eigenvectors (w_i) and that with respect to the the sequence of eigenvectors $(\tilde{w}_i)$ are related by a unitary similarity transformation. In particular, the transfer functions are the same.*

Proof. Since $\sigma_n > \sigma_{n+1}$ we have that $\mathscr{X}_n$ is the direct sum of eigenspaces. It follows that both $(w_i)_{i=1,\ldots,n}$ and $(\tilde{w}_i)_{i=1,\ldots,n}$ are bases for $\mathscr{X}_n$. Define the unitary operator $U \in \mathcal{L}(\mathscr{X}_n)$ by $Uw_i = \tilde{w}_i$ with $i = 1, \ldots, n$. It is easily seen that this is the desired unitary similarity transformation. $\square$

We now consider the distance between a discrete-time system and its truncated Lyapunov-balanced realizations. We measure this distance by the supremum norm of the difference of the transfer functions. To formulate the conditions needed we need the concept of a nuclear operator.

Remark 10.10. The **singular values** of an operator $T \in \mathcal{L}(\mathscr{H}_1, \mathscr{H}_2)$ are defined as follows. The k-th singular value of T is the distance, with respect to the norm in $\mathcal{L}(\mathscr{H}_1, \mathscr{H}_2)$, of T from the set of operators in $\mathcal{L}(\mathscr{H}_1, \mathscr{H}_2)$ of rank at most $k-1$. If T is compact, then the singular values are exactly the square roots of the eigenvalues of TT^*.

Remember that an operator $T \in \mathcal{L}(\mathscr{H}_1, \mathscr{H}_2)$ is **nuclear** if its singular values s_i satisfy $\sum_{i=1}^{\infty} s_i < \infty$. The sum of the singular values is called the nuclear norm. The set of nuclear operators is a linear subspace of $\mathcal{L}(\mathscr{H}_1, \mathscr{H}_2)$ and the nuclear norm is a norm on this subspace. Nuclear operators are compact. The operator T is called **Hilbert-Schmidt** if $\sum_{i=1}^{\infty} s_i^2 < \infty$. Equivalently, $T \in \mathcal{L}(\mathscr{H}_1, \mathscr{H}_2)$ is Hilbert-Schmidt if for all orthonormal bases (e_i) of $\mathscr{H}_1$ we have $\sum_i \|Te_i\|^2 < \infty$. Hilbert-Schmidt operators are compact.

Proposition 10.11. *Assume that $\mathsf{G} \in H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$, with $\mathscr{U}$ and $\mathscr{Y}$ finite-dimensional, has a nuclear Hankel map $\mathcal{H}$. Define G^n as the transfer function of a truncated Lyapunov-balanced realization of dimension n of G. Then we have*

$$\|\mathsf{G} - \mathsf{G}^n\|_\infty \leq 2 \sum_{i=n+1}^{\infty} \sigma_i.$$

In particular $\mathsf{G}^n \to \mathsf{G}$ in the H^∞ norm as $n \to \infty$.

Proof. The proof of this proposition is on page 110. □

Remark 10.12. Note that the error-bound does not depend on the choice made in the eigenvectors w_i used to define the truncated balanced realization of dimension n. This is due to the fact that by Proposition 10.9 the transfer function G^n does not depend on the choice of eigenvectors w_i.

Remark 10.13. Note that it follows from Proposition 10.11 that, under the conditions stated in that proposition, the Hankel operator of G^n converges to the Hankel operator of G in the nuclear norm.

The following proposition identifies a fundamental limitation to approximating a system by a system with a finite-dimensional state space.

Proposition 10.14. *Assume that $\mathsf{G} \in H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$ with $\mathscr{U}$ finite-dimensional. Let G^n be the transfer function of a discrete-time system with state space dimension n, input space $\mathscr{U}$ and output space $\mathscr{Y}$. Then $\mathsf{G}^n \to \mathsf{G}$ in the H^∞ norm as $n \to \infty$ only if G has a compact Hankel map.*

Proof. By assumption $\mathsf{G}^n \to \mathsf{G}$ in the H^∞ norm as $n \to \infty$. It follows that the Hankel operator of G^n converges to the Hankel operator of G in the operator norm. Since G^n is the transfer function of a discrete-time system with state space dimension n and input space dimension m its Hankel operator has rank at most mn. Since the Hankel operator of G is in the closure in the operator norm of the space of finite-rank operators it must be compact. □

Most of the remainder of this section on Lyapunov-balanced realizations is devoted to a proof of Proposition 10.11. Lemmas 10.15 and 10.16 however are included for the proof of Proposition 10.17, which gives a sufficient condition for nuclearity of the Hankel map.

Lemma 10.15. *Let Σ be a discrete-time system that is exponentially stable and has a finite-dimensional output space. Then its output map is Hilbert-Schmidt.*

Proof. Since the system is exponentially stable, there exist $M \geq 0$ and $r \in [0,1)$ such that for all $x \in \mathscr{X}$ we have $\|A^n x\| \leq M r^n \|x\|$ by Proposition 3.26.

Define $p := \dim \mathscr{Y}$ and let $(y_i)_{i\in\{1,\dots p\}}$ be a basis for the output space. Define $\mathcal{C}_i : \mathscr{X} \to l^2(\mathbb{Z}^+)$ by $(\mathcal{C}_i x)_n = \langle (\mathcal{C}x)_n, y_i\rangle_{\mathscr{Y}}$, where $\mathcal{C}$ is the output map of the system. We first prove that $\mathcal{C}_i$ is Hilbert-Schmidt.

For $n \in \mathbb{Z}^+$ define $\mathcal{C}_i^n : \mathscr{X} \to \mathbb{C}$ by $\mathcal{C}_i^n x = (\mathcal{C}_i x)_n$. This mapping is a continuous linear functional. By the Riesz representation theorem there exists, for each $n \in \mathbb{Z}^+$, a $w_n \in \mathscr{X}$ such that $\mathcal{C}_i^n x = \langle x, w_n\rangle_{\mathscr{X}}$ and

$$\|w_n\| = \|\mathcal{C}_i^n\| = \sup_{\|x\|=1} |\mathcal{C}_i^n x|.$$

We have

$$|\mathcal{C}_i^n x| = |\langle (\mathcal{C}x)_n, y_i\rangle_{\mathscr{Y}}| = |\langle CA^n x, y_i\rangle_{\mathscr{Y}}| \leq \|C\|\ M\ r^n\ \|x\|,$$

where in the last step we have used exponential stability. Thus we have

$$\begin{aligned}\sum_{n=0}^{\infty} \|\mathcal{C}_i^n\|^2 &= \sum_{n=0}^{\infty} \sup_{\|x\|=1} |\mathcal{C}_i^n x|^2 \leq \sum_{n=0}^{\infty} \sup_{\|x\|=1} \|C\|^2\ M^2\ r^{2n}\ \|x\|^2 \\ &= \|C\|^2\ M^2 \sum_{n=0}^{\infty} r^{2n} = \frac{\|C\|^2\ M^2}{1-r^2} < \infty.\end{aligned}$$

Let $(x_j)_{j\in\mathbb{Z}^+}$ be an orthonormal basis for $\mathscr{X}$. We compute

$$\sum_{j=0}^{\infty} \|\mathcal{C}_i x_j\|^2 = \sum_{j=0}^{\infty}\sum_{n=0}^{\infty} |\mathcal{C}_i^n x_j|^2 = \sum_{j=0}^{\infty}\sum_{n=0}^{\infty} |\langle x_j, w_n\rangle|^2 = \sum_{n=0}^{\infty}\sum_{j=0}^{\infty} |\langle x_j, w_n\rangle|^2.$$

From the Parseval relation we obtain that this is equal to

$$\sum_{n=0}^{\infty} \|w_n\|^2 = \sum_{n=0}^{\infty} \|\mathcal{C}_i^n\|^2.$$

We already saw that the right-hand side of this equation is finite. It follows that

$$\sum_{j=0}^{\infty} \|\mathcal{C}_i x_j\|^2 < \infty,$$

for all orthonormal sequences $(x_j)_{j\in\mathbb{Z}^+}$. This shows that $\mathcal{C}_i$ is Hilbert-Schmidt. We use the fact that $\mathcal{C}_i$ is Hilbert-Schmidt to show that $\mathcal{C}$ is. We have

$$\sum_{j=0}^{\infty} \|\mathcal{C} x_j\|^2 = \sum_{j=0}^{\infty}\sum_{i=1}^{p} \|\mathcal{C}_i x_j\|^2 = \sum_{i=1}^{p}\sum_{j=0}^{\infty} \|\mathcal{C}_i x_j\|^2 < \infty.$$

This shows that $\mathcal{C}$ is Hilbert-Schmidt. □

Lemma 10.16. *Let Σ be a discrete-time system that is exponentially stable and has a finite-dimensional input space. Then its input map is Hilbert-Schmidt.*

Proof. This follows from Lemma 10.15 applied to the dual system of Σ using that the adjoint of a Hilbert-Schmidt operator is Hilbert-Schmidt. □

Proposition 10.17. *Let Σ be a discrete-time system that is exponentially stable and has finite-dimensional input and output spaces. Then its Hankel map is nuclear.*

Proof. It follows from Lemmas 10.15 and 10.16 that the input and output maps are Hilbert-Schmidt. Using that the Hankel map is the product of these two maps (Lemma 2.4) and that the product of two Hilbert-Schmidt operators is nuclear we obtain the desired result. □

All results from the next proposition to the end of this section are included as building blocks for the proof of Proposition 10.11. We first prove a property of truncated Lyapunov-balanced realizations.

Proposition 10.18. *Let $\mathsf{G} \in H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$ have a compact Hankel map with at least n nonzero singular values. Then a n-dimensional truncated Lyapunov-balanced realization is exponentially stable.*

Proof. We decompose the state space $\mathscr{X} = \mathscr{X}_1 \oplus \mathscr{X}_2$, where $\mathscr{X}_1$ corresponds to the first n terms in the sequence (w_i). The system operator and the gramian are decomposed accordingly. The observation Lyapunov equation of $\Sigma^{\mathrm{bal}}_{(w_i)}$ gives

$$A_{11}^* L_1 A_{11} + A_{21}^* L_2 A_{21} - L_1 + C_1^* C_1 = 0.$$

Assume that $A_{11} v = \lambda v$. Using the above identity we obtain

$$(1 - |\lambda|^2)\ \|L_1^{1/2} v\|^2 = \|L_2^{1/2} A_{21} v\|^2 + \|C_1 v\|^2.$$

Since this is nonnegative, we obtain $|\lambda| \leq 1$. If $|\lambda| = 1$ then $C_1 v = 0$ and $L_2^{1/2} A_{21} v = 0$. Since $L_2 > 0$ (since the gramian of $\Sigma^{\mathrm{bal}}_{(w_i)}$ is positive), it follows that $A_{21} v = 0$. Define $V = [v; 0]$. Then (using $A_{21} v = 0$) $AV = \lambda V$, so $CA^k V = \lambda^k C_1 v = 0$. From the approximate observability of $\Sigma^{\mathrm{bal}}_{(w_i)}$ we obtain $V = 0$. We conclude that all eigenvalues of A_{11} are in the open unit disc. Since the state space is finite-dimensional, it follows that the n-dimensional truncated Lyapunov-balanced realization is exponentially stable. ☐

The following lemma shows continuous dependence of the eigenvalues and eigenvectors. Note that for continuity of the eigenvectors we might have to resort to a subsequence.

Lemma 10.19. *Let $T_m, T \in \mathcal{L}(\mathscr{H})$ be compact and nonnegative self-adjoint and such that $\|T_m - T\| \to 0$ as $m \to \infty$. Denote the eigenvalues of T by λ_i. Assume that the eigenvalues are ordered in decreasing magnitude and repeated according to their multiplicity. Then for the eigenvalues λ_i^m of T_m, also ordered in decreasing magnitude and repeated according to their multiplicity, we have $\lambda_i^m \to \lambda_i$. Let v_i^m be a basis of eigenvectors for T_m. There exists a subsequence T_{m_k} of T_m and a basis of eigenvectors (v_i) for T such that $\|v_i^{m_k} - v_i\| \to 0$ as $k \to \infty$.*

Proof. We recall Weyl's theorem on eigenvalues: if $A, B \in \mathcal{L}(\mathscr{H})$ are compact nonnegative self-adjoint operators, then their eigenvalues (ordered in decreasing magnitude) satisfy $\lambda_{i+j-1}(A+B) \leq \lambda_i(A) + \lambda_j(B)$. Taking $j = 1$ we obtain $\lambda_i(A + B) \leq \lambda_i(A) + \|B\|$.

From Weyl's theorem on eigenvalues we obtain for the given situation $\lambda_i^m \leq \lambda_i + \|T - T_m\|$ and $\lambda_i \leq \lambda_i^m + \|T - T_m\|$. This shows that $\lambda_i^m \to \lambda_i$ as $m \to \infty$.

We first show the statement on the eigenvectors for the set of leading eigenvectors (i.e. the ones corresponding to the largest eigenvalue). Denote the multiplicity of the largest eigenvalue of T by N. Further denote $v^m := [v_1^m, \dots v_N^m]$ and $v := [v_1, \dots v_N]$, where the v_i are orthonormal eigenvectors of T corresponding to λ_1 and the v_i^m are orthonormal eigenvectors of T_m for the largest N eigenvalues (counted according to their multiplicity) λ_i^m. Decompose $v^m = R^m v + x^m$ with R^m a $N \times N$ matrix of complex numbers and $\langle x_i^m, v_j \rangle = 0$ for all $i = 1, \ldots, N$ and $j = 1, \ldots, N$. Denote the rows of R by R_i. Then we have $v_i^m = R_i^m v + x_i^m$. Note that by the Pythagorean Theorem $\|x_i^m\|^2 = 1 - \|R_i^m\|_{\mathbb{C}^N}^2$. We have

$$\lambda_i^m = \|T_m v_i^m\| \leq \|T v_i^m\| + \|T - T_m\| = \|\lambda_1 R_i^m v + T x_i^m\| + \|T - T_m\|.$$

Note that the v_i are orthonormal and $\langle T x_i^m, v_j \rangle = \langle x_i^m, T v_j \rangle = 0$ since $\langle x_i^m, v_j \rangle = 0$. By the Pythagorean Theorem we then have

$$\|\lambda_1 R_i^m v + T x_i^m\|^2 = \lambda_1^2 \|R_i^m\|_{\mathbb{C}^N}^2 + \|T x_i^m\|^2.$$

Denote the restriction of T to the orthogonal complement of the eigenspace corresponding to λ_1 by $\check{T}$. Then $\check{T}$ is a compact nonnegative self-adjoint operator with eigenvalues $(\lambda_i)_{i \geq N+1}$ and so its norm is λ_{N+1}. Since x_i^m is in the domain of $\check{T}$ we have $\|T x_i^m\| \leq \lambda_{N+1} \|x_i^m\|$. Combining the above we obtain

$$\lambda_i^m \leq \sqrt{\lambda_1^2 \|R_i^m\|_{\mathbb{C}^N}^2 + \lambda_{N+1}^2 \left(1 - \|R_i^m\|_{\mathbb{C}^N}^2\right)} + \|T - T_m\|.$$

This gives

$$\frac{\left(\lambda_i^m - \|T - T_m\|\right)^2 - \lambda_{N+1}^2}{\lambda_1^2 - \lambda_{N+1}^2} \leq \|R_i^m\|_{\mathbb{C}^N}^2.$$

Since we have convergence of the eigenvalues, the left-hand side converges to 1. Since we have $\|R_i^m\|_{\mathbb{C}^N}^2 \leq 1$, we must have $\|R_i^m\|_{\mathbb{C}^N}^2 \to 1$. This implies that $\|x_i^m\| \to 0$. Hence $\|x^m\| \to 0$. The sequence of matrices (R^m) is bounded, which implies that it has a convergent subsequence. Denote the limit of such a subsequence by R^∞. Since $\|x^m\| \to 0$, we then have that the corresponding subsequence of (v^m) converges to $v^\infty := R^\infty v$. Since the components of (v^m) have norm one and are orthogonal to each other, the same holds for the components of v^∞. We now replace the set of orthonormal eigenvectors $(v_i)_{i=1,\ldots,N}$ by the components of v^∞. This gives another sequence of orthonormal eigenvectors of T. For the leading eigenvalue this sequence has the desired properties.

The general result follows by induction. Assume that we have proven the assertion for the first n eigenvalues (not counting multiplicity) with respective multiplicies N_n. Denote $N := \sum_{k=1}^n N_n$. We can apply the above to the operators $\tilde{T}$ and $\tilde{T}_m$, defined as the restriction of T respectively T_m, to the orthogonal complement of the eigenspaces corresponding to the first N eigenvalues (counting multiplicity). This gives the result for the first $n+1$ eigenvalues (not counting multiplicity). $\square$

Note that in Lemma 10.19 the eigenvectors v_i of T depend on the approximating sequence T_m: a different choice of approximating sequence may lead to a different orthonormal set of eigenvectors for T. The following lemma shows that we can obtain any desired orthonormal basis of eigenvectors for T by properly adjusting the approximating sequence. Lemma 10.20 is not needed for the proof of Proposition 10.11, but we give it for completeness.

Lemma 10.20. *Let $T_m, T \in \mathcal{L}(\mathscr{H})$ be compact and nonnegative self-adjoint and such that $\|T_m - T\| \to 0$ as $m \to \infty$. Let (v_i) be an orthonormal basis of eigenvectors of T, ordered according to decreasing magnitude of the corresponding eigenvalues. Then there exists a unitary operator $U \in \mathcal{L}(\mathscr{H})$ such that $U^*TU = T$ and there exists an orthonormal basis of eigenvectors of a subsequence of $\tilde{T}_m := U^*T_mU$ that converges to the given eigenvectors (v_i) as $m \to \infty$.*

Proof. From Lemma 10.19 we obtain an orthonormal basis of eigenvectors (w_i) of T and eigenvectors v_i^m of a subsequence of T_m such that $v_i^m \to w_i$. We define the unitary operator U by $Uv_i = w_i$. Note that since (w_i) and (v_i) are orthonormal bases this operator is indeed well-defined and unitary. Since both (w_i) and (v_i) are ordered the eigenvalues corresponding to the same index are equal. This gives $U^*TUv_i = Tv_i$, from which we obtain $U^*TU = T$ since (v_i) is a basis. Define $\tilde{v}_i^m := U^*v_i^m$. Then $\tilde{v}_i^m$ is an eigenvector of $\tilde{T}_m$ with eigenvalue λ_i^m. Since U is unitary the $\tilde{v}_i^m$ are orthonormal. We have $\tilde{v}_i^m \to U^*w_i = v_i$, since by assumption $v_i^m \to w_i$ and by definition of U we have $v_i = U^*w_i$. □

Using Lemma 10.19 we show the continuity of singular values and Schmidt pairs.

Lemma 10.21. *Let $T_m, T \in \mathcal{L}(\mathscr{H})$ be compact and such that $\|T_m - T\| \to 0$ as $m \to \infty$. Denote the singular values of T by σ_i. Assume that the singular values are ordered in decreasing magnitude and repeated according to their multiplicity. Then for the singular values σ_i^m of T_m, also ordered in decreasing magnitude and repeated according to their multiplicity, we have $\sigma_i^m \to \sigma_i$ as $m \to \infty$. Let (v_i^m, w_i^m) be Schmidt pairs for T_m. There exists a subsequence T_{m_k} of T_m and Schmidt pairs (v_i, w_i) for T such that $\|v_i^{m_k} - v_i\| \to 0$ and $\|w_i^{m_k} - w_i\| \to 0$ as $k \to \infty$. The w_i form an orthonormal basis of eigenvectors for TT^* and the v_i for T^*T.*

Proof. We first show that $T_m \to T$ implies $T_m^*T_m \to T^*T$. We have

$$\begin{aligned} \|T^*T - T_m^*T_m\| &= \|T^*T - T^*T_m + T^*T_m - T_m^*T_m\| \\ &\leq \|T\|\,\|T - T_m\| + \|T_m\|\,\|T - T_m\|. \end{aligned}$$

Since $T_m \to T$ we have $\|T_m\| \leq 2\|T\|$ for m large enough, which together with the above inequality gives the desired convergence. So we can apply Lemma 10.19 to obtain the convergence of the singular values. This lemma also gives a basis (v_i) of eigenvectors for T^*T and a basis (w_i) of eigenvectors for TT^* with the desired convergence properties. We only need to show that (v_i, w_i) is a Schmidt pair for T, i.e. $Tv_i = \sigma_i w_i$ and $T^*w_i = \sigma_i v_i$. We show this using that (v_i^m, w_i^m) is a Schmidt pair for T_m, i.e. $T_m v_i^m = \sigma_i^m w_i^m$ and $T_m^* w_i^m = \sigma_i^m v_i^m$. We have

$$\|Tv_i - \sigma_i w_i\| \leq \|T - T_m\| + \|T_m\| \, \|v_i - v_i^m\| + |\sigma_i^m - \sigma_i| + |\sigma_i| \|w_i^m - w_i\|,$$

which implies that $Tv_i = \sigma_i w_i$. The other equality is proven similarly. □

Similarly to Lemma 10.20 we can obtain any desired sequence of Schmidt pairs by changing the approximating sequence. Lemma 10.22 is not needed for the proof of Proposition 10.11, it is given for sake of completeness.

Lemma 10.22. *Let $T_m, T \in \mathcal{L}(\mathscr{H}_1, \mathscr{H}_2)$ be compact and such that $\|T_m - T\| \to 0$ as $m \to \infty$. Denote the singular values of T by σ_i and corresponding Schmidt pairs by (v_i, w_i). Then for the singular values σ_i^m of T_m we have $\sigma_i^m \to \sigma_i$. Furthermore, there exist unitary operators $V \in \mathcal{L}(\mathscr{H}_1)$ and $W \in \mathcal{L}(\mathscr{H}_2)$ such that $TV = WT$ and there exist Schmidt pairs $(v_i^{m_k}, w_i^{m_k})$ for a subsequence of $\tilde{T}_m := WT_mV^*$ such that $\|v_i^{m_k} - v_i\| \to 0$ and $\|w_i^{m_k} - w_i\| \to 0$ as $k \to \infty$.*

Proof. Convergence of the singular values follows immediately from Lemma 10.19. Let $(\tilde{v}_i^m, \tilde{w}_i^m)$ be a given sequence of Schmidt pairs of T_m. By Lemma 10.19 applied to T^*T with approximating sequence $T_m^*T_m$ there exist a basis of eigenvectors $(\tilde{v}_i)$ of T^*T such that $\tilde{v}_i^m \to \tilde{v}_i$. Similarly, there exist a basis of eigenvectors $(\tilde{w}_i)$ of TT^* such that $\tilde{w}_i^m \to \tilde{w}_i$. Since $(\tilde{v}_i^m, \tilde{w}_i^m)$ is a Schmidt pair we have $T_m\tilde{v}_i^m = \sigma_i^m \tilde{w}_i^m$ and $T_m^*\tilde{w}_i^m = \sigma_i^m \tilde{v}_i^m$. Taking limits we obtain $T\tilde{v}_i = \sigma_i \tilde{w}_i$ and $T^*\tilde{w}_i = \sigma_i \tilde{v}_i$, which shows that $(\tilde{v}_i, \tilde{w}_i)$ is a Schmidt pair of T. Define $V \in \mathcal{L}(\mathscr{H}_1)$ and $W \in \mathcal{L}(\mathscr{H}_2)$ by $V\tilde{v}_i = v_i$ and $W\tilde{w}_i = w_i$, respectively. Since $(\tilde{v}_i)$, (v_i), $(\tilde{w}_i)$, (w_i) are orthonormal bases V and W are unitary. Define $\tilde{T}_m := WT_mV^*$ and $v_i^m := V\tilde{v}_i^m$, $w_i^m := W\tilde{w}_i^m$. Then (v_i^m, w_i^m) is a Schmidt pair of $\tilde{T}_m$ since

$$\tilde{T}_m v_i^m = WT_mV^*v_i^m = WT_m\tilde{v}_i^m = \sigma_i^m W\tilde{w}_i^m = \sigma_i^m w_i^m,$$

$$\tilde{T}_m^* w_i^m = VT_m^*W^*w_i^m = VT_m^*\tilde{w}_i^m = \sigma_i^m V\tilde{v}_i^m = \sigma_i^m v_i^m.$$

We have $v_i^m = V\tilde{v}_i^m \to V\tilde{v}_i = v_i$ and $w_i^m = W\tilde{w}_i^m \to W\tilde{w}_i = w_i$. We have $TV = WT$ since $W^*TV\tilde{v}_i = W^*Tv_i = \sigma_i W^*w_i = \sigma_i\tilde{w}_i = T\tilde{v}_i$. □

The following lemma gives a series expansion of a function that has a nuclear Hankel map.

Lemma 10.23. *Assume that* $\mathsf{G} \in H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$*, with* $\mathscr{U}$ *and* $\mathscr{Y}$ *finite-dimensional, has a nuclear Hankel map. Then there exist* $c_n \in \mathcal{L}(\mathscr{U}, \mathscr{Y})$*,* $\lambda_n \in \mathbb{D}$ *such that*

$$\mathsf{G}(z) = \sum_{n=1}^{\infty} c_n \frac{1}{1-\lambda_n z}.$$

If $\mathscr{U} = \mathscr{Y} = \mathbb{C}$ *then we have*

$$\sum_{n=1}^{\infty} \frac{|c_n|}{1-|\lambda_n|} < \infty. \tag{10.1}$$

Proof. The scalar statement can be found in Peller [75, page 238]. The matrix statement follows from applying the scalar statement to components. □

Definition 10.24. Assume that $\mathsf{G} \in H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$, with $\mathscr{U}$ and $\mathscr{Y}$ finite-dimensional, has a nuclear Hankel map. Let c_n and λ_n be as in Lemma 10.23. For $m \in \mathbb{Z}^+$ define

$$\mathsf{G}^m(z) := \sum_{n=1}^{m} c_n \frac{1}{1-\lambda_n z}.$$

This is called the (m, c, λ) nuclear approximant of G.

Lemma 10.25. *With the assumptions and the notation as in Definition 10.24 we have*

$$\|\mathsf{G} - \mathsf{G}_m\|_\infty \to 0, \quad \|\mathcal{H} - \mathcal{H}_m\|_N \to 0,$$

where $\mathcal{H}$ *is the Hankel map of* G*,* $\mathcal{H}_m$ *is the Hankel map of* G_m *and* $\|\cdot\|_N$ *is the nuclear norm.*

Proof. We first compute the minimum of the absolute value of $z \mapsto 1 - \lambda_n z$ on the unit disc. Using the triangle inequality we have $|1 - \lambda_n z| \geq 1 - |\lambda_n|$. For $z = \bar{\lambda}_n / |\lambda_n|$ we have equality, so the minimum is $1 - |\lambda_n|$. It follows that

$$\left\| \frac{1}{1-\lambda_n \cdot} \right\|_\infty = \frac{1}{1-|\lambda_n|}.$$

We show convergence in the H^∞ norm for the scalar case. We have

$$\|\mathsf{G} - \mathsf{G}_m\|_\infty \leq \sum_{n=m+1}^{\infty} |c_n| \left\| \frac{1}{1-\lambda_n \cdot} \right\|_\infty = \sum_{n=m+1}^{\infty} |c_n| \frac{1}{1-|\lambda_n|} \to 0, \text{ as } m \to \infty$$

using Lemma 10.23. The matrix case follows from applying the above to each component of the matrix. We now turn to the case of the nuclear norm. It is easily seen that $[\lambda_n, 1; \lambda_n, 1]$ is a realization of $z \mapsto \frac{1}{1-\lambda_n z}$. The observation Lyapunov equation is $|\lambda_n|^2 L_C - L_C + |\lambda_n|^2$, which has as unique solution $L_C := \frac{|\lambda_n|^2}{1-|\lambda_n|^2}$. The control Lyapunov equation, $|\lambda_n|^2 L_B - L_B + 1$, has as unique solution $L_B := \frac{1}{1-|\lambda_n|^2}$. It follows that the unique positive Hankel singular value is $\frac{|\lambda_n|}{1-|\lambda_n|^2}$. So this is the nuclear norm of $\frac{1}{1-\lambda_n}$. Since $|\lambda_n| < 1$ we have that this nuclear norm is smaller than $\frac{1}{1-|\lambda_n|}$. Convergence in the nuclear norm is shown, using this, similarly to convergence in the H^∞ norm. □

The earlier proven continuity of singular values and Schmidt pairs applied to the nuclear approximants gives the following.

Lemma 10.26. *Assume that $\mathsf{G} \in H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{Y}))$, with $\mathscr{U}$ and $\mathscr{Y}$ finite-dimensional, has a nuclear Hankel map. Denote its (m, c, λ) nuclear approximant by G_m. Denote the singular values of the Hankel map of G by σ_i. Assume that the singular values are ordered in decreasing magnitude and repeated according to their multiplicity. Then for the singular values σ_i^m of the Hankel map of G_m, also ordered in decreasing magnitude and repeated according to their multiplicity, we have $\sigma_i^m \to \sigma_i$. Let (v_i^m, w_i^m) be Schmidt pairs for the Hankel map of G_m. There exists a subsequence G_{m_k} of G_m and Schmidt pairs (v_i, w_i) for the Hankel map of G such that $\|v_i^{m_k} - v_i\| \to 0$ and $\|w_i^{m_k} - w_i\| \to 0$ as $k \to \infty$.*

Proof. Lemma 10.25 show that G_m converges to G in the infinity norm. This implies that the Hankel map of G_m converges to the Hankel map of G in the operator norm. The result then follows from Lemma 10.21. □

Remark 10.27. Lemma 10.26 implies that, for m large enough and n such that $\sigma_n > \sigma_{n+1}$, we have $\sigma_n^m > \sigma_{n+1}^m$. So if the n-dimensional Lyapunov-balanced trucation of G is well-defined, then so is the n-dimensional Lyapunov-balanced trucation of G_m for m large enough. By Lemma 10.9 the transfer function of this n-dimensional Lyapunov-balanced trucation of G_m does not depend on the sequence of eigenvectors chosen. Denote this transfer function by G_m^n.

Lemma 10.28. *With the assumptions and the notation as in Definition 10.24 and Remark 10.27 we have (if G has at least n nonzero Hankel singular values)*

$$\|\mathsf{G}^n - \mathsf{G}_{m_k}^n\|_\infty \to 0 \text{ as } k \to \infty.$$

Proof. For notational convenience we will assume that G_m^n has been replaced by the subsequence $\mathsf{G}_{m_k}^n$. By Lemma 10.26 we have convergence of the singular values as $m \to \infty$. Given Schmidt pairs (v_i^m, w_i^m) of the Hankel map of G_m converge to certain Schmidt pairs (v_i, w_i) of the Hankel map of G by the same lemma. Denote by Σ_m the Lyapunov-balanced realization of G_m with respect to the eigenvectors (w_i^m) and by Σ_m^n its n-dimensional Lyapunov-balanced truncation. It follows from Remark 10.27 that the transfer function of Σ_m^n equals G_m^n. Denote by Σ the Lyapunov-balanced realization of G with respect to the sequence of eigenvectors (w_i) and by Σ^n its n-dimensional Lyapunov-balanced truncation. Using Proposition 10.9 it follows that the transfer function of Σ^n equals G^n. Note that we have

$$A_m^n = P_{\mathscr{X}_m^n} A_m|_{\mathscr{X}_m^n}, \quad A^n = P_{\mathscr{X}^n} A|_{\mathscr{X}^n}.$$

Let $x \in \mathscr{X}$. Then we have $x = \sum_{j=1}^{\infty} \langle x, w_j^m \rangle w_j^m$. It follows that

$$\langle A_m x, w_i^m \rangle = \sum_{j=1}^{\infty} \langle x, w_j^m \rangle \langle A_m w_j^m, w_i^m \rangle.$$

Using the explicit description from Remark 10.5 we obtain

$$\langle A_m x, w_i^m \rangle = \sum_{j=1}^{\infty} \langle x, w_j^m \rangle \sqrt{\frac{\sigma_j^m}{\sigma_i^m}} \langle A^{\mathrm{bs}} w_j^m, w_i^m \rangle.$$

Since we have $A_m^n x = \sum_{i=1}^{n} \langle A_m x, w_i^m \rangle w_i^m$ it follows that

$$A_m^n x = \sum_{i=1}^{n} \sum_{j=1}^{\infty} \langle x, w_j^m \rangle \sqrt{\frac{\sigma_j^m}{\sigma_i^m}} \langle A^{\mathrm{bs}} w_j^m, w_i^m \rangle w_i^m.$$

Similarly we have

$$A^n x = \sum_{i=1}^{n} \sum_{j=1}^{\infty} \langle x, w_j \rangle \sqrt{\frac{\sigma_j}{\sigma_i}} \langle A^{\mathrm{bs}} w_j, w_i \rangle w_i.$$

Along similar lines it follows that

$$B_m^n u = \sum_{k=1}^{n} \langle B_m^n, w_k^m \rangle w_k^m = \sum_{k=1}^{n} \sqrt{\sigma_k^m} \langle \underline{u}, v_k^m \rangle w_k^m,$$

$$B^n u = \sum_{k=1}^{n} \langle B^n, w_k \rangle w_k = \sum_{k=1}^{n} \sqrt{\sigma_k} \langle \underline{u}, v_k \rangle w_k,$$

where $\underline{u} : \mathbb{Z}^- \to \mathscr{U}$ is defined by $\underline{u}_{-1} = u$ and otherwise zero. It also follows that

$$C_m^n x = \sum_{i=1}^{p} \sum_{k=1}^{\infty} \sqrt{\sigma_k^m} \langle x, w_k^m \rangle \langle C^{\mathrm{bs}} w_k^m, y_i \rangle y_i,$$

$$C_m^n x = \sum_{i=1}^{p} \sum_{k=1}^{\infty} \sqrt{\sigma_k} \langle x, w_k \rangle \langle C^{\mathrm{bs}} w_k, y_i \rangle y_i,$$

where $p = \dim \mathscr{Y}$. From the above we obtain that

$$C_m^n A_m^n B_m^n u = \sum_{k=1}^{p} \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_j^m \langle \underline{u}, v_j^m \rangle \langle A^{\mathrm{bs}} w_j^m, w_i^m \rangle \langle C^{\mathrm{bs}} w_k^m, y_k \rangle y_k,$$

and

$$C^n A^n B^n u = \sum_{k=1}^{p} \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_j \langle \underline{u}, v_j \rangle \langle A^{\mathrm{bs}} w_j, w_i \rangle \langle C^{\mathrm{bs}} w_k, y_k \rangle y_k.$$

It follows that $C_m^n A_m^n B_m^n \to C^n A^n B^n$ as $m \to \infty$. Similarly we obtain for all $k \in \mathbb{Z}^+$ that $C_m^n (A_m^n)^k B_m^n \to C^n (A^n)^k B^n$ as $m \to \infty$. It follows that we have convergence of the Taylor coefficients of G_m^n to those of G^n.

Proposition 10.18 shows that Σ_m^n (for m large enough) and Σ^n are exponentially stable (note that since the Hankel singular values of Σ_m converge to those of Σ it follows that Σ_m has at least n nonzero Hankel singular values for m large enough). By dominated convergence it now follows that the transfer function G_m^n converges to G^n pointwise in $\overline{\mathbb{D}}$. Since $\overline{\mathbb{D}}$ is compact this is equivalent to uniform convergence and so we have $\|\mathsf{G}_m^n - \mathsf{G}^n\|_{H^\infty(\mathcal{L}(\mathscr{U},\mathscr{Y}))} \to 0$ as desired. □

Proof of Proposition 10.11. We have

$$\|\mathsf{G} - \mathsf{G}^n\|_\infty \leq \|\mathsf{G} - \mathsf{G}_m\|_\infty + \|\mathsf{G}_m - \mathsf{G}_m^n\|_\infty + \|\mathsf{G}_m^n - \mathsf{G}^n\|_\infty. \tag{10.2}$$

Let $\varepsilon > 0$. From Lemma 10.25 we have $\|\mathsf{G} - \mathsf{G}_m\|_\infty \to 0$ as $m \to \infty$ and so there exists a M_1 such that if $m \geq M_1$ then $\|\mathsf{G} - \mathsf{G}_m\|_\infty < \varepsilon$. From Lemma 10.28 we have $\|\mathsf{G}^n - \mathsf{G}_m^n\|_\infty \to 0$ as $m \to \infty$. Hence there exists a M_2 such that if $m \geq M_2$ then $\|\mathsf{G}_m^n - \mathsf{G}^n\|_\infty < \varepsilon$. We now consider the second term on the right-hand side of (10.2). From finite-dimensional theory (see Zhou, Doyle and Glover [103, page 566]) we obtain

$$\|\mathsf{G}_m - \mathsf{G}_m^n\|_\infty \leq 2 \sum_{i=n+1}^{m} \sigma_i^m.$$

Since $\sigma_i^m \to \sigma_i$ by Lemma 10.26 there exists an M_3 such that if $m \geq M_3$, then $|\sigma_i^m - \sigma_i| < \varepsilon/n$ from which it follows that

$$\sum_{i=1}^{n} \sigma_i^m > \sum_{i=1}^{n} \sigma_i - \varepsilon.$$

By the convergence in the nuclear norm of G_m to G (Lemma 10.25) we have $\|\mathsf{G}_m\|_N \to \|\mathsf{G}\|_N$, which implies the existence of a M_4 such that if $m \geq M_4$, then

$$\sum_{i=1}^{m} \sigma_i^m \leq \sum_{i=1}^{\infty} \sigma_i + \varepsilon.$$

Combining the above three inequalities we see that if $m \geq M_3$ and $m \geq M_4$, then

$$\|\mathsf{G}_m - \mathsf{G}_m^n\|_\infty \leq 2 \sum_{i=n+1}^{m} \sigma_i^m = 2\sum_{i=1}^{m} \sigma_i^m - 2\sum_{i=1}^{n} \sigma_i^m$$

$$\leq 2\sum_{i=1}^{\infty} \sigma_i + 2\varepsilon - 2\sum_{i=1}^{n} \sigma_i + 2\varepsilon = 2\sum_{i=n+1}^{\infty} \sigma_i + 4\varepsilon.$$

Define $M = \max\{M_1, M_2, M_3, M_4\}$. Then for $m \geq M$

$$\|\mathsf{G} - \mathsf{G}^n\|_\infty \leq 2 \sum_{i=n+1}^{\infty} \sigma_i + 6\varepsilon.$$

Since this holds for all $\varepsilon > 0$ we obtain

$$\|\mathsf{G} - \mathsf{G}^n\|_\infty \leq 2 \sum_{i=n+1}^{\infty} \sigma_i.$$

□

10.2 LQG-balanced realizations

In this section we study LQG-balanced realizations. The optimal closed-loop system from Definition 6.32 plays a crucial role in relating LQG-balanced realizations and Lyapunov-balanced realizations.

Proposition 10.29. *Let Σ be an input and output stabilizable discrete-time system. Let $Q^{\min}$ and $P^{\min}$ denote the optimal cost operators of the system and of its dual system, respectively, and let L_B and L_C denote the gramians of the optimal closed-loop system. Then $\lambda \in \sigma(P^{\min}Q^{\min})$ if and only if $\lambda/(1+\lambda) \in \sigma(L_B L_C)$.*

Proof. From Propositions 6.35 and 6.43 we obtain the equality $L_B L_C = (I + P^{\min}Q^{\min})^{-1}P^{\min}Q^{\min}$, from which it follows that $I - L_B L_C = (I + P^{\min}Q^{\min})^{-1}$. So $1 \in \rho(L_B L_C)$, and $P^{\min}Q^{\min} = (I - L_B L_C)^{-1} L_B L_C$. Let $\lambda \in \mathbb{C} - \{-1\}$ and define $\mu := \lambda/(1+\lambda)$, then $\lambda = \mu/(1-\mu)$. We have

$$\begin{aligned}
\lambda I - P^{\min}Q^{\min} &= \frac{\mu}{1-\mu} I - L_B L_C (I - L_B L_C)^{-1} \\
&= \frac{1}{1-\mu}\left[\mu I - (1-\mu) L_B L_C (I - L_B L_C)^{-1}\right] \\
&= \frac{1}{1-\mu}\left[\mu(I - L_B L_C) - (1-\mu) L_B L_C\right](I - L_B L_C)^{-1} \\
&= \frac{1}{1-\mu}(\mu I - L_B L_C)(I - L_B L_C)^{-1}.
\end{aligned}$$

This shows that $\lambda \in \sigma(P^{\min}Q^{\min})$ if and only if $\mu = \lambda/(1+\lambda) \in \sigma(L_B L_C)$. □

Proposition 10.30. *Let Σ_i with $i = 1, 2$ be two input and output stabilizable discrete-time systems. Let $Q_i^{\min}$ and $P_i^{\min}$ denote the optimal cost operators of the system and of its dual system, respectively. If the two systems have the same transfer function then, with the possible exception of zero, the spectra of $P_1^{\min}Q_1^{\min}$ and $P_2^{\min}Q_2^{\min}$ are equal.*

Proof. Denote the gramians of the optimal closed-loop system of Σ_i by L_{B_i} and L_{C_i}. Then according to Proposition 10.29, the proposition would be proved if the nonzero elements in the spectrum of $L_{B_1}L_{C_1}$ equal the nonzero elements in the spectrum of $L_{B_2}L_{C_2}$. Since the transfer function of both optimal closed-loop systems is a normalized weakly right-coprime factor of the transfer function of both the Σ_i by Proposition 7.11, there exists by Proposition 7.15 a unitary $V \in \mathcal{L}(\mathscr{U})$ such that $[\mathsf{M}_2; \mathsf{N}_2] = [\mathsf{M}_1; \mathsf{N}_1]V$. For the Hankel maps of the optimal closed-loop systems this implies $\mathcal{H}_2 = \mathcal{H}_1 V$, which implies that $\mathcal{H}_2\mathcal{H}_2^* = \mathcal{H}_1\mathcal{H}_1^*$. Since for arbitrary bounded operators S and T we have that the nonzero elements in the spectrum of ST equal the nonzero elements in the spectrum of TS (Lemma 3.16), we have that the nonzero elements in the spectrum of $L_B L_C = \mathcal{B}\mathcal{B}^*\mathcal{C}^*\mathcal{C}$ equal the nonzero elements in the spectrum of $\mathcal{H}\mathcal{H}^* = \mathcal{C}\mathcal{B}\mathcal{B}^*\mathcal{C}^*$. This shows that the nonzero elements in the spectrum of $L_{B_1}L_{C_1}$ equal the nonzero elements in the spectrum of $L_{B_2}L_{C_2}$. □

Definition 10.31. Let Σ be an input and output stabilizable system. Denote the optimal cost operator by $Q^{\min}$ and the optimal cost operator of the dual system by $P^{\min}$. The square roots of the points in the spectrum of $P^{\min}Q^{\min}$, with the exception of zero, are called the **LQG-characteristic values** of Σ.

Note that Proposition 10.30 shows that the LQG-characteristic values only depend on the transfer function, not on the particular realization.

Corollary 10.32. *Let Σ be an input and output stabilizable system. Denote the Hankel map of a realization of a normalized weakly right-coprime factor of the transfer function of Σ by $\mathcal{H}$. Then μ is a LQG-characteristic value of Σ if and only if $\mu \neq 0$ and $\mu^2/(1+\mu^2) \in \sigma(\mathcal{H}\mathcal{H}^*)$. In particular we have $\mu_1^2/(1+\mu_1^2) = \|\mathcal{H}\|^2$ for the largest LQG-characteristic value μ_1.*

Proof. The relationship between the LQG-characteristic values and the spectrum of $\mathcal{H}^*\mathcal{H}$ was proven in the proof of Proposition 10.30. The formula for the largest LQG-characteristic value follows using that since $\mathcal{H}\mathcal{H}^*$ is nonnegative self-adjoint its norm equals the largest eigenvalue and also equals $\|\mathcal{H}\|^2$. □

Definition 10.33. A discrete-time system is called **LQG-balanced** if it is input and output stabilizable and the optimal cost operator of the system and that of its dual system are equal.

The following result shows the existence and uniqueness of LQG-balanced realizations.

Proposition 10.34. *Let $\mathsf{G} : D(\mathsf{G}) \subset \mathbb{C} \to \mathcal{L}(\mathscr{U}, \mathscr{Y})$ be holomorphic with $0 \in D(\mathsf{G})$ and assume that G has a strongly right-coprime factorization. Then G has a minimal LQG-balanced realization. Conversely, the transfer function of a LQG-balanced realization has a strongly right coprime-factorization. Minimal LQG-balanced realizations are unique up to a unitary similarity transformation in the state space.*

Proof. Using Proposition 7.24 it follows from the assumption that G has a strongly right-coprime factorization that G has a normalized strongly right-coprime factorization. Denote a normalized strongly right-coprime factor by $[\mathsf{M}; \mathsf{N}]$. From Proposition 10.2 we obtain that $[\mathsf{M}; \mathsf{N}]$ has a minimal Lyapunov-balanced realization $\check{\Sigma}$. It follows from Proposition 5.7 that $\check{\Sigma}$ is energy preserving with storage operator L equal to the gramian. Define the discrete-time system Σ as in Proposition 2.23. Propositions 6.45 and 6.46 show that L is a solution of the control algebraic Riccati equation and $(I - L^2)^{-1}L$ is a solution of the filter algebraic Riccati equation of Σ. Obviously $\check{\Sigma}$ is the

Riccati closed-loop system of Σ corresponding to the solution $Q = L$. Since $\check{\Sigma}$ is strongly stable by Proposition 10.2 it follows from Proposition 6.38 that L is the unique solution of the control algebraic Riccati equation of Σ. We will show that $(I-L^2)^{-1}L$ is the unique solution of the filter algebraic Riccati equation of Σ. Suppose that there are two solutions P_1 and P_2. Then, by Proposition 6.39, $(I + P_iL)^{-1}P_i$, $i = 1, 2$, are both solutions of the control Lypunov equation of $\check{\Sigma}$. Since the dual system of the Lyapunov-balanced realization $\check{\Sigma}$ is strongly stable by Proposition 10.2 it follows from Proposition 3.14 that the control Lypunov equation of $\check{\Sigma}$ has a unique solution. This implies that $(I + P_1L)^{-1}P_1 = (I + P_2L)^{-1}P_2$, from which $P_1 = P_2$ easily follows. Now apply the similarity transformation $(I - L^2)^{1/4}$ to Σ to obtain a system Σ_{LQG}. It is easily seen that this system has $L(I-L^2)^{1/2}$ as the unique solution to both its control and filter algebraic Riccati equation. Hence Σ_{LQG} is LQG-balanced.

We now show that since $\check{\Sigma}$ is minimal, so is Σ_{LQG}. Since $\check{\Sigma}$ is minimal its gramian L is positive by Proposition 3.11. This implies that the optimal cost operator $L(I - L^2)^{1/2}$ of Σ_{LQG} is positive. It follows using Proposition 6.12 that Σ_{LQG} is approximately observable. Since the optimal cost operator of the dual system of Σ_{LQG} is positive it follows that Σ_{LQG} is approximately controllable.

We now show the uniqueness of minimal LQG-balanced realizations. Assume that Σ_i, $i = 1, 2$ are both minimal LQG-balanced realizations of the same transfer function. Denote the optimal cost operator of Σ_i by $Q_i^{\min}$. First apply the similarity transformation $(I + (Q_i^{\min})^2)^{1/4}$ to Σ_i and then construct the optimal closed-loop systems $\check{\Sigma}_i$. Using Propositions 6.35 and 6.43 it follows that $\check{\Sigma}_i$ is Lyapunov-balanced. The optimal control operator of Σ_i is positive by minimality (using Proposition 6.12), which implies that the gramian of $\check{\Sigma}_i$ is positive, which implies that $\check{\Sigma}_i$ is minimal using Propositions 3.11 and 3.15. Denote the transfer function of $\check{\Sigma}_i$ by $[\mathsf{M}_i; \mathsf{N}_i]$. It follows from Proposition 7.11 that these are normalized weakly right-coprime factors of the transfer function of the Σ_i. From Proposition 7.15 it follows that there exists a unitary $V \in \mathcal{L}(\mathscr{U})$ such that $[\mathsf{M}_1; \mathsf{N}_1] = [\mathsf{M}_2; \mathsf{N}_2]V$. It is easily seen that if we apply the input-space transformation V to $\check{\Sigma}_2$, then we obtain a minimal Lyapunov-balanced realization of $[\mathsf{M}_1; \mathsf{N}_1]$. Since $\check{\Sigma}_1$ is also a minimal Lyapunov-balanced realization of $[\mathsf{M}_1; \mathsf{N}_1]$ they are related by a unitary similarity transformation $U \in \mathcal{L}(\mathscr{X})$ by Proposition 10.2. From this it follows, using (2.5) which gives the system operator of Σ_i in terms of that of $\check{\Sigma}_i$, that Σ_1 and Σ_2 are related by the same unitary similarity transformation. Note that the operator V cancels when we apply (2.5). □

Corollary 10.35. *Let Σ be LQG-balanced with optimal cost operator $Q^{\min}$.*

Apply the similarity transformation $(I + (Q^{\min})^2)^{1/4}$ *to* Σ. *Denote by* Σ_{LYAP} *the optimal closed-loop system of this transformed system. Then* Σ_{LYAP} *is Lyapunov-balanced and its transfer function is a normalized strongly right-coprime factor of the transfer function of* Σ.

Proof. This follows from the proof of Proposition 10.34. □

Definition 10.36. The discrete-time system Σ_{LYAP} is called the Lyapunov-balanced system corresponding to the LQG-balanced discrete-time system Σ.

Definition 10.37. A discrete-time system is called **compact LQG-balanced** if it is LQG-balanced and its optimal cost operator is compact.

Definition 10.38. Given a compact LQG-balanced realization Σ, let (w_i) be an ordered sequence of eigenvectors of the optimal cost operator $Q^{\min}$ (the ordering is such that the corresponding eigenvalues μ_i form a non-increasing sequence). Let $n \in \mathbb{Z}^+$ be such that $\mu_n > \mu_{n+1}$. The **truncated LQG-balanced realization** of dimension n with respect to the sequence of eigenvectors (w_i) is defined as the restriction/projection of Σ onto $\mathscr{X}_n := \{w_i : i = 1, \ldots, n\}$.

Remark 10.39. The sequence (w_i) from Definition 10.38 is also an ordered basis of eigenvectors for the gramian L of the corresponding compact Lyapunov-balanced realization Σ_{LYAP}. Indeed, since $L = (I + (Q^{\min})^2)^{-1/2} Q^{\min}$ we have $\sigma_i = \mu_i / \sqrt{1 + \mu_i^2}$ for the corresponding eigenvalues.

Lemma 10.40. *Let* Σ *be a discrete-time system, let* $[F, G]$ *be an admissible feedback pair and denote by* $\Sigma_{[F,G]}$ *the corresponding closed-loop system. Let* $\tilde{\mathscr{X}} \subset \mathscr{X}$ *be a subspace. Let* $\tilde{\Sigma}$ *be the projection/restriction of* Σ *onto* $\tilde{\mathscr{X}}$. *Then* $[F|_{\tilde{\mathscr{X}}}, G]$ *is an admissible feedback pair for* $\tilde{\Sigma}$ *and the corresponding closed-loop system equals the projection/restriction of* $\Sigma_{[F,G]}$ *onto* $\tilde{\mathscr{X}}$.

Proof. This is easily seen from the definitions. □

Lemma 10.41. *Let* Σ *be a discrete-time system and* $\tilde{\mathscr{X}} \subset \mathscr{X}$ *a subspace. Let* $\tilde{\Sigma}$ *be the projection/restriction of* Σ *onto* $\tilde{\mathscr{X}}$. *Let* $T \in \mathcal{L}(\mathscr{X})$ *have a bounded inverse and map* $\tilde{\mathscr{X}}$ *onto itself. Denote the discrete-time system obtained from* Σ *by applying the similarity transformation* T *by* Σ_T. *Denote the discrete-time system obtained from* $\tilde{\Sigma}$ *by applying the similarity transformation* $T|_{\tilde{\mathscr{X}}}$ *by* $\tilde{\Sigma}_T$. *Then* $\tilde{\Sigma}_T$ *is the projection/restriction of* Σ_T *onto* $\tilde{\mathscr{X}}$.

Proof. This follows easily. □

Proposition 10.42. *Let Σ be a compact LQG-balanced realization. Let (w_i) be an ordered sequence of eigenvectors of the optimal cost operator $Q^{\min}$ (the ordering is such that the corresponding eigenvalues μ_i form a nonincreasing sequence). Denote the optimal feedback pair for Σ as given in Proposition 6.33 by $[F^{\min}, G^{\min}]$. Denote the truncated LQG-balanced realization with respect to the sequence (w_i) by Σ_n. Define Σ_n^{cl} as the closed-loop system of Σ_n with the feedback pair $[F^{\min}|_{\mathscr{X}_n}, G^{\min}]$. Apply the similarity transformation $T := (I + Q^{\min}|^2_{\mathscr{X}_n})^{1/4}$ to Σ_n^{cl} to obtain $\Sigma_{n,\mathrm{LYAP}}$. Define $\Sigma_{\mathrm{LYAP},n}$ as the truncation of the Lyapunov-balanced realization Σ_{LYAP} corresponding to the LQG-balanced discrete-time system Σ. Then $\Sigma_{n,\mathrm{LYAP}} = \Sigma_{\mathrm{LYAP},n}$.*

Proof. This follows using Lemmas 10.40 and 10.41. □

Corollary 10.43. *Using the assumptions and notation of Proposition 10.42 we have that the transfer function of $\Sigma_{\mathrm{LYAP},n}$ is a right factor of the transfer function of Σ_n.*

Proof. The discrete-time system $\Sigma_{\mathrm{LYAP},n}$ is input-output stable by Proposition 10.18 combined with Proposition 3.28. Since $\Sigma_{n,\mathrm{LYAP}}$ is obtained from Σ_n by feedback and a similarity transformation we have the relation in (2.5) (up to the similarity transformation) between their system operators. The relationship between their transfer functions as in Proposition 2.23 follows. Since $\Sigma_{n,L} = \Sigma_{L,n}$ by Proposition 10.42 we obtain the desired result. □

Remark 10.44. We use the notation of Proposition 10.42. The feedback pair $[F^{\min}|_{\mathscr{X}_n}, G^{\min}]$ is in general not the optimal feedback pair for Σ_n. It follows from Proposition 10.42 that Σ_n^{cl} is input stable, output stable and input-output stable. It follows that Σ_n is output stabilizable. Hence it has an optimal cost operator $Q_n^{\min}$. We have for every $x_0 \in \mathscr{X}_n$ that $\langle Q_n^{\min} x_0, x_0\rangle \leq \langle Q^{\min}|_{\mathscr{X}_n} x_0, x_0\rangle$, where $Q^{\min}$ is the optimal cost operator of Σ. Since $\mu_1 = \|Q^{\min}\| = \|Q^{\min}|_{\mathscr{X}_n}\|$ and for μ_1^n, the largest LQG characteristic value of Σ_n, we have $\mu_1^n = \|Q_n^{\min}\|$ we obtain $\mu_1^n \leq \mu_1$.

Proposition 10.45. *Let Σ be a compact LQG-balanced realization with $\mathscr{U}$ and $\mathscr{Y}$ finite-dimensional. Assume that the LQG-characteristic values (μ_i) form a summable sequence. Define G^n as the transfer function of a truncated LQG-balanced realization of dimension n of Σ. Then we have*

$$\vec{\delta}_g(\mathsf{G}, \mathsf{G}^n) \leq 2 \sum_{i=n+1}^{\infty} \frac{\mu_i}{\sqrt{1+\mu_i^2}}.$$

Proof. Denote, as in Proposition 10.42, by Σ_{LYAP} the Lyapunov-balanced realization corresponding to the LQG-balanced realization Σ and by $\Sigma_{\mathrm{LYAP},n}$

its truncation. It follows from Corollary 10.43 that the transfer function $[\mathsf{M}^n; \mathsf{N}^n]$ of $\Sigma_{\text{LYAP},n}$ is a right factor of G^n. Let $[\mathsf{M}; \mathsf{N}]$ denote the transfer function of Σ_{LYAP}. It follows from Corollary 10.35 that $[\mathsf{M}; \mathsf{N}]$ is a normalized strongly right-coprime factor of G. Proposition 9.15 shows that

$$\vec{\delta}_g(\mathsf{G}, \mathsf{G}^n) \leq \left\| \begin{bmatrix} \mathsf{M} \\ \mathsf{N} \end{bmatrix} - \begin{bmatrix} \mathsf{M}^n \\ \mathsf{N}^n \end{bmatrix} \right\|.$$

Using the relation between the LQG-characteristic values of Σ and the Hankel singular values of Σ_{LYAP} from Remark 10.39 we see that Σ_{LYAP} has a nuclear Hankel map. Proposition 10.11 shows that

$$\left\| \begin{bmatrix} \mathsf{M} \\ \mathsf{N} \end{bmatrix} - \begin{bmatrix} \mathsf{M}^n \\ \mathsf{N}^n \end{bmatrix} \right\| \leq 2 \sum_{i=n+1}^{\infty} \frac{\mu_i}{\sqrt{1+\mu_i^2}}.$$

The desired result follows. □

Proposition 10.46. *Let Σ be a compact LQG-balanced realization with $\mathscr{U}$ and $\mathscr{Y}$ finite-dimensional. Assume that the LQG-characteristic values (μ_i) form a summable sequence. Define G^n as the transfer function of a truncated LQG-balanced realization of dimension n of Σ. Then there exists a $N \in \mathbb{Z}^+$ such that*

$$2 \sum_{i=N+1}^{\infty} \frac{\mu_i}{\sqrt{1+\mu_i^2}} < \frac{1}{\sqrt{1+\mu_1^2}}. \tag{10.3}$$

For $n \geq N$ we have

$$\delta_g(\mathsf{G}, \mathsf{G}^n) \leq 2 \sum_{i=n+1}^{\infty} \frac{\mu_i}{\sqrt{1+\mu_i^2}}.$$

Proof. The existence of N such that (10.3) holds follows from the assumption that (μ_i) is a summable sequence.

From the proof of Proposition 10.45 we obtain that the function G has a normalized strongly right-coprime factor $[\mathsf{M}; \mathsf{N}]$ and G^n has a right factor $[\mathsf{M}^n; \mathsf{N}^n]$ such that

$$\left\| \begin{bmatrix} \mathsf{M} - \mathsf{M}_n \\ \mathsf{N} - \mathsf{N}_n \end{bmatrix} \right\| \leq 2 \sum_{i=n+1}^{\infty} \frac{\mu_i}{\sqrt{1+\mu_i^2}}.$$

It follows from Proposition 7.32 that since $n \geq N$ we have that $[\mathsf{M}^n; \mathsf{N}^n]$ is strongly right-coprime. Here we have used that the right-hand side of

(10.3) is exactly $\sqrt{1-\|\mathcal{H}\|^2}$, where $\mathcal{H}$ is the Hankel map corresponding to $[\mathsf{M};\mathsf{N}]$, a formula that follows from the one given in Corollary 10.32 for the largest LQG-characteristic value. It follows using Proposition 9.18 that for all $n \geq N$ we have $\delta(\mathsf{G},\mathsf{G}^n) = \vec{\delta}(\mathsf{G},\mathsf{G}^n) = \vec{\delta}(\mathsf{G}^n,\mathsf{G})$. The result now follows using Proposition 10.45. □

Proposition 10.47. *Let Σ be a compact LQG-balanced realization with $\mathscr{U}$ and $\mathscr{Y}$ finite-dimensional. Assume that the LQG-characteristic values (μ_i) form a summable sequence. Define G^n as the transfer function of a truncated LQG-balanced realization of dimension n of Σ. Then there exists a $N \in \mathbb{Z}^+$ such that for all $n \geq N$*

$$2\sum_{i=n+1}^{\infty} \frac{\mu_i}{\sqrt{1+\mu_i^2}} < \frac{1}{\sqrt{1+(\mu_1^n)^2}}, \tag{10.4}$$

where μ_1^n is the largest LQG-characteristic value of G^n. For given $n \geq N$ choose ε_n such that

$$2\sum_{i=n+1}^{\infty} \frac{\mu_i}{\sqrt{1+\mu_i^2}} < \varepsilon_n < \frac{1}{\sqrt{1+(\mu_1^n)^2}}.$$

Then the ε_n-robust right factor stabilizing feedback function for G^n stabilizes G.

Proof. We have $\mu_1^n \leq \mu_1$ by Remark 10.44. This implies that the right hand side of (10.4) is bounded from below by $1/\sqrt{1+\mu_1^2}$. Formula (10.4) then follows from the fact that (μ_i) forms a summable sequence. With the indicated choice of ε_n we obtain from Proposition 10.46 that $\delta_g(\mathsf{G},\mathsf{G}^n) < \varepsilon_n$. The result then follows from Proposition 9.25 using that the right hand side of (10.4) equals $\sqrt{1-\|\mathcal{H}_n\|^2}$, where $\mathcal{H}_n$ is the Hankel map of a normalized strongly right-coprime factor of G^n. □

Remark 10.48. Note that since G^n is rational it has a finite-dimensional state space realization. Consequently, Corollary 8.19 implies that the ε_n-robust right factor stabilizing feedback function mentioned in Proposition 10.47 can be chosen to be rational.

Remark 10.49. Consider the situation as in Proposition 10.47. Let G_Δ be such that $\delta_g(\mathsf{G},\mathsf{G}_\Delta) < 1/\sqrt{1+\mu_1^2}$, where μ_1 is the largest LQG-characteristic value of G. Then there exists a $N \in \mathbb{Z}^+$ such that for $n \geq N$ the ε_n-robust right factor stabilizing feedback function for G^n, with ε_n chosen sufficiently close to $1/\sqrt{1+(\mu_1^n)^2}$, stabilizes G_Δ. This follows using the triangle inequality.

We give a sufficient condition for the LQG-characteristic values (μ_i) to form a summable sequence.

Proposition 10.50. *Let Σ be an exponentially stabilizable and detectable system with finite-dimensional input and output spaces. Then its LQG-characteristic values (μ_i) form a summable sequence.*

Proof. It follows from Corollary 4.13 that the optimal closed-loop system Σ^{opt} of Σ is exponentially stable. Proposition 10.17 then shows that Σ^{opt} has a nuclear Hankel map. It follows from Remark 10.39 that the Hankel singular values of Σ^{opt} equal $\mu_i/\sqrt{1+\mu_i}$. So $(\mu_i/\sqrt{1+\mu_i})$ forms a summable sequence. It is easily seen that this is equivalent to (μ_i) being a summable sequence. □

Notes

Lyapunov-balanced realization were introduced by Moore [57] for finite dimensional systems. LQG-balanced realizations were introduced by Verriest [93], also in the context of finite dimensional systems. See also Jonckheere and Silverman [41] for LQG-balanced realizations for finite-dimensional systems. Propositions 10.2 and 10.4 are due to Young [99], except for the statement on strong stability in Proposition 10.2, which is due to Ober and Wu [63]. Proposition 10.11 is due to Glover, Curtain and Partington [35] in the continuous-time case and based on these ideas by Bonnet [5] in the discrete-time case. Both references treat only the case with nonrepeating eigenvalues, but the general case considered here follows along the same lines as was already indicated in [35]. Proposition 10.17 is based on Curtain and Sasane [9]. The results on LQG-balanced realizations are based on Opmeer and Curtain [71] and Opmeer [68].