On the Connection between Balanced Proper Orthogonal Decomposition, Balanced Truncation, and Metric Complexity Theory for Infinite Dimensional Systems

Seddik M. Djouadi

Abstract—In this paper, the connection between two important model reduction techniques, namely balanced proper orthogonal decomposition (POD) and balanced truncation is investigated for infinite dimensional systems. In particular, balanced POD is shown to be optimal in the sense of distance minimization in a space of integral operators under the Hilbert-Schmidt norm. Whereas balanced truncation is shown to be a particular case of balanced POD for infinite dimensional systems for which the impulse response satisfies certain finite energy constraints. POD and balanced truncation are related to certain notions of metric complexity theory. In particular both are shown to minimize different n-widths of partial differential equation solutions including the Kolmogorov, Gelfand, linear and Bernstein n-widths. The n-widths quantify inherent and representation errors due to lack of data and loss of information.

I. INTRODUCTION

In this paper, we study the possible connections between two important model reduction techniques, namely, balanced truncation introduced in [14], and balanced proper orthogonal decomposition (POD) introduced in [2] for finite dimensional systems. Both model reduction techniques have been widely applied in diverse areas, see for example, [1][9][6][2][3][4][5]. Balanced truncation has been extended to infinite dimensional systems in [7][18]. Balanced POD has been extended to the infinite dimensional setting in [8]. Balanced POD is a tractable method for computing approximate balanced truncations, that has computational cost similar to POD [2], where connections between balanced POD and balanced truncation are discussed. In particular, it is pointed that balanced truncation may be viewed as POD of a particular dataset, using the observability Gramian in an inner product [2].

In this paper, we take a different direction and show that the balanced POD and balanced truncations coincide when using a particular balanced realization for infinite dimensional systems. This is carried out by first showing that balanced POD is optimal, as far as best approximation of a certain class of impulse responses of stable linear infinite dimensional systems by stable impulse responses of finite dimensional systems in some \( L^2 \)-norm. The latter is shown to be equivalent to optimal approximation of Hilbert-Schmidt Hankel operators by finite rank Hankel operators in the Hilbert-Schmidt norm.

We relate balanced POD and balanced truncation to certain notions in metric complexity theory. In particular both are shown to minimize different n-widths of infinite dimensional systems, including the Kolmogorov, Gelfand, linear and Bernstein n-widths. The n-widths quantify inherent and representation errors due to lack of data and loss of information. For example, the Kolmogorov n-width quantifies the representation error due to inaccurate representation of the range space of the system Hankel operator. It represents the loss of information in the information processing stage. The Kolomogorov n-width characterizes the representation complexity of the model reduction problem. The Gelfand n-width characterizes the experimental complexity of the information collecting stage using simulation or identification. It is related to the inherent error due to lack of data and inaccurate measurements. The inverse of the Gelfand n-width gives the least number of measurements needed to reduce the modeling uncertainty to a predetermined value.

The rest of the paper is organized as follows. Section II contains the problem formulation. In section III a key balanced truncation is presented and is based on [18]. In section IV truncation of the balanced realization is shown to yield balanced POD. Section V discusses the relationship between balanced POD, balanced truncation and metric complexity theory. Section VI provides concluding remarks.

II. PROBLEM FORMULATION

In this paper the class of stable linear infinite dimensional systems \( G \) with \( m \) inputs, \( p \) outputs described by the following,

\[
\begin{align*}
G : \dot{x}(t) &= Ax(t) + Bu(t) \\
y(t) &= Cx(t)
\end{align*}
\]

(1)

where \( A \) is the generator of a strongly continuous semigroup \( T(\cdot) \) (see for e.g. [17] for definition), \( B \) a linear bounded operator from \( \mathbb{R}^m \) into \( H \), \( C \) a linear bounded operator from \( H \) into \( \mathbb{R}^p \), \( H \) a Hilbert space representing the state space. The Hilbert space \( H \) will be taken to be the (universal) Hilbert space \( \ell^2 \) for concreteness.

In what follows we assume that the initial state \( x_0 = 0 \). We
have then (see for e.g. [11])
\[ x(t) = \int_0^t T(t - \tau)Bu(\tau)d\tau, \quad t \geq 0 \quad (2) \]
and consequently
\[ y(t) = \int_0^t CT(t - \tau)Bu(\tau)d\tau, \quad t \geq 0 \quad (3) \]
The corresponding impulse response \( h(t) \) is then defined by
\[ h(t) = CT(t)B, \quad t \geq 0. \]
We assume that the impulse response satisfies
\[ h(\cdot) \in L^1(\mathbb{R}^{p \times m}) \cap L^2(\mathbb{R}^{p \times m}) \quad (4) \]
The output is given by the convolution
\[ y(t) = \int_0^t h(t - s)u(s)ds \quad (5) \]
We would like to approximate the behavior of the infinite dimensional system (5) by finite dimensional models of order, say \( n \). To do so we consider the impulse response of the system (5), and formulate the problem as a distance minimization as follows
\[ \mu = \inf \left\{ \int_0^t \| h(t) - h_n(t) \|_2^2 dt \right\}^{\frac{1}{2}} \quad (6) \]
where \( h_n \) is an impulse response of order \( n < \infty \), and the norm \( \| \cdot \|_2 \) is defined as
\[ \| h(t) \|_2 = (\text{tr}(h^*h)(t))^{\frac{1}{2}} \quad (7) \]
where \( \text{tr}(\cdot) \) denotes the trace, and \( h^* \) denotes the adjoint of \( h \).

The corresponding Hankel operator to \( h(t) \), denoted \( \Gamma : L^2(\mathbb{R}^m) \to L^2(\mathbb{R}^p) \), is defined as
\[ \Gamma u(t) = \int_{-\infty}^0 h(t - \tau)u(\tau)d\tau \quad (8) \]
The change of variables \( \tau \mapsto -s \) transforms (44) into
\[ \Gamma u(t) = \int_0^\infty h(t + s)u(s)ds \quad (9) \]
Note that (9) is an integral operator with kernel \( h(t + s) \).

The assumption (4) guarantees that the operator \( \Gamma \) is a Hilbert-Schmidt operator defined on \( L^2(\mathbb{R}^m) \). Moreover, the Hilbert-Schmidt norm of \( \Gamma \), denoted \( \| \cdot \|_{HS} \), can be written as [16]
\[ \|\Gamma\|_{HS} = \left( \int_0^\infty \| h(t) \|_2^2 dt \right)^{\frac{1}{2}} \quad (10) \]
Let its spectral factorization be given by
\[ \Gamma = \sum_{i=1}^\infty \sigma_i \chi_i \otimes \zeta_i, \quad \chi_i \in L^2(\mathbb{R}^m), \quad \zeta_i \in L^2(\mathbb{R}^p) \quad (11) \]
where \( \sigma_i \) are the Hankel singular values of the system \( G \) ordered in decreasing order, i.e., \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq \sigma_n \geq \cdots \), \( \otimes \) denotes the tensor product (see e.g. [16] for the definition), and \( \{ \chi_i \}_{i=1}^n \) and \( \{ \zeta_i \}_{i=1}^n \) are orthonormal sets in \( L^2(\mathbb{R}^m) \) and \( L^2(\mathbb{R}^p) \), respectively. The Hilbert Schmidt norm of \( \Gamma \) is also given by
\[ \|\Gamma\|_{HS} = \left( \sum_{i=1}^\infty \sigma_i^2 \right)^{\frac{1}{2}} = \left( \int_0^\infty \text{tr}(h^*h)(s)ds \right)^{\frac{1}{2}} = \left( \sum_{i=0}^\infty \| \Gamma \phi_i \|_2^2 \right)^{\frac{1}{2}} \quad (12) \]
where \( \{ \phi_i \} \) is any orthonormal basis in \( L^2(\mathbb{R}^m) \).

Next, consider the optimal distance minimization
\[ \mu_n := \inf_{n<\infty} \| \Gamma - \Gamma_n \|_{HS} \quad (13) \]
where \( \Gamma_n \) is an operator acting from \( L^2(\mathbb{R}^m) \) into \( L^2(\mathbb{R}^p) \) of rank at most \( n \). The optimization (13) is the shortest distance from the operator \( \Gamma \) to the subspace of finite rank Hilbert-Schmidt operators acting from \( L^2(\mathbb{R}^m) \) into \( L^2(\mathbb{R}^p) \). The latter will be denoted by \( \mathcal{S}_{2n} \). Note that this distance problem is posed in an infinite-dimensional space. The latter is the space of Hilbert-Schmidt operators with the HS-norm which is Hilbert operator space. If the minimizing operator is itself a Hankel operator then the corresponding impulse response will solve the optimization (6).

A Theorem in [16] asserts that the subspace of finite rank Hilbert-Schmidt operators defined from \( L^2(\mathbb{R}^m) \) into \( L^2(\mathbb{R}^p) \), \( \mathcal{S}_{2n} \), can be represented in terms of tensors as
\[ \mathcal{S}_{2n} = \{ S = \sum_{j=1}^n \alpha_j v_j(\tau) \otimes \varphi_j(t), \alpha_j \geq 0, \] \[ v_j(\tau) \in L^2(\mathbb{R}^m), \varphi_j(t) \in L^2(\mathbb{R}^p) \} \quad (14) \]
where \( \{ v_j \} \) and \( \{ \varphi_j \} \) can be chosen to be orthonormal sets in \( L^2(\mathbb{R}^m) \) and \( L^2(\mathbb{R}^p) \), respectively.

The space of Hilbert-Schmidt operators is in fact a Hilbert space with the inner product [16], denoted \( \langle \cdot, \cdot \rangle \), if \( A \) and \( B \) are two Hilbert-Schmidt operators defined on \( L^2(\mathbb{R}^m) \),
\[ \langle A, B \rangle := \text{tr}(B^*A) \quad (15) \]
where \( \text{tr} \) denotes the trace, which in this case is given by the sum of the eigenvalues of the operator \( B^*A \) which is necessarily finite [16]. Note that the inner product (16) induces the Hilbert-Schmidt norm \( \| A \|_{HS} = (\text{tr}(A^*A))^{\frac{1}{2}} \). As it stands by the principle of orthogonality the shortest distance minimization is solved by the orthogonal projection of \( \Gamma \) onto the subspace \( \mathcal{S}_{2n} \). Let us call this orthogonal projection \( P_S \).

Using the polar representation of compact operators [16], \( \Gamma = U(\Gamma^*) \frac{1}{2} \), where \( U \) is a partial isometry and \( (\Gamma^*)^2 \) the square root of \( \Gamma \), which is also a Hilbert-Schmidt operator, and admits a spectral factorization of the form [16]
\[ (\Gamma^*\Gamma)^{\frac{1}{2}} = \sum_i \sigma_i \chi_i \otimes \chi_i \quad (16) \]
where $\sigma_i > 0$, $\sigma_i \downarrow 0$ as $i \uparrow \infty$, are the eigenvalues of $\langle \Gamma^* \Gamma \rangle^{\frac{1}{2}}$, and $\nu_i$ form the corresponding orthonormal sequence of eigenvectors, i.e., $\langle \Gamma^* \Gamma \rangle^{\frac{1}{2}} \chi_i = \sigma_i \chi_i$, $i = 1, 2, \ldots$. Note that $U \chi_i = \zeta_i$. Both $\{\chi_i\}$ and $\{\zeta_i\}$ are orthonormal sequences in $L^2(\mathbb{R}^m)$ and $L^2(\mathbb{R}^p)$, respectively. The sum (16) has either a finite or countably infinite number of terms. The above representation is unique.

Our objective now is to compute the orthogonal projection $P_S$. To do so, note that the subspace $S_{2n}$ can be re-written as

$$S_{2n} = \text{Span}\{v_j(\tau) \otimes \varphi_j(t), \; j = 1, 2, \ldots, n, \; v_j \in L^2(\mathbb{R}^m), \; \varphi_j \in L^2(\mathbb{R}^p)\}$$  \hspace{1cm} (17)

First, we show that the minimizing tensors $v_j(\tau) \otimes \varphi_j(t)$, $j = 1, 2, \ldots, n$, correspond to the Hankel operator tensors $\chi_i \otimes \zeta_i$, $i = 1, 2, \ldots, n$. To see this let $\{\phi_j\}$ be a basis in $L^2(\mathbb{R}^m)$, then we have

$$\|\Gamma\|_{HS}^2 = \sum_{i=1}^{\infty} \|\Gamma^i\|_2^2$$  \hspace{1cm} (18)

Let $\alpha_j$ be scalars, with $v_j = \phi_j$, $j = 1, \ldots, n$. The minimization (13) can be written as

$$\mu_n = \inf_{\alpha_j, v_j, \varphi_j} \{\|\Gamma - \sum_{j=1}^{n} \alpha_j v_j \otimes \varphi_j\|_2\}$$  \hspace{1cm} (19)

It follows from (18) that the RHS of (19) squared satisfies

$$\|\Gamma - \sum_{j=1}^{n} \alpha_j v_j \otimes \varphi_j\|_2^2 = \sum_{i=0}^{\infty} \|\Gamma - \sum_{j=1}^{n} \alpha_j v_j \otimes \varphi_j\phi_i\|_2^2$$

$$= \sum_{i=1}^{\infty} \|\Gamma\phi_i\|_2^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_j \langle \phi_i, \varphi_j \rangle^2$$

$$= \|\Gamma\|_{HS}^2 - \|\sum_{i=1}^{n} \Gamma\phi_i\|_2^2 + \|\sum_{j=1}^{n} \Gamma\varphi_j\|_2^2$$

Therefore, we have

$$\inf_{\alpha_j, v_j, \varphi_j} \{\|\Gamma - \sum_{j=1}^{n} \alpha_j v_j \otimes \varphi_j\|_2^2\} = \inf_{\alpha_j, v_j, \varphi_j} \|\Gamma\|_{HS}^2 - \|\sum_{i=1}^{n} \Gamma\phi_i\|_2^2 + \|\sum_{j=1}^{n} \Gamma\varphi_j\|_2^2\}$$  \hspace{1cm} (20)

The minimization can be accomplished by maximizing $\|\sum_{i=1}^{n} \Gamma\phi_i\|_2$ with respect to $v_i$, $\|v_i\|_2 = 1$, $i = 1, \ldots, n$, and minimizing $\|\sum_{j=1}^{n} \Gamma\varphi_j\|_2$, under $\|\varphi_j\|_2 = 1$. The maximization is solved by taking $v_j = \chi_j$, and the minimization by putting $\alpha_j = \sigma_j$, $\varphi_j = \zeta_j$, $j = 1, \cdots, n$, since

$$\sum_{j=1}^{n} \Gamma\chi_j = \sum_{i=1}^{\infty} \sum_{j=1}^{n} \sigma_i < \chi_j, \chi_i > \zeta_i = \sum_{j=1}^{n} \sigma_j \zeta_j$$

Hence, the optimal subspace $S_{2n}$ can be written as

$$S_{2n} = \text{Span}\{\chi_j(\tau) \otimes \zeta_j(t), \; j = 1, 2, \ldots, n\}$$  \hspace{1cm} (23)

The optimal rank $n$ operator, denoted $\Gamma_n$, in (13) is therefore given by

$$\Gamma_n = \sum_{i=1}^{n} \sigma_i \chi_i \otimes \zeta_i$$  \hspace{1cm} (24)

From (24) we see that $\Gamma_n$ is a Hilbert-Schmidt operator

$$\Gamma_n : L^2(\mathbb{R}^m) \rightarrow L^2(\mathbb{R}^p)$$

$$(\Gamma_n f)(t) = \int_{0}^{\infty} \sum_{i=1}^{n} \sigma_i \chi_i(\tau) f(\tau) \, d\tau \zeta_i(t)$$

The corresponding Hilbert-Schmidt kernel is

$$k(\tau, t) := \sum_{i=1}^{n} \sigma_i \chi_i(\tau) \otimes \zeta_i(t)$$

The optimal distance minimization (13) is then

$$\mu_n = \|\Gamma - \Gamma_n\|_{HS} = \left( \sum_{i=n+1}^{\infty} \sigma_i^2 \right)^{\frac{1}{2}}$$  \hspace{1cm} (27)

The optimal $\Gamma_n$ is a not necessarily a Hankel operator. In fact, $\Gamma_n$ is a Hankel operator if there exists an integral kernel $N$ such that [11]

$$k(\tau, t) = N(\tau + t), \quad \tau, \; t \geq 0$$  \hspace{1cm} (28)

In the next section, we show that a particular balanced realization realizes the optimal operator $\Gamma_n$ as a Hankel operator. That is, the corresponding kernel $k(\tau, t)$ of $\Gamma_n$ may be taken to satisfy (28). Moreover, truncating this balanced realization will be shown to correspond to balanced POD.

### III. A PARTICULAR BALANCED REALIZATION

In this section, we introduce a particularly important balanced realization proposed in [18], and which play a crucial role in the sequel. The importance of this realization stems from the fact that it uses the Hankel singular vectors and values in the state space representation.

Let $\tilde{A}, \tilde{B}$ and $\tilde{C}$ be linear bounded operators: $\tilde{A} : H \rightarrow H$, $\tilde{B} : \mathbb{R}^m \rightarrow H$, $\tilde{C} : H \rightarrow \mathbb{R}^p$, where $\tilde{A}$ is the infinitesimal generator of a strongly continuous semigroup, $\tilde{T}(t)$, on $H$ such that $h(t) = \tilde{C} \tilde{T}(t) \tilde{B}$, a.e. in $t$. That is $(\tilde{C}, \tilde{A}, \tilde{B})$ is a realization of $h(t)$. We will assume that $\tilde{T}(t)$ is exponentially stable.

Define the controllability and observability operators denoted $\Psi_c$ and $\Psi_o$, respectively by [11]

$$\Psi_c : L^2(\mathbb{R}^m) \rightarrow H$$

$$\Psi_c u := \int_{0}^{\infty} \tilde{T}(t) \tilde{B} u(\tau) d\tau$$

$$\Psi_o : H \rightarrow L^2(\mathbb{R}^p)$$

$$\Psi_o x_0 := \tilde{C} \tilde{T}(t) x_0, \quad t \geq 0$$  \hspace{1cm} (29)

The operators $\Psi_c$ and $\Psi_o$ realize the Hankel operator $\Gamma$ as

$$\Gamma = \Psi_o \Psi_c$$  \hspace{1cm} (30)
The realization $(\tilde{A}, \tilde{T}(t), \tilde{B})$ is balanced if the following controllability and observability gramians

$$\Psi_c \Psi_c^* = \int_0^\infty \tilde{T}(t) \tilde{B} \tilde{B}^* \tilde{T}(t)^* dt$$  \hspace{1cm} (31)

$$\Psi_o \Psi_o^* = \int_0^\infty \tilde{T}(t)^* \tilde{C}^* \tilde{C} \tilde{T}(t) dt$$  \hspace{1cm} (32)

are both equal to the same positive diagonal operator, i.e.,

$$\Psi_c \Psi_c^* = \Psi_o \Psi_o^* = \Sigma := \text{diag}\{\sigma_1, \sigma_2, \cdots\}$$  \hspace{1cm} (33)

Following [18] we use the following realization on $\ell^2$, for $i, j = 1, 2, \cdots$,

$$T_{ij}(t) = \left(\frac{\sigma_j}{\sigma_i}\right)^{\frac{1}{2}} \int_0^\infty \zeta_i^*(\tau) \zeta_j(t+\tau) d\tau$$  \hspace{1cm} (34)

$$\tilde{B} = (\sqrt{\sigma_1 \chi_1(0)}, \cdots, \sqrt{\sigma_i \chi_i(0)}, \cdots)^T$$  \hspace{1cm} (35)

$$\tilde{C} = (\sqrt{\sigma_1 \zeta_1(0)}, \cdots, \sqrt{\sigma_i \zeta_i(0)}, \cdots)$$  \hspace{1cm} (36)

Note that $\tilde{T}(t) = (\tilde{T}_{ij}(t))$ is a strongly continuous contraction semigroup on $\ell^2$. Moreover, the corresponding infinitesimal generator $\tilde{A} = (\tilde{A}_{ij})$ satisfies [18]

$$\tilde{A}_{ij} = \left(\frac{\sigma_j}{\sigma_i}\right)^{\frac{1}{2}} \int_0^\infty \zeta_i^*(\tau) \frac{d\zeta_j(\tau)}{d\tau} d\tau$$  \hspace{1cm} (37)

This realization produces bounded controllability and observability operators which satisfy [18]

$$\Psi_c \Psi_c^* = \Psi_o \Psi_o^* = \Sigma := \text{diag}\{\sigma_1, \sigma_2, \cdots\}$$  \hspace{1cm} (38)

That is the realization $(\tilde{C}, \tilde{A}, \tilde{B})$ is a balanced realization. In the next section, we show that truncating this balanced realization results in balanced POD.

IV. TRUNCATION OF THE BALANCED REALIZATION YIELDS BALANCED POD

Following [18] and as in the finite dimensional case, we define the $n$-th order truncations of $(\tilde{C}, \tilde{A}, \tilde{B})$ as

$$\tilde{A}_r := (\tilde{A}_r)_{i,j} := \left(\frac{\sigma_j}{\sigma_i}\right)^{\frac{1}{2}} \int_0^\infty \zeta_i^*(\tau) \frac{d\zeta_j(\tau)}{d\tau} d\tau$$

$$i, j = 1, 2, \cdots, n$$

$$\tilde{B}_r := (\sqrt{\sigma_1 \chi_1(0)}, \cdots, \sqrt{\sigma_n \chi_n(0)})^T$$

$$\tilde{C}_r := (\sqrt{\sigma_1 \zeta_1(0)}, \cdots, \sqrt{\sigma_n \zeta_n(0)})$$  \hspace{1cm} (39)

The truncated system $(\tilde{A}_r, \tilde{B}_r, \tilde{C}_r)$ has state space $\mathbb{R}^n$, is stable and balanced [18], i.e., the corresponding controllability and observability gramians satisfy

$$\Psi_r \Psi_r^* = \Psi_o \Psi_o^* = \Sigma := \text{diag}\{\sigma_1, \sigma_2, \cdots, \sigma_n\}$$  \hspace{1cm} (40)

Therefore, the reduced model (39) corresponds exactly to balanced truncation. Now its impulse response, denoted $h_r(t)$, is given by $h_r(t) = \tilde{C}_r e^{\tilde{A}_r t} \tilde{B}_r$, a.e. $t$.

In terms of the singular vectors $\chi_i$ and $\zeta_i$, $i = 1, 2, \cdots, n$, the semigroup $e^{\tilde{A}_r t}$ can be written as

$$e^{\tilde{A}_r t} = \left(\frac{\sigma_j}{\sigma_i}\right)^{\frac{1}{2}} \int_0^\infty \zeta_i^*(\tau) \zeta_j(t+\tau) d\tau$$  \hspace{1cm} (41)

The Hankel operator, denoted $\Gamma_r$, associated to the truncated balanced system, $(\tilde{A}_r, \tilde{B}_r, \tilde{C}_r)$, is given by

$$\Gamma_r : L^2(\mathbb{R}^m) \longrightarrow L^2(\mathbb{R}^p)$$

$$(\Gamma_r u)(t) = \int_0^\infty h_r(t+\tau) u(\tau) d\tau$$  \hspace{1cm} (42)

The corresponding spectral decomposition is clearly

$$\Gamma_r = \sum_{i=1}^n \sigma_i \chi_i \otimes \zeta_i$$  \hspace{1cm} (43)

and therefore, $\Gamma_r = \Gamma_r^o$ and $\Gamma_r$ solves the optimal distance minimization (13). In addition, since $\Gamma_r$ is Hankel its impulse response solves the optimization (6), i.e.,

$$\mu = \inf_{h_n} \left(\int_0^\infty \|h(t) - h_n(t)\|_2^2 dt\right)^{\frac{1}{2}}$$

$$= \left(\int_0^\infty \|h(t) - h_r(t)\|_2^2 dt\right)^{\frac{1}{2}}$$  \hspace{1cm} (44)

and both optimizations (6) and (13) are equal, i.e., $\mu = \mu_n$.

The above derivation shows that balanced POD and balanced truncation are both the result of the optimal approximation (13), and in fact coincide when using the realization (36) proposed in [18].

In the next section, balanced POD and balanced truncation are related to key notions in metric complexity theory.

V. BALANCED POD, BALANCED TRUNCATION AND METRIC COMPLEXITY THEORY

In this section, we discuss role of balanced POD and truncation in optimizing different $n$-widths defined in [13]. The main object here is the Hankel operator $\Gamma$ associated to the system. Recall that $\Gamma$ is defined as

$$\Gamma : L^2(\mathbb{R}^m) \longrightarrow L^2(\mathbb{R}^p)$$

$$(\Gamma u)(t) = \int_0^\infty h(t+\tau) u(\tau) d\tau$$  \hspace{1cm} (45)

Under the assumption on the impulse response $h(t)$ stated above the operator $\Gamma$ is Hilbert-Schmidt.

Since in model reduction we are interested in approximation by finite dimensional models, and in particular, $n$-parameter affine models such as in balanced POD, we introduce the representation error of $\Gamma$ as the best (optimal) set distance between the range space of $\Gamma$, $(\Gamma L^2(\mathbb{R}^m))$ and the finite dimensional subspaces, $X_n$, of $L^2(\mathbb{R}^p)$, i.e.,

$$\text{dist}((\Gamma L^2(\mathbb{R}^m)), X_n) := \sup_{f \in (\Gamma L^2(\mathbb{R}^m))} \inf_{g \in X_n} \|f - g\|_2$$  \hspace{1cm} (46)

where the infimum is taken over the $n$-dimensional subspaces $X_n$ of $L^2(\mathbb{R}^p)$.

By definition of the representation error (46), the minimum
representation of the range space \((\Gamma L^2(\mathbb{R}^m))\) by an \(n\)-dimensional subspace \(X_n\) is

\[
d_n \left( \left( \Gamma L^2(\mathbb{R}^m) \right), L^2(\mathbb{R}^p) \right) = \inf_{X_n \subset L^2(\mathbb{R}^p)} \text{dist} \left( \left( \Gamma L^2(\mathbb{R}^m) \right), X_n \right)
\]

(47)

This is exactly the Kolmogorov \(n\)-width of \((\Gamma L^2(\mathbb{R}^m))\). If the infimum in (47) is attained for some subspace \(X_n^*\), then \(X_n^*\) is said to be an optimal subspace for \(d_n \left( \left( \Gamma L^2(\mathbb{R}^m) \right), L^2(\mathbb{R}^p) \right)\). The optimal subspace gives the optimal \(n\)-dimensional affine model for \((\Gamma L^2(\mathbb{R}^m))\).

The Kolmogorov \(n\)-width characterizes the representation complexity of the model reduction problem. The inverse function of \(d_n(\cdot)\) was called metric dimension function by Zames [19], and viewed as an appropriate measure of the metric complexity of uncertain systems. It is the dimension of the smallest subspace whose elements can approximate arbitrary points of \((\Gamma L^2(\mathbb{R}^m))\) to a specified tolerance.

The \(n\)-width in the sense of Gel’fand, is defined as [13]

\[
d^n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right) := \inf_{L^n(\mathbb{R}^p)} \sup_{\|f\|_2 \leq 1} \|\Gamma f\|_2
\]

(48)

where the infimum is taken over all subspaces \(L^n\) of \(\Gamma(L^2(\mathbb{R}^m))\) of codimension at most \(n\). A subspace is said to be of co-dimension \(n\) if there exist \(n\) independent bounded linear functionals \(\{f_i\}_{i=1}^n\) on \(L^2(\mathbb{R}^p)\) for which

\[
L^n = \{g : g \in L^2(\mathbb{R}^p), f_i(g) = 0, i = 1, 2, \ldots, n\}
\]

(49)

If

\[
d^n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right) = \sup \{\|f\|_2 : f \in \Gamma(L^2(\mathbb{R}^m)) \cap L^n\}
\]

(50)

where \(L^n\) is a subspace of codimension at most \(n\), then \(L^n\) is an optimal subspace for

\[
d^n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right)
\]

(51)

The Gel’fand \(n\)-width characterizes the experimental complexity of the information collecting stage using simulation or identification. It is related to the inherent error due to lack of data and inaccurate measurements. The inverse of the Gel’fand \(n\)-width gives the least number of measurements needed to reduce the modeling uncertainty to a predetermined value.

The linear \(n\)-width is defined by

\[
\delta_n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right) := \inf_{P_n} \sup_{\|\phi\|_2 \leq 1, \phi \in L^2(\mathbb{R}^m)} \|\Gamma \phi - P_n \phi\|_2
\]

(52)

where \(P_n\) is any continuous linear operator from \(L^2(\mathbb{R}^m)\) into \(L^2(\mathbb{R}^p)\) of rank at most \(n\).

The \(n\)-width in the sense of Bernstein of \(\Gamma(L^2(\mathbb{R}^m))\) is defined by [13]

\[
b_n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right) = \sup \sup \{\lambda : \lambda S(X_{n+1}) \subseteq L^2(\mathbb{R}^p)\} = \sup \inf_{X_{n+1} \in \partial(\Gamma(L^2(\mathbb{R}^m)) \cap X_{n+1})} \|x\|_2
\]

(53)

where \(X_{n+1}\) is any \((n+1)\)-dimensional subspace of \(L^2(\mathbb{R}^p)\), and \(S(X_{n+1})\) is the unit ball in \(X_{n+1}\), i.e.,

\[
S(X_{n+1}) = \{y \in X_{n+1} : \|y\|_2 \leq 1\}
\]

(54)

The basic results of this section are summarized in the following theorem which tells us that balanced POD and balanced truncation optimize the different \(n\)-widths.

**Theorem 1:** Let the operator \(\Gamma\) and \(\{\sigma_i\}, \{\chi_i\}, \{\zeta_i\}\) be defined as above. Then

\[
d_n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right) = d^n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right) = \delta_n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right) = b_n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right)
\]

(55)

\[
\sigma_{n+1}, n = 0, 1, 2, \ldots
\]

Furthermore, the singular vectors \(\{\chi_i\}\) and \(\{\zeta_i\}\) are optimal for the \(n\)-widths in the following sense

i) The subspace spanned by the vectors \(\{\zeta_i\}\), \(X_n = \text{Span}\{\zeta_1, \ldots, \zeta_n\}\), is optimal for

\[
d_n \left( \Gamma(L^2(\mathbb{R}^p)); L^2(\mathbb{R}^p) \right)
\]

(56)

ii) The subspace

\[
L^n = \left\{\chi \in L^2(\mathbb{R}^m), \int_0^\infty \chi(\tau) \chi_i(\tau) d\tau = 0, i = 1, 2, \ldots, n\right\}
\]

(57)

is optimal for \(d^n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right)\).

iii) The linear operator

\[
Q_n \phi = \sum_{i=1}^n \int_0^\infty \phi(\tau) \chi_i(\tau) d\tau \chi_i
\]

(59)

is optimal for \(\delta_n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right)\).

iv) The subspace

\[
X_{n+1} = \text{Span}\{\zeta_1, \ldots, \zeta_{n+1}\}
\]

(60)

is optimal for \(b_n \left( \Gamma(L^2(\mathbb{R}^m)); L^2(\mathbb{R}^p) \right)\).
Sketch of the Proof. From Chapter II \cite{13}, we have
\[
\delta_n \left( \Gamma \left( L^2(\mathbb{R}^m) \right); L^2(\mathbb{R}^p) \right) \geq d_n \left( \Gamma \left( L^2(\mathbb{R}^m) \right); L^2(\mathbb{R}^p) \right) = \delta_n \left( \Gamma \left( L^2(\mathbb{R}^m) \right); L^2(\mathbb{R}^p) \right) \geq b_n \left( \Gamma \left( L^2(\mathbb{R}^m) \right); L^2(\mathbb{R}^p) \right)
\]
(61)

Note that from (53) we have
\[
b_n \left( \Gamma \left( L^2(\mathbb{R}^m) \right); L^2(\mathbb{R}^p) \right) = \sup_{X_{n+1} \in \Gamma \left( L^2(\mathbb{R}^m) \right) \cap X_{n+1}} \inf_{\|x\|_2} \|x\|_2
\]
(62)

Let \( \{\psi_1, \ldots, \psi_{n+1}\} \) be a basis for \( X_{n+1} \), and since for \( x \in \Gamma \left( L^2(\mathbb{R}^m) \right) \cap X_{n+1}, \ x = \Gamma y, \ \exists y \in L^2(\mathbb{R}^m) \) with \( \Gamma y \in X_{n+1} = \text{Span}\{\psi_1, \ldots, \psi_{n+1}\} \), and therefore for some \( \varphi \in X_{n+1} \), we have
\[
\|x\|_2 = \|\Gamma y\|_2 = \frac{\langle (\Gamma^*)^{\frac{1}{2}} \varphi, \varphi \rangle}{\langle \varphi, \varphi \rangle} \leq \sigma_{n+1}
\]
(63)

where \( \langle \cdot, \cdot \rangle \) represents the inner product in \( L^2(\mathbb{R}^p) \), and it follows that
\[
\inf_{x \in \partial(\Gamma \left( L^2(\mathbb{R}^m) \right) \cap X_{n+1})} \|x\|_2 \|x\|_2 = \frac{\langle (\Gamma^*)^{\frac{1}{2}} \varphi, \varphi \rangle}{\langle \varphi, \varphi \rangle} = \sigma_{n+1}
\]
(64)

and thus
\[
b_n \left( \Gamma \left( L^2(\mathbb{R}^m) \right); L^2(\mathbb{R}^p) \right) = \sup_{\psi_1, \ldots, \psi_{n+1}, \varphi \in \text{Span}\{\psi_1, \ldots, \psi_{n+1}\}, \varphi \neq 0} \frac{\langle (\Gamma^*)^{\frac{1}{2}} \varphi, \varphi \rangle}{\langle \varphi, \varphi \rangle} = \sigma_{n+1}
\]
(65)

where the supremum is taken w.r.t. all sets of \( (n+1) \) linearly dependent vectors \( \{\psi_i\} \) in \( L^2(\mathbb{R}^p) \). Moreover, the supremum is attained for \( \psi_i = \zeta_i, \ i = 1, \ldots, n+1, \) and \( \varphi = \zeta_{n+1} \).

The linear \( n \)-width satisfies \cite{13}
\[
\delta_n \left( \Gamma \left( L^2(\mathbb{R}^m) \right); L^2(\mathbb{R}^p) \right) \leq \sigma_{n+1}
\]
(66)

Thus all the \( n \)-widths equal \( \sigma_{n+1}, \) and i), iii) and iv) hold.

VI. CONCLUSION

In this paper, using tools from operator and system theory we showed that balanced POD and balanced truncations coincide when using a particular balanced realization for infinite dimensional systems. In this regard, Hilbert-Schmidt Hankel operator associated to the system when it exists plays a crucial role. Connections to different \( n \)-widths which are important notions in metric complexity theory are explored, and show that balanced POD and truncations optimize the \( n \)-widths.