

A Unified View of Balanced Truncation and Singular Perturbation Approximations

Philip E. Paré, Alma T. Wilson, Mark K. Transtrum and Sean C. Warnick

Abstract—This paper demonstrates that both Balanced Truncation and Balanced Singular Perturbation Approximations can be viewed as limiting approximations of the same parameterization of Linear Time Invariant (LTI) systems. First, we introduce a specific parameterization of LTI systems that distinguishes dynamic and structural parameters. Next, we apply the Model Boundary Approximation Method (MBAM) [1] to particular parameters to achieve different approximations. This unified view of these popular model reduction techniques, which themselves can result in quite different approximations, illustrates that each approximation corresponds to a particular boundary element on a manifold, the “model manifold,” which is associated with the specific choice of model parameterization and is embedded in a sample space of measured outputs.

I. INTRODUCTION: MOTIVATION

Our modern world has focused its attention on systems of unprecedented size and complexity [2]. Typical examples include the Internet [3], biological systems [4], and economic networks [5]. Analyzing and designing such systems has made the need for simplified models all the more urgent, since detailed descriptions of such systems are unwieldy and defy comprehension. We focus on two common motivations for building simplified approximations of complex systems.

First, simplified models are necessary when attempting to learn a system from limited data. First principles models typically have many parameters that must be tuned correctly for the model to reflect the behavior of a real system. Using data to learn the correct values of parameters is the purview of system identification, and a rich theory has developed quantifying when data is informative enough to accurately estimate parameter values [6]. Typically we have much less data than needed to learn all the parameters in a first-principles model, so simplifying the model to yield one with fewer parameters can help identify the system from data.

Another reason for simplified models arises when designing feedback controllers for complex systems. Often the complexity of an optimal controller may mirror that of the system to be controlled, so controlling complex systems may often suggest the need for complicated controllers. Nevertheless, when engineering such complicated systems is unreasonable,

Philip Paré is at the University of Illinois at Urbana-Champaign and can be reached at philip.e.pare@gmail.com. Alma T. Wilson, Mark K. Transtrum and Sean C. Warnick are at Brigham Young University, Provo, Utah, and can be reached at alba_wilzod@yahoo.com, mkttranstrum@byu.edu, and sean.warnick@gmail.com, respectively. This material is based on research sponsored by the Department of Homeland Security (DHS) Science and Technology Directorate, Homeland Security Advanced Research Projects Agency (HSARPA), Cyber Security Division (DHS S&T/HSARPA/CSD), LRBA 12-07 via contract number HSHQDC-13-C-B0052. All material in this paper represents the position of the authors and not necessarily that of DHS.

designing controllers for simplified approximations can lead to acceptable trade-offs between complexity savings and performance degradation.

In this paper we explore the underlying structure of system approximation and show that two common model reduction methods for LTI systems actually arise from a similar limiting approximation applied to a common parametrization. In Section II we review two prominent model reduction techniques; Balanced Truncation (BT) and Balanced Singular Perturbation Approximation (BSPA). We present MBAM in Section III. In Section IV, we present a parameterization of LTI systems and then use it to derive BT and BSPA using MBAM. In Section V we present a simple example that illustrates the result and how MBAM can give insight into model reduction of linear systems.

II. BACKGROUND

For this work we use LTI systems,

$$\begin{aligned}\dot{\hat{x}}(t) &= \hat{A}\hat{x} + \hat{B}u \\ y(t) &= \hat{C}\hat{x} + \hat{D}u,\end{aligned}\tag{1}$$

where $\hat{x}(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^p$, and $u(t) \in \mathbb{R}^m$. These systems may not capture all dynamic behavior but it may be applicable to any system near equilibria by virtue of Lyapunovs Indirect Method [7]. We will assume the system in Equation (1) is minimal and stable.

A. Model Reduction

Model reduction is an important, well-studied problem in controls. It can be argued that there are two main types of model reduction for LTI systems: BT and BSPA. We will review these two methods and their accompanying literature.

1) *Balanced Truncation*: Balanced Truncation was first proposed in [8] and has been well developed since then, with a clear presentation in [9]. Consider a minimal and stable system as in Equation (1). There exists a state transformation from these matrices to an input–output equivalent balanced realization

$$\begin{aligned}\dot{x}(t) &= Ax + Bu \\ y(t) &= Cx + Du,\end{aligned}\tag{2}$$

where A satisfies the Lyapunov equations:

$$A^T X + X A = -C^T C \quad \text{and} \quad A X + X A^T = -B B^T,\tag{3}$$

where $X = \text{diag}(\theta_1, \dots, \theta_n)$; the θ_i 's are the Hankel singular values (HSVs) [9]. This makes each state equally observable and controllable; they are ordered from strongest

to weakest. Once balanced, we partition the states giving

$$A = \begin{bmatrix} A_{11} & A_{1k} \\ A_{k1} & A_{kk} \end{bmatrix}, B = \begin{bmatrix} B_1 \\ B_k \end{bmatrix}, C = [C_1 \ C_k], \quad (4)$$

where $A_{11} \in \mathbb{R}^{(n-k) \times (n-k)}$, $A_{kk} \in \mathbb{R}^{k \times k}$, and the rest of the blocks are the appropriate dimensions. Then BT of k states gives the system

$$\begin{aligned} \dot{x}(t) &= A_{11}x + B_1u \\ y(t) &= C_1x + Du. \end{aligned} \quad (5)$$

This approximation has well defined error bounds [9]–[11]. This method has been extended to several different classes of systems including time-varying, multidimensional, and uncertain systems [12]–[14]. Many structure preserving model reduction techniques, with varying definitions, have been considered as extensions of BT [15]–[18]. Extensions have also been developed for nonlinear systems [19]–[21].

2) *Singular Perturbation Approximation*: Perturbation theory is a well studied area and has a rich background in linear operator theory and the controls field [22], [23]. It is commonly applied in the context of well-separated time scales. In this case the ratio of time-scales identifies an explicit “small” parameter in which a series expansion can be computed. The theory has also been applied to balanced realizations, which we will refer to as Balanced Singular Perturbation Approximation [24]–[26]. Given a partitioned balanced realization as in Equation (4), the reduced system becomes

$$\begin{aligned} \tilde{A} &= A_{11} - A_{1k}A_{kk}^{-1}A_{k1}, \quad \tilde{B} = B_1 - A_{1k}A_{kk}^{-1}B_k, \\ \tilde{C} &= C_1 - C_kA_{kk}^{-1}A_{k1}, \quad \tilde{D} = D - C_kA_{kk}^{-1}B_k. \end{aligned} \quad (6)$$

The matrix \tilde{A} is the Schur complement, which we will denote by A/A_{kk} .

In many respects, BT and BSPA are complementary types of approximations. They are both derived from the same block partition of a balanced realization. They share the same error bounds in terms of the HSVs. It is well known that BT typically provides better approximations at high frequencies while BSPA works well at low frequencies [26]. In what follows, we will see how these similarities allow us to unify both approximations as limiting approximations of a balanced system.

III. MANIFOLD BOUNDARY APPROXIMATION METHOD

The Manifold Boundary Approximation Method (MBAM) was originally described in [1]. Here we present a formalization of MBAM. The basis for this approach is the observation that a parameterized model is a mapping between a parameter space and a prediction space, data space, or a sample space of measured outputs. As such, it is natural to interpret a model as a manifold embedded in the space of possible predictions. We refer to this manifold as the *model manifold*, denoted by $\mathcal{M}(D)$.

Definition 1: The model mapping is defined as $\mathcal{M} : D \subset \mathbb{R}^N \rightarrow \mathbb{R}^M$, where \mathbb{R}^N is the parameter space, \mathbb{R}^M is the data space or prediction space, and $N < M$.

Note that in Definition 1, we have assumed that $N < M$, with no bound on M . It is possible for the prediction space to be infinite dimensional (e.g. a function space).

For a given model manifold $\mathcal{M}(D)$, we denote the closure of the manifold as $\overline{\mathcal{M}(D)}$.

Definition 2: Given a model mapping $\mathcal{M} : D \subset \mathbb{R}^N \rightarrow \mathbb{R}^M$, a point $x \in \overline{\mathcal{M}(D)}$ is interior if there exists an open neighborhood of $\mathcal{M}(D)$ centered at x . If a point is not interior then it is a boundary point.

The set of boundary points of the manifold closure defines the boundary of the manifold, denoted by $\partial\mathcal{M}(D)$.

Definition 3: Given a model mapping $\mathcal{M} : D \subset \mathbb{R}^N \rightarrow \mathbb{R}^M$, a model mapping $\tilde{\mathcal{M}}$ is a k^{th} order manifold boundary approximation of \mathcal{M} if:

- 1) $\tilde{\mathcal{M}} : \tilde{D} \subset \mathbb{R}^{N-k} \rightarrow \mathbb{R}^M$ and
- 2) $\tilde{\mathcal{M}}(\tilde{D}) \subset \partial\mathcal{M}(D)$.

That is to say that a k^{th} order MBAM approximation is defined as being on the boundary of the manifold and having k less parameters than the original system.

Note that up to this point, we have not employed any notion of distance on the model manifold. In practice, it is necessary to define a metric on the data space in order to quantify the error of the approximation, and MBAM can accommodate any choice of metric on the data space [27]. The chosen metric will determine which approximation is optimal and the associated error bound.

In the original presentation of MBAM [1], an information distance was imposed on the data space defined by the Fisher Information Matrix (FIM). Assuming the prediction space corresponds to a probability distribution, the FIM defines a Riemannian metric on the data space (a Riemannian metric is the metric defined by the inner product on the tangent space to a curved manifold [28]). From this Riemannian metric, computational differential geometry can be used to identify candidate boundary approximations [1]. Nonetheless, MBAM is a topological operation and is agnostic to the actual choice of metric [27].

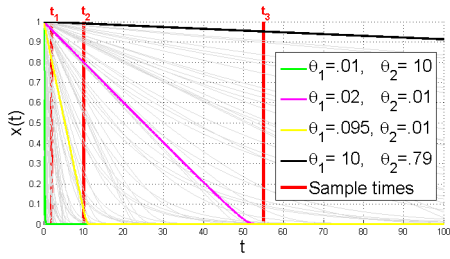
A. Simple MBAM Example

The MBAM concept is best illustrated by example. Consider the dynamics of the Michaelis Menten Reaction (MMR),

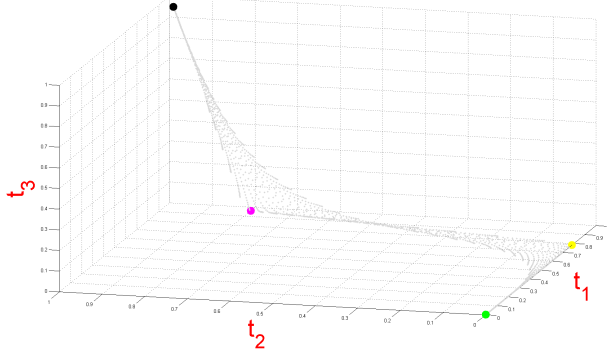
$$\dot{x} = \frac{-\theta_1 x}{\theta_2 + x}, \quad (7)$$

where θ_i indicates a parameter in the dynamic equations characterizing the family of models under consideration [29], [30]. The parameters θ_i are only physically relevant for positive values. We therefore restrict our attention to the domain D as the positive quadrant of the parameter space, that is where $\theta_i \geq 0 \forall i$. Note Equation (7) is stable for all parameter values in D . Assuming a fixed initial condition, $x_0 = 1$, and varying θ_1 and θ_2 among all possible elements of the domain, D , we observe time series for this model as shown in Figure 1a.

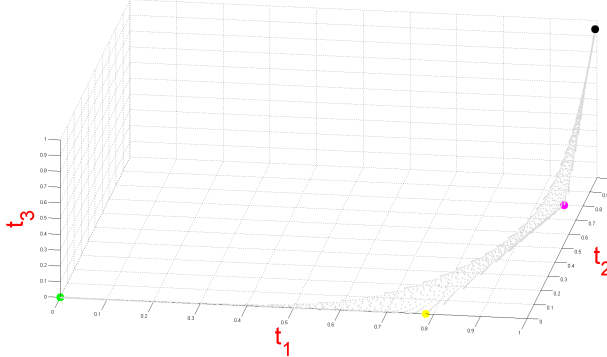
Consider now three observations of this system at time points, (t_1, t_2, t_3) , indicated by the red vertical lines in Figure 1a. The space of all possible experimental data at these



(a) Potential time series for the MMR model. Different curves are calculated by varying the parameters θ_1 and θ_2 .



(b) First view of the Model Manifold for the MMR model. Each point in the data space corresponds to a curve in Fig. 1a.



(c) Second view of the Model Manifold for the MMR model.

Fig. 1: MBAM works by sampling outputs at particular times for a variety of parameter values (Fig. 1a) and analyzing the resulting manifold in “Data space” (Fig. 1b-1c).

three time points forms a three-dimensional “data space.” All possible model predictions, for different values of θ_1 and θ_2 , correspond to a two-dimensional subset of this space. This two-dimensional surface is the model manifold from Definition 1. The manifold for the MMR model illustrated in Figures 1b-1c is bounded, a feature shown to be common among model manifolds with many parameters [31], [32]. The existence of these boundaries is the crucial element that enables the manifold boundary approximation, where each boundary corresponds to a different model reduction of the original model in Equation (7). The smoother looking boundary corresponds to the limit where $\theta_1, \theta_2 \rightarrow \infty$ and the ratio θ_1/θ_2 becomes the new parameter. The other boundary corresponds to the limit where $\theta_2 \rightarrow 0$, leaving θ_1 as the sole parameter. These two limiting approximations lead to

reduced models of the form:

$$\begin{aligned} \dot{x} &= -\frac{\theta_1}{\theta_2}x, \\ \dot{x} &= -\theta_1, \end{aligned} \quad (8)$$

respectively. A well-known interpretation of the MMR in biochemistry is that it describes a saturable reaction. That is, at low concentrations (small values of x), the reaction rate is nearly first order (i.e., linear in x), while at high concentrations the rate saturates into a zeroth order (i.e., constant with respect to x). Notice how these two limits are naturally identified as the boundaries of the model manifold in Figures 1b-1c.

IV. MAIN RESULT

We start with the presentation of a unified parameterization for LTI systems. Consider a minimal, stable system as in Equation (1). It is well known that there exists a state transformation from these matrices to an input–output equivalent balanced realization (A, B, C, D) as in Equation (2), where X is the diagonal matrix of HSVs satisfying Equation (3). The HSVs give a measure of the energy of the dynamics of the states; therefore they offer themselves as natural parameters for the system. The state transformation also becomes part of the parameterization because it maps to the balanced realization from the physical system or network structure [33]. Notice the following simple statement is true for balanced realizations.

Lemma 1: Given a balanced realization (A, B, C, D) , $\text{diag}(BB^T) = \text{diag}(C^T C)$.

Proof: Since (A, B, C, D) is balanced, the observability and controllability grammians are equal with the HSVs, $\theta_1, \dots, \theta_n$, on the diagonal. By inspecting the diagonals of the Lyapunov equations we see

$$\begin{aligned} a_{ii}\theta_i + \theta_i a_{ii} &= -(C^T C)_{ii} \\ a_{ii}\theta_i + \theta_i a_{ii} &= -(BB^T)_{ii}, \end{aligned} \quad (9)$$

where the subscript ii indicates the i th diagonal entry of the matrix and a_{ii} is the i th diagonal entry of the A matrix. Therefore $\text{diag}(BB^T) = \text{diag}(C^T C)$. ■

We introduce n parameters to denote the common diagonal entries and denote them by r_1^2, \dots, r_n^2 . We can then write the B matrix as

$$B = \begin{bmatrix} r_1 \beta_1^T \\ \vdots \\ r_n \beta_n^T \end{bmatrix} \quad (10)$$

where the β_i 's are a collection of normalized column vectors in \mathbb{R}^m satisfying $\beta_i^T \beta_i = 1$ for all $i = 1, \dots, n$. Denoting the $n \times m$ matrix whose rows correspond to β_i^T as β^T and introducing $R = \text{diag}(r_1, \dots, r_n)$, it follows that $B = R\beta^T$. Clearly by construction $\text{diag}(BB^T) = (r_1^2, \dots, r_n^2)$.

Similarly, we can write C as

$$C = [r_1 \gamma_1 \quad \dots \quad r_n \gamma_n] \quad (11)$$

where the γ_i 's are a collection of normalized column vectors in \mathbb{R}^p satisfying $\gamma_i^T \gamma_i = 1$ for all $i = 1, \dots, n$. We write $C = \gamma R$; note that $\text{diag}(C^T C) = (r_1^2, \dots, r_n^2) = \text{diag}(BB^T)$, which is consistent with Lemma 1.

Note that plugging r_i^2 into Equation (9) gives $a_{ii} = -\frac{r_i^2}{2\theta_i}$, for the diagonal elements of the A matrix. From the off-diagonals of the Lyapunov equations we find,

$$a_{ij} = r_i r_j \alpha_{ij}, \quad (12)$$

for $i \neq j$, where

$$\alpha_{ij} = \frac{\theta_j(\beta^T \beta)_{ij} - \theta_i(\gamma^T \gamma)_{ij}}{\theta_i^2 - \theta_j^2}. \quad (13)$$

The β_i 's and γ_i 's offer themselves as natural parameters of the system. Note that because of the normalization, each β_i can be uniquely identified by $m - 1$ independent parameters (for example, using the generalized spherical coordinates), so that the matrix β can be specified by $n(m - 1)$ independent parameters. Similarly, the matrix γ can be uniquely specified by $n(p - 1)$ independent parameters. Combining these numbers with the $2n$ parameters $\theta_1, \dots, \theta_n$ and r_1, \dots, r_n implies that an (A, B, C) system in a balanced realization can be uniquely identified with $n(p + m)$ parameters. The D matrix (which is independent of the other three) adds an additional pm parameters. Combining these with the parameters corresponding to the $n \times n$ state transformation T gives that any system $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ can be parameterized by the $n(p + m) + pm + n^2$ parameters corresponding to the HSVs, the r_i 's, the β_i 's, the γ_i 's, and the elements of T . Note that this is the same number of parameters as there are elements of the matrices $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$.

We call the elements of the state transformation the *structural parameters* because they reveal the physical system or network structure from the balanced realization and because they do not change the input–output, or dynamic, behavior of the system. We call the rest of the parameters *dynamic parameters*, because if changed (unless a systemic change of sign in the columns of β and γ), the input–output behavior of the system changes. Because the model manifold is defined only in terms of physically observable quantities, i.e. the outputs, different points on the model manifold are uniquely labeled by the dynamic parameters. In contrast, the structural parameters are non-identifiable from any physical observation of the system and some a priori knowledge of structure is required [27], [33]. In all cases the dimensionality of the model manifold will be given by the number of dynamic parameters that are allowed to vary in the system.

A. Balanced Truncation from MBAM

For the first theorem of this section we will restrict ourselves to considering the HSVs as the parameters, holding R , β , γ and D fixed.

Theorem 1: Balanced Truncation of k states on an n^{th} -order realization is equivalent to k MBAM approximations for the parameterization in which the Hankel singular values are parameters and the matrices R , β , γ and D are fixed.

Proof: Consider the equation for $\dot{x}_n(t)$,

$$\dot{x}_n(t) = \sum_{i=1}^{n-1} r_i r_n \alpha_{ni} x_i(t) - \frac{r_n^2 x_n(t)}{2\theta_n} + r_n \sum_{i=1}^m \beta_{ni} u_i(t). \quad (14)$$

Multiplying through by θ_n gives

$$\theta_n \dot{x}_n(t) = \theta_n \sum_{i=1}^{n-1} r_i r_n \alpha_{ni} x_i(t) - \frac{r_n^2 x_n(t)}{2} + \theta_n r_n \sum_{i=1}^m \beta_{ni} u_i(t). \quad (15)$$

Performing an MBAM approximation by taking the limit $\theta_n \rightarrow 0$ gives that $x_n = 0$. Plugging this back into the rest of the system gives BT of one state. Iterating this limit k times, always choosing the smallest HSV, completes the proof. ■

Theorem 2: Balanced Truncation of k states on an n^{th} -order realization is equivalent to k MBAM approximations for the parameterization in which the elements of R (the square root of the diagonal elements of BB^T and $C^T C$) are parameters and the Hankel singular values, β , γ , and D are fixed.

Proof: Consider the equation for $\dot{x}_n(t)$. Taking the limit $r_n \rightarrow 0$ gives that $\dot{x}_n = 0$, which effectively truncates the n^{th} state. This gives BT of one state. Iterating this limit k times, always choosing the r_i with the largest subscript, completes the proof. ■

B. Singular Perturbation Approximation from MBAM

Theorem 3: Balanced Singular Perturbation Approximation of k states on an n^{th} -order realization is equivalent to k MBAM approximations for the parameterization in which the elements of R (the square root of the diagonal elements of BB^T and $C^T C$) are parameters and the Hankel singular values, β , γ , and D are fixed.

Proof: We will prove this by induction, starting with the one state case. Consider the $\dot{x}_n(t)$ equation of the parameterization of the balanced realization. Dividing the equation by r_n gives

$$\frac{1}{r_n} (\dot{x}_n(t)) = \sum_{i=1}^{n-1} r_i \alpha_{ni} x_i(t) - \frac{r_n x_n(t)}{2\theta_n} + \sum_{i=1}^m \beta_{ni} u_i(t). \quad (16)$$

Taking the limit as $r_n \rightarrow \infty$ with $r_n x_n$ remaining finite (by letting $x_n \rightarrow 0$) gives

$$r_n x_n(t) = 2\theta_n \sum_{i=1}^{n-1} r_i \alpha_{ni} x_i(t) + 2\theta_n \sum_{i=1}^m \beta_{ni} u_i(t). \quad (17)$$

Furthermore, in the remaining equations for $\dot{x}_i(t)$ ($i = 1, \dots, n - 1$) and $y_i(t)$ ($i = 1, \dots, p$), we find that r_n and x_n always appear in the combination $r_n x_n$. Therefore, this limit is a well-defined boundary approximation for this parameterization. Plugging (17) into the rest of the system, i.e. $\dot{x}_i(t)$, $i < n$, gives

$$\begin{aligned} \dot{x}_i(t) &= \sum_{j=1, j \neq i}^{n-1} (r_i r_j \alpha_{ij} + 2\theta_n r_i r_j \alpha_{in} \alpha_{nj}) x_j(t) \\ &+ (2\theta_n r_i^2 \alpha_{in} \alpha_{ni} - \frac{r_i^2}{2\theta_i}) x_i(t) \\ &+ \sum_{j=1}^m (r_i \beta_{ij} + 2\theta_n r_i \alpha_{in} \beta_{nj}) u_j(t) \\ &= \sum_{j=1}^{n-1} (a_{ij} - \frac{a_{in} a_{nj}}{a_{nn}}) x_j(t) \\ &+ \sum_{j=1}^m (b_{ij} - \frac{a_{in} b_{nj}}{a_{nn}}) u_j(t), \end{aligned} \quad (18)$$

which is the system in Equation (6) for $k = 1$, that is, BSPA of one state.

Now assume BSPA of k states is equivalent to k MBAM approximations, giving the system in Equation (6), denoted by $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$. Let the $k+1$ state BSPA be denoted by $(\tilde{A}^{k+1}, \tilde{B}^{k+1}, \tilde{C}^{k+1}, \tilde{D}^{k+1})$. By the Quotient Formula for Schur complements (Theorem 1.4 in [34]),

$$\tilde{A}^{k+1} = A/A_{k+1,k+1} = (A/A_{kk})/(a_{k+1,k+1}), \quad (19)$$

where

$$A_{k+1,k+1} = \left[\begin{array}{c|c} a_{k+1,k+1} & A_{k+1,k} \\ \hline A_{k,k+1} & A_{kk} \end{array} \right]. \quad (20)$$

This means that \tilde{A}^{k+1} from BSPA of $k+1$ states (the left hand side of Equation (19)) is equivalent to the \tilde{A}^{k+1} from one MBAM on the k^{th} MBAM/BSPA approximation (the right hand side).

Now we will show this for \tilde{B}^{k+1} , which is slightly more laborious because we cannot appeal to Schur complement properties. By Equation (6) for $k = k+1$, Equation (20), and partitioning $A_{1,k+1}$ and B_{k+1}

$$\tilde{B}^{k+1} = B_1 - [a_{k+1} \quad \tilde{A}_{1k}] \left[\begin{array}{c|c} a_{k+1,k+1} & A_{k+1,k} \\ \hline A_{k,k+1} & A_{kk} \end{array} \right]^{-1} \begin{bmatrix} b_{k+1}^T \\ B_k \end{bmatrix} \quad (21)$$

where \tilde{A}_{1k} indicates truncating off the last row of A_{1k} and a_{k+1} and b_{k+1}^T are vectors. By the block matrix inversion formula, matrix multiplication, and some rearranging of terms, the second term becomes

$$A_{1k} A_{kk}^{-1} B_k + \frac{1}{c} (a_{k+1} - \tilde{A}_{1k} A_{kk}^{-1} A_{k,k+1}) d \quad (22)$$

where $c = a_{k+1,k+1} - A_{k+1,k} A_{kk}^{-1} A_{k,k+1}$ and $d = b_{k+1}^T - A_{k+1,k} A_{kk}^{-1} B_k$. Note that c is just $\tilde{a}_{n-k,n-k}$, the last entry of the \tilde{A} matrix from the k^{th} MBAM/BSPA approximation. Similarly note $(a_{k+1} - \tilde{A}_{1k} A_{kk}^{-1} A_{k,k+1}) = \tilde{A}_{1,n-k}$, the last column of \tilde{A} excluding $\tilde{a}_{n-k,n-k}$, and $d = \tilde{B}_{n-k}$, the last row of \tilde{B} (with a slight abuse of notation $n-k$ indicates the row/column number and not the size for the tilde system). Also note that $B_1 - A_{1k} A_{kk}^{-1} B_k = \tilde{B}_1$, the first $n-k-1$ rows of \tilde{B} . Therefore Equation (21) becomes

$$\tilde{B}^{k+1} = \tilde{B}_1 - \frac{1}{\tilde{a}_{k-1,k-1}} \tilde{A}_{1,n-k} \tilde{B}_{n-k}, \quad (23)$$

which is one MBAM on the k^{th} MBAM/BSPA approximation. It can be shown similarly that the same relationship holds for \tilde{C}^{k+1} and \tilde{D}^{k+1} . Therefore, given that BSPA of k states is equivalent to k MBAM approximations, then BSPA of $k+1$ states is equivalent to $k+1$ MBAM approximations. Therefore, by induction, it holds for all k . ■

The resulting BSPA system is Hurwitz and is still a balanced realization with the $n-1$ HSVs, the same as the largest $n-1$ HSVs of the original system [24].

Note: In Equation (18) if θ_n , which is clearly no longer a HSV, is equal to its original value, then the reduced system is BSPA. However if θ_n is set to zero, then the resulting reduction is BT. Therefore, Theorem 3 offers a whole new class of reduced systems ranging between BSPA and BT and given a metric or goal, the optimal reduction can be found.

Consider the parameterization, presented in Section IV, of a two-state system

$$\begin{aligned} \dot{x} &= \begin{bmatrix} -r_1^2/(2\theta_1) & \frac{r_1 r_2 (\theta_2 - \theta_1)}{\theta_1^2 - \theta_2^2} \\ \frac{r_2 r_1 (\theta_1 - \theta_2)}{\theta_2^2 - \theta_1^2} & -r_2^2/(2\theta_2) \end{bmatrix} x + \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} u \\ y &= [r_1 \quad r_2] x + du, \end{aligned} \quad (24)$$

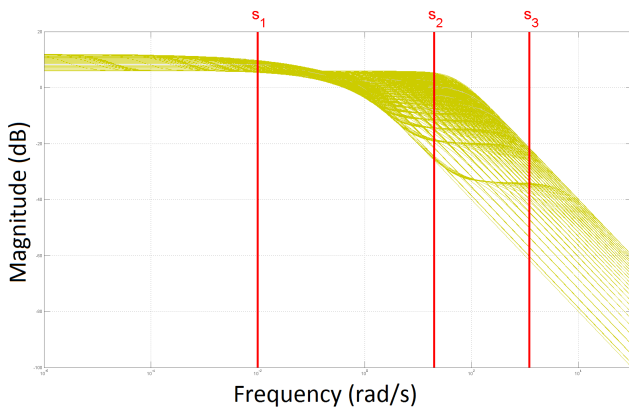
where $\theta_1 = r_1 = 1$ and θ_2 and r_2 are the free parameters. We plot the frequency responses of the different system in Figure 2a and use (s_1, s_2, s_3) to create the manifold. The model manifold, in Figures 2b-2c, has two boundaries. The top blue boundary indicates where $r_2 \rightarrow \infty$ (we use ten since it is an order of magnitude bigger than one), the bottom red boundary shows where $\theta_1 = \theta_2$, and where they meet at the bottom, the magenta dot, is BT, where $\theta_2 \rightarrow 0$ and/or $r_2 \rightarrow 0$.

Consider the system where $\theta_2 = .7$, $r_2 = 8$, depicted on the manifold by the red triangle. It is clear that BT is not going to be a good approximation of the system by the distance between the two systems on the manifold. However, the BSPA denoted by the cyan “plus” symbol is a very close approximation. Consider also another system where $\theta_2 = .01$, $r_2 = .8$, depicted on the manifold by the green circle. The BSPA, shown as a black “x”, is still fairly close to the true system but BT is much closer.

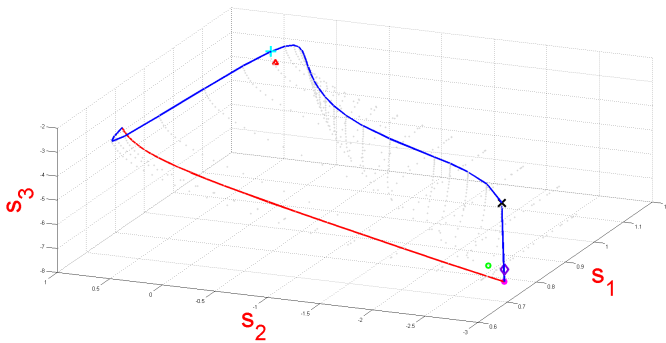
Although BT and BSPA have the same a priori H_∞ error bound, BT typically gives better results at high frequencies while SPA excels at low frequencies [26]. The optimal approximation requires one to identify a metric customized to a context of interest. As a concrete example, returning to the model in which $\theta_2 = .01$, $r_2 = .8$, the blue curve connecting the black “x” to the magenta dot represents a family of candidate reduced models that interpolates between BT and BSPA. The optimal reduced model relative to the three frequencies sampled is the point on this blue curve that is closest to the green dot, in the 2-norm sense. It is identified numerically as $\theta_2 = 5.3958 \times 10^{-4}$ (purple diamond). Clearly this point is much closer to the original system than either BSPA and BT and therefore a better approximation for this metric.

VI. CONCLUSION

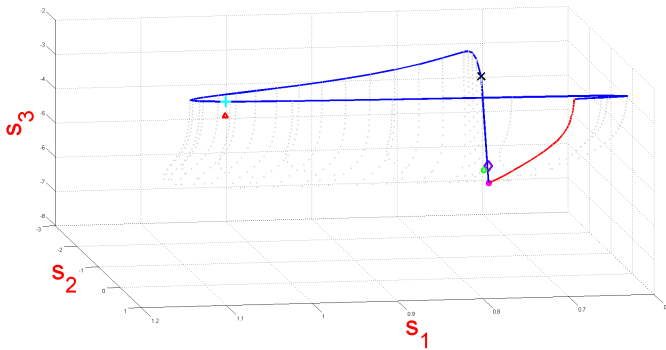
We have presented a complete parameterization of LTI systems that naturally partitions parameters into two groups: dynamic and structural. We have seen that the dynamic parameters are identifiable from physical observations and act as coordinates on a model manifold while structural parameters require a priori knowledge. By applying the Manifold Boundary Approximation Method to subsets of the dynamical parameters, Balanced Truncation and Balanced Singular Perturbation Approximation were shown to be limiting approximations of this parameterization. It is well known that these popular model reduction techniques can result in quite different approximations. This unified view of model reduction of LTI system illustrates that each approximation corresponds to a particular boundary element on the model



(a) Magnitude plot for the two-state space model. Different curves are calculated by varying the parameters θ_2 and r_2 .



(b) The Model Manifold for the two-state space model: Each point in the data space corresponds to a curve in Figure 2a. The top blue boundary indicates where $r_2 \rightarrow \infty$, the bottom red boundary shows where $\theta_1 = \theta_2$, and the magenta dot at the bottom is BT.



(c) Alternate view of the Model Manifold for the two-state model.

Fig. 2: Magnitude plot and the resulting manifold for the two-state model.

manifold, $\overline{\mathcal{M}(\mathcal{D})}$, embedded in the data space. Depending on the choice of metric, the manifold can give insight into which approximation is the best and give alternative approximations interpolating between BT and BSPA.

REFERENCES

[1] M. K. Transtrum and P. Qiu, "Model reduction by manifold boundaries," *Phys. Rev. Lett.*, vol. 113, Aug 2014.
 [2] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.

[3] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, "Graph structure in the web," *Computer networks*, vol. 33, 2000.
 [4] L. H. Hartwell, J. J. Hopfield, S. Leibler, and A. W. Murray, "From molecular to modular cell biology," *Nature*, vol. 402, 1999.
 [5] M. O. Jackson and A. Wolinsky, "A strategic model of social and economic networks," *Journal of economic theory*, vol. 71, no. 1, 1996.
 [6] N. Shimkin and A. Feuer, "Persistency of excitation in continuous-time systems," *Systems & control letters*, vol. 9, no. 3, 1987.
 [7] K. J. Aström and R. M. Murray, *Feedback systems: an introduction for scientists and engineers*. Princeton university press, 2010.
 [8] B. Moore, "Principal component analysis in linear systems: Controlability, observability, and model reduction," *Automatic Control, IEEE Transactions on*, vol. 26, no. 1, pp. 17–32, 1981.
 [9] G. E. Dullerud and F. Paganini, *A course in robust control theory*. Springer New York, 2000, vol. 6.
 [10] K. Glover, "All optimal hankel-norm approximations of linear multi-variable systems and their l_∞ -error bounds," *International Journal of Control*, vol. 39, no. 6, 1984.
 [11] D. F. Enns, "Model reduction with balanced realizations: An error bound and a frequency weighted generalization," in *CDC*, 1984.
 [12] C. L. Beck, J. Doyle, and K. Glover, "Model reduction of multidimensional and uncertain systems," *Automatic Control, IEEE Transactions on*, vol. 41, 1996.
 [13] S. Lall and C. Beck, "Model reduction of complex systems in the linear-fractional framework," in *Computer aided control system design, 1999*.
 [14] C. Beck, "Coprime factors reduction methods for linear parameter varying and uncertain systems," *Systems & control letters*, vol. 55, no. 3, pp. 199–213, 2006.
 [15] C. S. Hsu, "Decomposition of large scale systems—a numerical approach," in *ACC*, 1983.
 [16] A. Vandendorpe and P. Van Dooren, "On model reduction of interconnected systems," in *Proceedings International Symposium Math. Th. Netw. Syst., Belgium*, 2004.
 [17] L. Li and F. Paganini, "Structured coprime factor model reduction based on lmis," *Automatica*, vol. 41, 2005.
 [18] H. Sandberg and R. M. Murray, "Model reduction of interconnected linear systems," *Optimal Control Applications and Methods*, vol. 30, no. 3, pp. 225–245, 2009.
 [19] J. M. Scherpen, "Balancing for nonlinear systems," *Systems & Control Letters*, vol. 21, no. 2, 1993.
 [20] —, " h_∞ balancing for nonlinear systems," *Int. journal of robust and nonlinear control*, vol. 6, 1996.
 [21] S. Lall, J. E. Marsden, and S. Glavaški, "A subspace approach to balanced truncation for model reduction of nonlinear control systems," *International journal of robust and nonlinear control*, vol. 12, 2002.
 [22] P. V. Kokotovic et al., "Singular perturbations and order reduction in control theory—an overview," *Automatica*, vol. 12, no. 2, 1976.
 [23] T. Kato, *Perturbation theory for linear operators*. Springer, 1995.
 [24] K. Fernando and H. Nicholson, "Singular perturbational model reduction of balanced systems," *Automatic Control, IEEE Transactions on*, vol. 27, no. 2, 1982.
 [25] —, "Singular perturbational approximations for discrete-time balanced systems," *Automatic Control, IEEE Transactions on*, vol. 28, no. 2, pp. 240–242, 1983.
 [26] Y. Liu and B. D. Anderson, "Singular perturbation approximation of balanced systems," *International Journal of Control*, vol. 50, no. 4, pp. 1379–1405, 1989.
 [27] M. Transtrum, G. Hart, and P. Qiu, "Information topology identifies emergent model classes," 2014.
 [28] M. Spivak, *A comprehensive introduction to differential geometry*. Publish or Perish, 1979.
 [29] L. Michaelis and M. M. L. Menten, "Die kinetik der invertinwirkung," *Biochem Z*, vol. 49, pp. 333–369, 1913.
 [30] G. E. Briggs and J. B. S. Haldane, "A note on the kinetics of enzyme action," *Biochem Journal*, vol. 19, 1925.
 [31] M. Transtrum, B. B. Machta, and J. P. Sethna, "Why are nonlinear fits to data so challenging?" *Physical Review Letters*, vol. 104, 2010.
 [32] —, "Geometry of nonlinear least squares with applications to sloppy models and optimization," *Physical Review E*, 2011.
 [33] P. E. Paré, "Necessary and Sufficient Conditions for State-Space Network Realization," Master's thesis, Brigham Young University, Provo, UT, 2014.
 [34] F. Zhang, *The Schur complement and its applications*. Springer, 2006.