# Model Order Reduction in the Alignment Distance and Metrization of the Kalman Decomposition

Bijan Afsari and René Vidal, *Fellow, IEEE,*

***Abstract*—In this paper, we formulate the problem of model order reduction for LTI (MIMO) dynamical systems in terms of the alignment distance. The significance of this formulation is that it establishes a natural quantitative link between the "Kalman canonical decomposition" and "model order reduction," and fills an existing gap in this regard. The alignment distance includes a large class of state-space based distances on the manifold of systems of fixed minimal order $n$ and output-input dimension $(p, m)$; it is a natural distance associated with the quotient space structure of this manifold. The intuition behind our formulation is to consider systems of orders lower than $n$ as points on the boundary of the mentioned manifold in an appropriate ambient space; and the goal is to find a system of order at most $r$ (on the boundary) "closest" to a given system of order $n$, where closeness is measured in the alignment distance. In materializing this idea, certain theoretical and computational challenges arise, which will be addressed (e.g., while we have to extend the alignment distance to the boundary, the entire of the boundary is not metrizable; hence, we pass to a subset of the boundary, called diagonalizable s-balanced systems, and establish its metrizability). Ultimately, a computationally-friendly problem is formulated, for which, using methods of optimization on manifolds, we introduce an efficient algorithm called Align, Truncate, and Project (ATP). We also give some a-priori error bounds in terms of the Hankel singular values of the system. Interesting connections emerge with the popular balanced truncation method, which is a method not based on any optimality criterion. Our approach is applicable to both stable and unstable systems, and we establish robustness of feedback stability in the alignment distance.**

***Index Terms*—Linear dynamical systems, model order reduction, balanced realization, Kalman decomposition, alignment distance, metrizability, optimization on manifolds.**

## I. Introduction

In minimal realization theory of linear dynamical systems, the celebrated Kalman canonical decomposition [1] plays a similar role as the singular value decomposition (SVD) plays in the case of matrices, i.e., in the same way that the SVD of an $n \times n$ matrix reveals its *rank* $r$, the Kalman decomposition reveals the *minimal order* $r$ of a linear dynamical system of (state-space) order $n$. However, contrary to the SVD that solves or is instrumental in solving the problem of best rank-$r$ ($r < n$) approximation of a rank-$n$ matrix, there is *no* machinery that *quantitatively* relates the Kalman decomposition to the problem of *model order reduction*, namely, finding the best approximation of a minimal system of order $n$ with one of minimal order $r < n$. In the language of B. C. Moore [2], there is a *gap* between "minimal realization theory" and the model order reduction problem.[1] In this paper, we close this gap and show how to quantify or *metrize* the Kalman canonical decomposition.[2] The main question of interest to us is: how to make sense of "comparing" the Kalman decomposition of a minimal system of order $n$ and that of non-minimal systems of order $n$ and minimal order $r < n$, and thereby formulating model order reduction as finding a Kalman decomposition of minimal order $r$ "closest" to a given one of minimal order $n$.

The core idea in our solution is the notion of "realization alignment," which is the main idea behind in the recently introduced *alignment distance*, and refers to finding the "best" state-space change of basis that brings given realizations of two minimal systems "as close as possible" [3], [4]. Indeed, the alignment distance is nothing but a way to compare the Kalman canonical decompositions of two minimal systems. Recall that a linear system of order $n$ has an equivalence class of realizations, all related by the so-called *similarity action* or *transformation*, i.e., a state-space change of basis under $GL(n)$, the Lie group of non-singular $n \times n$ matrices. This leads to the *quotient* space structure of the space of systems of fixed order $n$ (and input-output size $(m, p)$, which is immaterial for our purposes). To find the alignment distance, one first *aligns* given realizations of two systems and then compares them. We have established that by considering *balanced* realizations of the systems, the alignment can be done by only an *orthogonal* change of basis—something which brings about significant computational advantages [3]. The alignment distance, however, is defined on the *manifold* of minimal systems of fixed order $n$. The (intuitive) geometric picture that we base our approach on is that of the manifold of minimal systems of order $n$ sitting inside the space of all systems of order $n$, as an *open* subset, with non-minimal systems as *boundary* points. Thus, our main goal is to formulate a "model order reduction" problem, by extending—in a useful way—the notion of "realization alignment" and the "alignment distance" to the boundary of the manifold of minimal systems—thereby comparing the Kalman decompositions of minimal and non-minimal systems, in a meaningful way.[3]

[1] The "gap," indeed, is something that many students studying control theory feel once they are taught the beautiful "Kalman decomposition;" as afterwards they are left in vacuum, in the sense that nothing is done with it.

[2] In topology, the term "metrization" has a very specific meaning having to do with equipping a topological space with a metric or distance function that matches the original topology of the space. Here, we use the term "metrization" more in sense of "quantification" with an eye on the exact topological meaning, which becomes relevant too.

[3] As it turns out this geometric picture is not quite accurate, since not all non-minimal systems can be included in this metric setting; instead, only a subset of non-minimal systems that we call diagonalizable s-balanced systems can be metrized, see Definitions 8 and 14.

Both authors are with the Center for Imaging Science, The Johns Hopkins University, Baltimore, MD USA e-mail: ({bijan,rvidal}@cis.jhu.edu).

This work has been supported by grants NSF 1335035 and ONR N000141310116.

Manuscript received August X, 2015; revised August X, 2015.

### A. Prior Work: Diagonally (D-) Balanced Truncation

Attempts at closing the mentioned gap have been made before, by Moore [2] and others, leading to the notion of "diagonally balanced (d-balanced) realizations"[4] and the method of "d-balanced truncation" for model order reduction. The method simply involves retaining "the strong sub-realization" in a d-balanced realization and throwing out the weak sub-realization. The strong sub-realization is the one which is associated with the largest $r$ Hankel singular values of the system. This method is perhaps the most popular model order reduction method. This is in particular interesting, since while all other model order reduction approaches are based on the input-output description of linear systems, the "d-balanced truncation" method is based on the state-space description. Its popularity can be attributed to its simplicity, effectiveness, and computational efficiency as well as certain neat theoretical error guarantees (see [5], [6]). The d-balanced truncation method, however, is *not* quite an answer to the question of how to quantitatively relate the Kalman canonical decomposition and "model order reduction;" firstly, because in it there is no *quantification* of "distance to non-minimality" beyond the use of Hankel singular values as a measure of minimality—in a loose similarity with how matrix singular values can quantify a distance to rank deficiency. Stated differently, the "d-balanced truncated realization" is *not* optimal in any quantitative sense. Secondly, in this approach the role of "state-space change of basis" is ignored, as an attempt is made to identify a minimal *system* with *a* d-balanced *realization*, whereas it is well known that *topologically* this is not possible [7]—this is manifested by the fact that a d-balanced realization has an orthogonal ambiguity, when the Hankel singular values are not distinct. In [2], Moore acknowledges "theoretical gaps" in the justification for selecting the "strong sub-realization," as he leaves it as a conjecture that "internal dominance" of the "strong sub-realization"—a property which is realization or state-space basis specific—implies "external dominance" —which is realization-independent. However, it is shown by Kabamba [8] that "internal dominance" does *not* imply "external dominance," i.e., the magnitudes of the Hankel singular values alone do not determine the weight of a sub-realization in the overall impulse response.

### B. Scope of the Paper, Some History, and Contributions

This paper can be considered as a sequel to [3], and its scope is primarily theoretical and devoted to (both stable and unstable) Multi-Input Multi-Output (MIMO) discrete-time deterministic systems. Nevertheless, in principle, our results and methodology should be extendable to other cases with little effort (see e.g., [9] and [3]). The primary goal is to rigorously establish how "minimal realization theory" can lead to "model order reduction" using the alignment distance. The alignment distance includes a large class of distances, although so far the simplest one, namely, the Frobenius norm based distance has been studied (see (6)). Thus, effectively, we are introducing a large toolbox of model order reduction methods; however, the effectiveness of such a toolbox, in practice, has to be examined extensively, and that is beyond the scope of this paper. Although the related optimization is *not* a convex problem, we believe that our proposed algorithm is quite efficient and remains practical for systems of moderate order (around $40-50$). It is important to mention that this work, in some sense, belongs to a line of research started in the 1970s by Kalman and others in which the *algebraic* or *differential* geometry of the state-space description of linear dynamical systems was studied (see e.g., [7], [10]–[13] and references cited in [3]). That line of research, however, did not result in significant computational tools, and was more or less abandoned by the end of 1980s. In that regard, we believe that the current work is an important step forward.

The main idea of model order reduction using the alignment distance was introduced in [14]. However, the current paper contains major improvements, refinements, and extensions to that work, including proofs, new results and new algorithms.

### C. Summary and Outline

The rest of this paper is organized as follows. In Section II, notation is established, and the basic theory of alignment distance is reviewed. In Section III, we study the boundaries of manifolds of minimal realizations and systems, and distinguish between the standard *boundary* and the *elevated boundary*—a relaxed form of boundary that will be more practical to work with. We also introduce the notions of *balanced Kalman decomposition*, and *s-balanced realizations* and state the Helmke-Moore lemma. This paves the way to extend the alignment distance to non-minimal systems. In Section IV, we give several possible formulations for "model order reduction" based on the alignment distance. For computational reasons, our choice is Definition 15, which is based on the elevated boundary, and some of its basic properties are studied. In Section V, more general ways of metrization of the Kalman decomposition are studied. Theorem 23 is an important result which identifies a metrizable space that includes both the minimal and non-minimal systems—this is the space of diagonalizable s-balanced systems. In Section VI, an efficient alternating minimization algorithm called Align, Truncate, and Project (ATP) is derived to solve the model order reduction problem. Optimization methods on manifolds are used to this end, and yield significant improvements over an earlier version in [14]. In Section VII, an a-priori bound on the model order reduction error, similar to the well-known results for d-balanced truncation, is given, and connection with d-balanced truncation is studied. In Section VIII, as an application, we show that internal stability under constant gain feedback is a robust property in the alignment distance. In Section IX, an example is given for reduction of an unstable system; and Section X concludes the paper. Appendices A and B contain certain proofs and derivations.

## II. Preliminaries on the Manifolds of Systems and the Alignment Distance

The reader is referred to [3] for a detailed and rigorous introduction to the alignment distance. We consider a deterministic discrete-time LTI dynamical (or state-space) system

[4]Since we will deal with a slightly more general form of balancing, which contrary to d-balancing has a differential geometric meaning, we explicitly distinguish between "balancing" and "d-balancing" (see Sections II-B).

$M$ of *order* $n$ and input-output *size* $(m, p)$ described by:

$$\begin{cases} x_t = Ax_{t-1} + Bu_t \\ y_t = Cx_t, \end{cases} \quad (1)$$

where $R = (A, B, C) \in \widetilde{\mathcal{L}}_{m,n,p} = \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}$ is called *a realization* of $M$. Here, $u_t$ is the $m$-dimensional input assumed to be a *deterministic* stimulus. We denote the subset of (internally) asymptotically stable (a.s.) realizations by $\widetilde{\mathcal{L}}^{\mathrm{a}}_{m,n,p}$. Given a positive integer $r < n$, partition the matrices conformaly as $A = \left[\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right]$, $B = \left[\begin{smallmatrix} B_{11} \\ B_{21} \end{smallmatrix}\right]$ $C = \left[\begin{smallmatrix} C_{11} & C_{12} \end{smallmatrix}\right]$, where $A_{11}$ is $r \times r$. We call $R_{11} = (A_{11}, B_{11}, C_{11})$ and $R_{22} = (A_{22}, B_{21}, C_{12})$ the top (or 11) and the bottom (or 22) sub-realizations of $R$, respectively. For a (possibly unstable) realization denote by $\mathcal{O}_k = [C^\top, (CA)^\top, \ldots, (CA^{k-1})^\top]^\top$ and $\mathcal{C}_k = [B, AB, \ldots, A^{k-1}B]$ the observability and controllability matrices of order $k$ ($n \leq k \leq \infty$). Here $^\top$ denotes the transpose operation. The observability and controllability Gramians of order $k \geq n$ are defined as $W_{\mathrm{o},k} = \mathcal{O}_k^\top \mathcal{O}_k$ and $W_{\mathrm{c},k} = \mathcal{C}_k \mathcal{C}_k^\top$, respectively. For $k = \infty$ a.s. of $A$ is needed; in this case depending on the situation we may use $W_{\mathrm{o}}$ and $W_{\mathrm{c}}$ or $W_{\mathrm{o},k}$ and $W_{\mathrm{c},k}$ while letting $k = \infty$. For an a.s. realization $R$ the controllability and observability Gramians satisfy the Lyapunov equations

$$W_{\mathrm{c}} = BB^\top + AW_{\mathrm{c}}A^\top, \quad (2a)$$
$$W_{\mathrm{o}} = C^\top C + A^\top W_{\mathrm{o}} A. \quad (2b)$$

The Hankel singular values of $M$ which we denote by $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_n$ are the square roots of the eigenvalues of $W_{\mathrm{o}}W_{\mathrm{c}}$ (they are invariant under a state-space change of coordinates). Let $\Lambda$ be a diagonal matrix with diagonal $\lambda_1, \ldots, \lambda_n$. If the realization $R$ is d-balanced, then we can assume $W_{\mathrm{c}} = W_{\mathrm{o}} = \Lambda$. Given a d-balanced realization $R$, we always assume that the singular values are non-increasingly ordered, and call $R_{11}$ and $R_{22}$ the *strong* and *weak* sub-realizations, respectively [2].

### A. Quotient Manifolds of Minimal Systems of Fixed Order

We distinguish between a realization $R$ and $M$, the system realized by $R$, which has an equivalence class of realizations, all related by a change of coordinates under $GL(n)$. Specifically, denote the state-space change of basis or the so-called *similarity action* by $\circ$, where for any $P \in GL(n)$

$$P \circ R = (P^{-1}AP, P^{-1}B, CP). \quad (3)$$

Then $R$ and $P \circ R$ are indistinguishable from an input-output point of view. Thus, the *space of systems* is the *quotient* of the space of realizations under the action $\circ$. We write $\mathcal{L}_{m,n,p} = \widetilde{\mathcal{L}}_{m,n,p}/GL(n)$ and $\mathcal{L}^{\mathrm{a}}_{m,n,p} = \widetilde{\mathcal{L}}^{\mathrm{a}}_{m,n,p}/GL(n)$ for the a.s. case, and assume that both are equipped with the *natural quotient topology*. We call $[R] = \{P \circ R | P \in GL(n)\}$ (also denoted as $GL(n) \circ R$) the $GL(n)$-orbit of $R$, and identify it with $M$.

The space $\mathcal{L}_{m,n,p}$ is not a nice mathematical object (e.g., it is not even Hausdorff, see Section III-B). However, if we restrict attention to the space (manifold) of minimal realizations $\widetilde{\Sigma}^{\min}_{m,n,p}$ or a.s. minimal realizations $\widetilde{\Sigma}^{\min,\mathrm{a}}_{m,n,p}$, then their respective quotient spaces (namely $\Sigma^{\min}_{m,n,p} \triangleq \widetilde{\Sigma}^{\min}_{m,n,p}/GL(n)$ and $\Sigma^{\min,\mathrm{a}}_{m,n,p} \triangleq \widetilde{\Sigma}^{\min,\mathrm{a}}_{m,n,p}/GL(n)$) are smooth manifolds of dimension $n(m+p)$. Here, smoothness comes from the usual notion of smoothness in the Euclidean space $\widetilde{\mathcal{L}}_{m,n,p}$. The realization-space pairs $(\widetilde{\Sigma}^{\min}_{m,n,p}, \Sigma^{\min}_{m,n,p})$ and $(\widetilde{\Sigma}^{\min,\mathrm{a}}_{m,n,p}, \Sigma^{\min,\mathrm{a}}_{m,n,p})$ form an object called *principal fiber bundle* with *structure group* $GL(n)$.

Defining a (group action induced) distance on the bottom or base space $\Sigma$ of a generic principal bundle $(\widetilde{\Sigma}, \Sigma)$ with structure group $GL(n)$ is conceptually simple: Given a $GL(n)$-invariant distance $\tilde{d}_{\widetilde{\Sigma}}$ on the top space one defines $d_\Sigma(M_1, M_2) = \inf_{P \in GL(n)} \tilde{d}_{\widetilde{\Sigma}}(P \circ R_1, R_2)$, where $R_i (i = 1, 2)$ is *any* realization (or representation) of $M_i$. The distance $\tilde{d}_{\widetilde{\Sigma}}(P \circ R_1, R_2)$ is $GL(n)$-invariant if $\tilde{d}_{\widetilde{\Sigma}}(P \circ R_1, P \circ R_2) = \tilde{d}_{\widetilde{\Sigma}}(R_1, R_2)$ for $\forall P \in GL(n)$ and $\forall R_1, R_2 \in \widetilde{\Sigma}$. The simple intuition here is to align the two realizations (bring them as *close* as possible) by sliding one along the fiber it belongs to. We call such a distance a *noncompact alignment distance*. The main difficulty with the noncompact alignment distance is that (due to noncompactness of $GL(n)$) constructing a $GL(n)$-invariant distance $\tilde{d}_{\widetilde{\Sigma}}$ is complicated (see [3], for details).

### B. Bundle Reduction and the Alignment Distance

One might wonder if we could somehow replace $GL(n)$ with $O(n)$, its (compact) subgroup of orthogonal matrices. The answer is positive, and in a general setting it is called *reduction of the structure group*, a notion which has a precise meaning in differential geometry [3], [15]. In the context of control applications certain forms of realization (Gramian) *balancing* [16] and [17] can be linked to the reduction of structure group [3]. Define the set of a.s. balanced minimal realizations as

$$\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a,bl}}_{m,n,p} = \{(A, B, C) \in \widetilde{\Sigma}^{\min,\mathrm{a}}_{m,n,p} | W_{\mathrm{o}} = W_{\mathrm{c}} \succ 0\}, \quad (4)$$

and the set of $k$-balanced minimal realizations ($n \leq k < \infty$) as

$$\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p} = \{(A, B, C) \in \widetilde{\Sigma}^{\min}_{m,n,p} | W_{\mathrm{o},k} = W_{\mathrm{c},k} \succ 0\}, \quad (5)$$

where $X \succ 0$ means that $X$ is a positive definite matrix. In comparison with the sets of d-balanced realizations, in these sets, diagonality of the Gramians is not enforced. If $R$ belongs to $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a,bl}}_{m,n,p}$ so does $Q \circ R$ for every $Q \in O(n)$. Conversely if $R$ and $P \circ R$ belong to $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a,bl}}_{m,n,p}$ for some $P \in GL(n)$, then $P \in O(n)$ [3]. The same holds for $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p}$. One can show that $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a,bl}}_{m,n,p}$ and $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p}$ are smooth submanifolds (called *reduced subbundles*) of $\widetilde{\Sigma}^{\min,\mathrm{a}}_{m,n,p}$ and $\widetilde{\Sigma}^{\min}_{m,n,p}$, respectively [3]. The key point is that $\Sigma^{\min,\mathrm{a}}_{m,n,p} = \widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a,bl}}_{m,n,p}/O(n)$ and $\Sigma^{\min}_{m,n,p} = \widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p}/O(n)$, where equality is in the sense of *diffeomorphism* (see [3] for more details). Other forms of balancing and reduction of the structure group are possible [3]. We may use "reduction" and "standardization," interchangeably, and call a realization in a reduced subbundle a standardized realization. In the rest of the paper, we may write $(\widetilde{\mathcal{O}\Sigma}_{m,n,p}, \Sigma_{m,n,p})$ to denote either $(\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a,bl}}_{m,n,p}, \Sigma^{\min,\mathrm{a}}_{m,n,p})$ or $(\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p}, \Sigma^{\min}_{m,n,p})$. By $\overline{\widetilde{\mathcal{O}\Sigma}}_{m,n,p}$ we mean the closure of $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p}$ in $\widetilde{\mathcal{L}}_{m,n,p}$ or the closure of $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a,bl}}_{m,n,p}$ in $\widetilde{\mathcal{L}}^{\mathrm{a}}_{m,n,p}$. Also in certain cases (e.g., in Section VI) we let $k = \infty$ in $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p}$, in which case $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},\infty}_{m,n,p} \equiv \widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a,bl}}_{m,n,p}$.

The alignment distance can be defined in a very general setting [3]. Here, we limit ourselves to some specific choices. The starting point is an $O(n)$-invariant distance on the *realization* space. Our choice is the Frobenius norm based distance:

$$\tilde{d}^2_F(R_1, R_2) = \|A_1 - A_2\|_F^2 + \|B_1 - B_2\|_F^2 + \|C_1 - C_2\|_F^2, \quad (6)$$

where $R_i = (A_i, B_i, C_i), i = 1, 2$. Then we define:

*Definition 1 (Alignment Distance):* Let $M_1$ and $M_2$ be two systems in $\Sigma_{m,n,p}$. The alignment distance subordinate to (standardization) $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ is defined as

$$d_F(M_1, M_2) = \min_{Q\in O(n)} \tilde{d}_F(Q\circ R_1, R_2), \tag{7}$$

where $R_i (i = 1, 2)$ is any realization of $M_i$ in $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$. The above minimization problem is called the realization alignment problem subordinate to $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$.

Thus the realization alignment problem is simply defined as *aligning* two standardized realizations by an orthogonal state-space change of basis so as to minimize the above distance. The alignment distance is a bona fide distance, i.e., it is symmetric, positive definite and obeys the triangle inequality [3]. Computing the alignment distance amounts to a (non-convex) optimization on the manifold $O(n)$, for which efficient algorithms can be devised (see [18]).

### C. Positive Definite (p.d-) Balancing

By a ($k$-)balancing matrix (transformation or change of basis) for $R \in \widetilde{\Sigma}^{\min}_{m,n,p}$, we mean $P \in GL(n)$ such that $P\circ R \in \widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p}$. This means that $P^\top W_{\mathrm{o},k}P = P^{-1}W_{\mathrm{c},k}P^{-\top}$. If $P$ is a balancing matrix for $R$ so is $PQ$ for any $Q \in O(n)$. However, as it can be shown, the symmetric positive definite (p.d.) matrix $S = PP^\top$ which solves the balancing equation:

$$SW_{\mathrm{o},k}S = W_{\mathrm{c},k} \tag{8}$$

is *unique*; thus the *p.d-balancing* matrix, $P = \sqrt{S}$, which is the (unique) square root of $S$, is unique [19, Ch. 8], [3]. Interestingly, $S$ is the Riemannian average of $W^{-1}_{\mathrm{o},k}$ and $W_{\mathrm{c},k}$ on $\mathcal{S}(n)$ the manifold of $n\times n$ symmetric p.d. matrices (see [3] for details). Finding $S$ is as simple as finding any other balancing transformation $P$, i,e., we just set $S = PP^\top$, although we also have: $S = W^{1/2}_{\mathrm{c},k}(W^{-1/2}_{\mathrm{c},k}W^{-1}_{\mathrm{o},k}W^{-1/2}_{\mathrm{c},k})^{1/2}W^{1/2}_{\mathrm{c},k}$. P.d-balancing is useful in establishing certain important results about balanced realizations (e.g., in Sections III and VI-A). If $R \in \widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p}$, then we have $S = I_n$. As mentioned earlier, this property is closely related to the notion of "reduction of the structure group" in a principal fiber bundle. Indeed, the *smooth* map $\nu : \widetilde{\Sigma}^{\min}_{m,n,p} \to \mathcal{S}(n)$, where $\nu(R) = S^{-1}$ and $S$ solves (8), is called a *bundle reduction* map, and $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p} = \nu^{-1}(I_n)$ is a reduced subbundle of $\widetilde{\Sigma}^{\min}_{m,n,p}$ [3].

## III. The Balanced Kalman Decomposition and the Boundaries of Manifolds of Minimal Realizations

In this section, we first describe the *boundaries* of the manifolds of balanced minimal realizations. We distinguish between "boundary" and "elevated boundary," the latter being a *relaxed* form of boundary, leading to a computationally simpler model order reduction formulation. In Subsection III-B, we introduce s-balanced realizations, a generalization of balanced realizations. We also state the Helmke-Moore Lemma, which helps us to understand why the quotient system spaces $\mathcal{L}_{m,n,p}$ and $\mathcal{L}^{\mathrm{a}}_{m,n,p}$ are non-Hausdorff, and how to deal with it. Most of our results are expressed in terms of balanced Kalman standard (or canonical) realizations defined as:

*Definition 2 (Balanced Kalman Standard Realization):* We call a realization $\bar{R} \in \widetilde{\mathcal{L}}_{m,n,p}$ of the form

$$\bar{A} = \begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & \bar{A}_{22} \end{bmatrix}, \bar{B} = \begin{bmatrix} \bar{B}_{11} \\ 0 \end{bmatrix}, \bar{C} = \begin{bmatrix} \bar{C}_{11} & 0 \end{bmatrix}, \tag{9}$$

with $\bar{R}_{11} = (\bar{A}_{11}, \bar{B}_{11}, \bar{C}_{11}) \in \widetilde{\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ a $k$-balanced Kalman standard (or canonical) realization of minimal order $0 \le r \le n$. If $\bar{R}_{11} \in \widetilde{\Sigma}^{\min,\mathrm{a}}_{m,r,p}$ and $\bar{A}_{11}$ is a.s., then we call $R$ an a.s. balanced Kalman standard realization of minimal order $r$. If $\bar{R}$ is of the form (9) but $\bar{R}_{11}$ is not minimal, then we call $\bar{R}$ a balanced Kalman standard realization of minimal order not larger than $r$. In general, if $R \in \widetilde{\mathcal{L}}_{m,n,p}$ has a Kalman canonical decomposition in which the standard realization is balanced, then we call the decomposition a balanced Kalman canonical decomposition of $R$.

### A. Boundaries of Manifolds of Balanced Minimal Realizations

First, we study the boundaries of $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a,bl}}_{m,n,p}$ (in $\widetilde{\mathcal{L}}^{\mathrm{a}}_{m,n,p}$) and $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p}$ (in $\widetilde{\mathcal{L}}_{m,n,p}$). We start by recalling that in a metric space $\mathcal{M}$, a point $x \notin \mathcal{S}$ is a boundary point of $\mathcal{S} \subset \mathcal{M}$ if and only if (iff) it is the limit of a sequence of points in $\mathcal{S}$.

*Proposition 3:* A realization $R \in \widetilde{\mathcal{L}}^{\mathrm{a}}_{m,n,p}$ of minimal order $r < n$ belongs to the boundary of $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a,bl}}_{m,n,p}$ iff $R = Q\circ\bar{R}$, where $Q \in O(n)$ and $\bar{R} = (\bar{A}, \bar{B}, \bar{C})$ is an a.s. Kalman standard realization with $\bar{A}_{22}$ being the $A$-matrix of a balanced realization in $\widetilde{\Sigma}^{\min,\mathrm{a}}_{m,n-r,p}$ (independent of $\bar{R}_{11}$).

*Proof:* Since $R$ is on the boundary of $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a,bl}}_{m,n,p}$ and of minimal order $r$, the Gramians of $R$ are equal of rank $r$. Thus with an orthogonal matrix $\bar{Q} \in O(n)$ the Gramians, $W_{\mathrm{c}}$ and $W_{\mathrm{o}}$, can be converted to diagonal form $\bar{Q}^\top W_{\mathrm{c}}\bar{Q} = \bar{Q}^\top W_{\mathrm{o}}\bar{Q} = \Lambda = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix}$, where the diagonal $r\times r$ matrix $\Lambda_1$ is positive definite (with diagonal elements in non-increasing order) and the diagonal $(n-r)\times(n-r)$ matrix $\Lambda_2$ is zero. Accordingly, the realization $\bar{R} = \bar{Q}\circ R$ is d-balanced and we have $R = Q\circ\bar{R}$ with $Q = \bar{Q}^\top$. So it suffices to prove a d-balanced version of the claim and characterize the boundary points which are d-balanced, namely where $\bar{R}_{11}$ is d-balanced and $\bar{A}_{22}$ is the $A$ matrix of a d-balanced realization. Next, we try construct the (d-balanced) boundary points as limits of (d-balanced) minimal realizations. To this end, first we consider the continuous-time case in which parameterization of a d-balanced realization $R^{\mathrm{c}} = (A^{\mathrm{c}}, B^{\mathrm{c}}, C^{\mathrm{c}})$ is straightforward [8], [20] and translate the result to the discrete-time case via the well-known Mobious transformation [20]–[22]:

$$\begin{aligned}\varphi(A^{\mathrm{c}}, B^{\mathrm{c}}, C^{\mathrm{c}}) =&((I_n - A^{\mathrm{c}})^{-1}(I_n + A^{\mathrm{c}}), \\ &\sqrt{2}(I_n - A^{\mathrm{c}})^{-1}B^{\mathrm{c}}, \sqrt{2}C^{\mathrm{c}}(I_n - A^{\mathrm{c}})^{-1}),\end{aligned} \tag{10}$$

which is a homeomorphism between the manifolds of continuous-time and discrete-time a.s. minimal realizations of order $n$ and size $(p, m)$ with the inverse

$$\begin{aligned}\varphi^{-1}(A, B, C) =&((I_n + A)^{-1}(A - I_n), \\ &\sqrt{2}(I_n + A)^{-1}B, \sqrt{2}C(I_n + A)^{-1}),\end{aligned} \tag{11}$$

Crucially, $R^{\mathrm{c}}$ and $\varphi(R^{\mathrm{c}})$ have the same respective Gramians [20], [21]. Thus $\varphi$ and $\varphi^{-1}$ remain continuous and bounded up to the (bounded) boundary points (non-minimal realizations).

In continuous-time, the free parameters for d-balanced realizations are: the Hankel singular values $\lambda_i > 0$ $(1 \le i \le n)$, matrices $\dot{B} \in \mathbb{R}^{n\times m}$ and $\dot{C} \in \mathbb{R}^{p\times n}$, with the restriction that $\|\dot{B}_i\| = 1$ and $\|\dot{C}_i\| = 1$ and if $\lambda_i = \lambda_j$ then $\dot{B}_i\dot{B}_j^\top = \dot{C}_i^\top\dot{C}_j$ ($\dot{B}_i$ and $\dot{C}_i$ being the $i^{th}$ row and column of $\dot{B}$ and $\dot{C}$, respectively), and the balanced gains $\gamma_i \ge 0$ $(1 \le i \le n)$. Define $\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)$. Then we have:

$$B^{\mathrm{c}} = \Gamma\dot{B}, \quad C^{\mathrm{c}} = \dot{C}\Gamma \tag{12a}$$

$$[A^{\mathrm{c}}]_{ij} = a^{\mathrm{c}}_{ij} = \begin{cases} -\frac{\gamma_i^2}{2\lambda_i} & i=j, 1 \le i \le n \\ \frac{\gamma_i\gamma_j(\dot{B}_i\dot{B}_j^\top\lambda_j - \dot{C}_i^\top\dot{C}_j\lambda_i)}{\lambda_i^2-\lambda_j^2} & i \ne j, \lambda_i \ne \lambda_j, \end{cases} \tag{12b}$$

where if $\lambda_i = \lambda_j$ then $a^{\mathrm{c}}_{ij} + a^{\mathrm{c}}_{ji} = \frac{\gamma_i\gamma_j\dot{B}_i\dot{B}_j^\top}{\lambda_i}$. This latter condition means that if $\lambda_i = \lambda_j$, then the skew-symmetric part of the matrix $\left[\begin{smallmatrix} \bar{a}^{\mathrm{c}}_{ii} & \bar{a}^{\mathrm{c}}_{ij} \\ \bar{a}^{\mathrm{c}}_{ji} & \bar{a}^{\mathrm{c}}_{jj}\end{smallmatrix}\right]$ is left arbitrary. In (12b), By sending the last $n-r$ singular values $\lambda_{r+1},\dots,\lambda_n$ to zero and keeping the other free variables constant, $A^{\mathrm{c}}$ will become unbounded, which is not of interest to us (see below). To get bounded $A^{\mathrm{c}}$, $\gamma_i^2$ has to be sent to zero with (at least) the same rate as $\lambda_i$. Let $\lambda_i = \varepsilon^2\ddot{\lambda}_i$ and $\gamma_i = \varepsilon\dot{\gamma}_i$ for $r < i \le n$ ($\ddot{\lambda}_i > 0, \dot{\gamma}_i \ge 0$). Denote the limit realization as $\varepsilon \to 0$ by $\bar{R}^{\mathrm{c}} = (\bar{A}^{\mathrm{c}}, \bar{B}^{\mathrm{c}}, \bar{C}^{\mathrm{c}})$. Observe that the last $n-r$ rows and columns of $\bar{B}^{\mathrm{c}}$ and $\bar{C}^{\mathrm{c}}$ are zero, i.e., $\bar{B}^{\mathrm{c}} = \left[\begin{smallmatrix}\bar{B}^{\mathrm{c}}_{11} \\ 0\end{smallmatrix}\right]$, $\bar{C}^{\mathrm{c}} = [\,\bar{C}^{\mathrm{c}}_{11}\ 0\,]$. For the index pair $(i,j)$ with $i > r$ and $j \le r$ and the index pair $(i,j)$ with $j > r$ and $i < j$ we have $\bar{a}^{\mathrm{c}}_{ij} = 0$. This means that blocks $\bar{A}^{\mathrm{c}}_{12}$ and $\bar{A}^{\mathrm{c}}_{21}$ are zero. The elements of the block $\bar{A}^{\mathrm{c}}_{11}$ are determined based only on the free parameters with index $i,j \le r$, hence the sub-realization $(\bar{A}^{\mathrm{c}}_{11}, \bar{B}^{\mathrm{c}}_{11}, \bar{C}^{\mathrm{c}}_{11})$ is d-balanced and minimal (order $r$). Next note that for $r+1 \le i,j \le n$

$$\bar{a}^{\mathrm{c}}_{ij} = \begin{cases} -\frac{\dot{\gamma}_i^2}{2\ddot{\lambda}_i} & i=j \\ \frac{\dot{\gamma}_i\dot{\gamma}_j(\dot{B}_i\dot{B}_j^\top\ddot{\lambda}_j - \dot{C}_i^\top\dot{C}_j\ddot{\lambda}_i)}{\ddot{\lambda}_i^2-\ddot{\lambda}_j^2} & i \ne j, \ddot{\lambda}_i \ne \ddot{\lambda}_j, \end{cases} \tag{13}$$

and if $\ddot{\lambda}_i = \ddot{\lambda}_j$, then $\bar{a}^{\mathrm{c}}_{ij} + \bar{a}^{\mathrm{c}}_{ji} = \frac{\dot{\gamma}_i\dot{\gamma}_j\dot{B}_i\dot{B}_j^\top}{\ddot{\lambda}_i}$. Thus $\bar{A}^{\mathrm{c}}_{22}$ is determined as if it were the $A$-matrix of a d-balanced minimal realization of order $n-r$ determined by $\dot{B}_{21}, \dot{C}_{12}$ and $\ddot{\lambda}_i, \dot{\gamma}_i$ $(r+1 \le i \le n)$. (This realization is indeed $\dot{\bar{R}}^{\mathrm{c}}_{22} = (\bar{A}^{\mathrm{c}}_{22}, \dot{\Gamma}_2\dot{B}_{21}, \dot{C}_{12}\dot{\Gamma}_2)$, where $\dot{\Gamma}_2 = \mathrm{diag}(\dot{\gamma}_{r+1},\dots,\dot{\gamma}_n)$.) Thus as $R^{\mathrm{c}}$ approaches a boundary point $\bar{R}^{\mathrm{c}}$, $\varphi(R^{\mathrm{c}})$ also approaches a boundary point $\bar{R}$ which is also d-balanced. (Since $\varphi$ maps an unbounded $A^{\mathrm{c}}$ to an unstable discrete-time $A$, we only need to consider bounded boundary points $\bar{R}^{\mathrm{c}}$.) Since $\bar{A}^{\mathrm{c}}$ is block-diagonal, we note that the top realization of $\bar{R}$, $\bar{R}_{11}$, is d-balanced, $\bar{A}_{21} = 0$, $\bar{A}_{12} = 0$, $\bar{B}_{21} = 0$, and $\bar{C}_{12} = 0$. Moreover, we have $\bar{A}_{22} = (I_{n-r} - \bar{A}^{\mathrm{c}}_{22})^{-1}(I_{n-r} + \bar{A}^{\mathrm{c}}_{22})$, which means that $\bar{A}_{22}$ is the $A$-matrix of a discrete-time d-balanced realization of order $n-r$ (that realization being $\varphi(\dot{\bar{R}}^{\mathrm{c}}_{22})$). This completes the proof of the d-balanced version of the result; as explained, the balanced version follows immediately. ∎

*Proposition 4:* Any realization $R \in \widetilde{\mathcal{L}}_{m,n,p}$ of minimal order $r < n$ on the boundary of $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$ $(n \le k < \infty)$ can be written as $Q \circ \bar{R}$, where $\bar{R} \in \widetilde{\mathcal{L}}_{m,n,p}$ is a $k$-balanced Kalman standard realization of minimal order $r$.

*Proof:* Similar to the proof of Proposition 3 we see that $R = Q \circ \bar{R}$, where it suffices to consider a d-balanced realization $\bar{R}$ of minimal order $r$ (on the boundary of $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$). In this case, we have

$$W_{\mathrm{c},k} = \Lambda = \textstyle\sum_{i=0}^{k-1} \bar{A}^i\bar{B}\bar{B}^\top\bar{A}^{i\top}, \tag{14a}$$
$$W_{\mathrm{o},k} = \Lambda = \textstyle\sum_{i=0}^{k-1} \bar{A}^{i\top}\bar{C}^\top\bar{C}\bar{A}^{i}, \tag{14b}$$

where $\Lambda = \left[\begin{smallmatrix}\Lambda_1 & 0 \\ 0 & 0\end{smallmatrix}\right]$ and the $r \times r$ diagonal matrix $\Lambda_1$ is positive definite. For every $i$, in the $i^{th}$ term inside both the summations the $(2,2)^{th}$ block must be zero. Thus we have $\bar{B}_{21} = 0$ and $\bar{C}_{12} = 0$. Now, considering the term for $i = 1$ in the summation in (14a), we must have $\bar{A}_{21}\bar{B}_{11} = 0$. This implies that for $i = 1$ the summand term is $\left[\begin{smallmatrix}\bar{A}_{11}\bar{B}_{11}\bar{B}_{11}^\top\bar{A}_{11}^\top & 0 \\ 0 & 0\end{smallmatrix}\right]$. More generally, for $i$ $(1 \le i \le k-1)$ we have $\bar{A}_{21}\bar{A}_{11}^{i-1}\bar{B}_{11} = 0$ and the summand term is $\left[\begin{smallmatrix}\bar{A}_{11}^i\bar{B}_{11}\bar{B}_{11}^\top(\bar{A}_{11}^\top)^i & 0 \\ 0 & 0\end{smallmatrix}\right]$ for $1 \le i \le k-1$ $(k \ge n)$. Thus $\bar{A}_{21}[\bar{B}_{11}|\bar{A}_{11}\bar{B}_{11}|\dots|\bar{A}_{11}^{k-2}\bar{B}_{11}] = 0$ and the $k-1^{th}$ order controllability Gramian of the top sub-realization $\bar{R}_{11}$ is $\Lambda_1$, hence it is full-rank (note that $k-1 \ge r$). It follows that in the above equation $\bar{A}_{21} = 0$. Similarly, from (14b) we get $\bar{A}_{12} = 0$. This completes the proof. ∎

Contrary to Proposition 3, this proposition is silent on the explicit form of the $\bar{A}_{22}$-matrix of $\bar{R}$. However, we conjecture that a similar result as in the a.s. case also should hold.

Next, we extend the definition of balanced realizations to *non-minimal* realizations and define the *elevated boundary* of the manifold of balanced minimal realizations:

*Definition 5 (Balanced Realizations):* We call $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,bl},k}_{m,n,p}$ and $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,a,bl}}_{m,n,p}$ defined as

$$\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,bl},k}_{m,n,p} \triangleq \{R \in \widetilde{\mathcal{L}}_{m,n,p} \mid W_{\mathrm{o},k} = W_{\mathrm{c},k} \succeq 0\}, \tag{15}$$

and

$$\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,a,bl}}_{m,n,p} \triangleq \{R \in \widetilde{\mathcal{L}}^{\mathrm{a}}_{m,n,p} \mid W_{\mathrm{o}} = W_{\mathrm{c}} \succeq 0\}, \tag{16}$$

respectively, the sets of $k$-balanced realizations and balanced a.s. realizations of order $n$ and size $(p,m)$. Alternatively, we call boundary of $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,bl},k}_{m,n,p}$ as the *elevated boundary* of $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$ (and $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,bl},k}_{m,n,p}$ itself as the *elevated closure* of $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$). A similar terminology applies to the case of $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,a,bl}}_{m,n,p}$ and $\widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,n,p}$.

Non-minimal balanced realizations form the boundaries of $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,a,bl}}_{m,n,p}$ and $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,bl},k}_{m,n,p}$. We leave the proof of the next proposition as an exercise:

*Proposition 6:* A realization $R \in \widetilde{\mathcal{L}}_{m,n,p}$ of minimal order $r < n$ belongs to the boundary of $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,bl},k}_{m,n,p}$ (resp. $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,a,bl}}_{m,n,p}$) iff it can be written as $R = Q \circ \bar{R}$, where $R$ is as in (9) with $\bar{R}_{11} \in \widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,r,p}$ (resp. $\bar{R}_{11} \in \widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,r,p}$) and $\bar{A}_{22}$ arbitrary (resp. a.s. but otherwise arbitrary).

In comparing $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,bl},k}_{m,n,p}$ with $\overline{\widetilde{\mathcal{O}\Sigma}}^{\text{min,bl},k}_{m,n,p}$ (the closure of $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$), we see that both have the same interiors (namely $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$), but different boundaries. However, the two boundaries can be related:

*Proposition 7:* Any balanced realization $R$ of minimal $r$ on the elevated boundary of $\widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,n,p}$ (resp. $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$) is of the form $R = PQ \circ \bar{R}$, where $Q \in O(n)$, $P = \left[\begin{smallmatrix} I_r & 0 \\ 0 & P_{22}\end{smallmatrix}\right] \in GL(n)$, and $\bar{R}$ is a balanced Kalman standard realization of minimal order $r$ on the boundary of $\widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,n,p}$ (resp. $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$).

*Proof:* For $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,a,bl}}_{m,n,p}$ and $\widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,n,p}$ the proof follows from Propositions 3 and 6: Since the $\bar{A}_{22}$-matrix of $\bar{R}$ has the form as if coming from a balanced realization in $\widetilde{\Sigma}^{\text{min,a}}_{m,n-r,p}$, then as $\bar{A}_{22}$ and $P_{22} \in GL(n-r)$ vary, $P_{22}^{-1}\bar{A}_{22}P_{22}$ generates all a.s. $n-r \times n-r$ matrices. The proof for $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,bl},k}_{m,n,p}$ and $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$ is indirect and relies on Propositions 9 and 10 which relate to s-balanced realization defined in the next subsection: By Proposition 9, $R$ is s-balanced, and by Proposition 10, it belongs to the $GL(n)$-orbit of a balanced realizations of minimal order $r$ on the boundary of $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}^{\text{min,bl},k}_{m,n,p}$, which can be assumed to be a d-balanced Kalman standard realization $\bar{R}$ with $\bar{W}_{\mathrm{o},k} = \bar{W}_{\mathrm{c},k} = \left[\begin{smallmatrix}\Lambda_1 & 0\\ 0 & 0\end{smallmatrix}\right]$ with $\Lambda_1 \succ 0$. If $\bar{P}\circ\bar{R}$ ($\bar{P}\in GL(n)$) is to be balanced, then $\bar{P}^\top \bar{W}_{\mathrm{o},k}\bar{P} = \bar{P}^{-1}\bar{W}_{\mathrm{c},k}\bar{P}^{-\top}$, which implies that $\bar{P} = \left[\begin{smallmatrix}\bar{P}_{11} & 0\\ 0 & \bar{P}_{22}\end{smallmatrix}\right]$ with $\bar{P}_{11}\in O(n)$. ■

Thus, in calling $\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,bl},k}_{m,n,p}$ the "elevated closure" of $\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,bl},k}_{m,n,p}$, "elevated" means that $PQ$ belongs to a *larger* subset of $GL(n)$ than $O(n)$ (cf. Propositions 3 and 4).

*B. S-Balanced Realizations and the Helmke-Moore Lemma*

We start by defining s-balanced realizations:

*Definition 8:* (S-balanced Realizations) We call a realization $R\in\widetilde{\mathcal{L}}_{m,n,p}$ a subspace-balanced (s-balanced) realization of order $n$ and minimal order $r$ if its Kalman canonical realization (structure) is comprised of only a minimal realization of order $r\le n$ and an uncontrollable-and-unobservable realization of order $n-r$, i.e., it can be written as $R = P\circ\bar{R}$, where $\bar{R} = (\left[\begin{smallmatrix}\bar{A}_{11} & 0\\ 0 & \bar{A}_{22}\end{smallmatrix}\right], [\,\bar{B}_{11}\; 0\,], \left[\begin{smallmatrix}\bar{C}_{11}\\ 0\end{smallmatrix}\right])$ and $\bar{R}_{11} = (\bar{A}_{11},\bar{B}_{11},\bar{C}_{11})$ is minimal of order $r$. For such an $R$, we call the system realized by $R$ a balanced system of order $n$ and minimal order $r$. We denote the set of all (resp. a.s.) s-balanced realizations by $\overline{\overline{\Sigma}}^{\mathrm{min}}_{m,n,p}$ (resp. $\overline{\overline{\Sigma}}^{\mathrm{min,a}}_{m,n,p}$). The respective system spaces will be denoted by $\overline{\Sigma}^{\mathrm{min}}_{m,n,p} \triangleq \overline{\overline{\Sigma}}^{\mathrm{min}}_{m,n,p}/GL(n)$ and $\overline{\Sigma}^{\mathrm{min,a}}_{m,n,p} \triangleq \overline{\overline{\Sigma}}^{\mathrm{min,a}}_{m,n,p}/GL(n)$. We denote either of these realization-system pairs by $(\overline{\overline{\Sigma}}_{m,m,p}, \overline{\Sigma}_{m,n,p})$.

It can be shown that $R$ being s-balanced is equivalent to requiring its controllability and observability Gramians of some order $k$ ($n\le k\le\infty$), $W_{\mathrm{c},k}$ and $W_{\mathrm{o},k}$, be of rank $r\le n$ and $\mathcal{N}(W_{\mathrm{o},k})\cap\mathcal{R}(W_{\mathrm{c},k}) = \{0\}$, where $\mathcal{R}(X)$ and $\mathcal{N}(X)$, respectively, denote the range and null spaces of matrix $X$. This is where the name "subspace-balanced" is taken from. Any balanced or $k$-balanced realization—minimal or non-minimal—is s-balanced, but controllable or observable realizations may not be. It is easy to verify that:

*Proposition 9:* A realization $R\in\widetilde{\mathcal{L}}_{m,n,p}$ belongs to the $GL(n)$-orbit of a balanced realization iff it is s-balanced.

The next proposition relates the boundary of the manifolds of balanced minimal realizations to s-balanced realizations.

*Proposition 10:* Every s-balanced realization $\bar{R}\in\widetilde{\mathcal{L}}_{m,n,p}$ (resp. $\bar{R}\in\widetilde{\mathcal{L}}^{\mathrm{a}}_{m,n,p}$) of minimal order $r<n$ can be written as $\bar{R} = P\circ\bar{R}^{\mathrm{bl},k}$ (resp. $\bar{R} = P\circ\bar{R}^{\mathrm{bl}}$) for some $P\in GL(n)$ and $\bar{R}^{\mathrm{bl},k}$ (resp. $\bar{R}^{\mathrm{bl}}$) at the boundary of $\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,bl},k}_{m,n,p}$ (resp. $\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,a,bl}}_{m,n,p}$) and of minimal order $r$ ($n\le k<\infty$).

*Proof:* We only prove the case $\bar{R}\in\widetilde{\mathcal{L}}_{m,n,p}$, as the other case is similar. It suffices to prove the claim for when $\bar{R}$ is a Kalman standard realization, i.e., $\bar{R} = (\left[\begin{smallmatrix}\bar{A}_{11} & 0\\ 0 & \bar{A}_{22}\end{smallmatrix}\right], \left[\begin{smallmatrix}\bar{B}_{11}\\ 0\end{smallmatrix}\right], [\,\bar{C}_{11}\; 0\,])$, where $\bar{R}_{11} = (\bar{A}_{11},\bar{B}_{11},\bar{C}_{11})$ is minimal of order $r$. Next, consider a sequence $\{R_\varepsilon\}_\varepsilon$ ($\varepsilon>0$) with $\lim_{\varepsilon\to0}R_\varepsilon = \bar{R}$ in $\overline{\overline{\Sigma}}^{\mathrm{min}}_{m,n,p}$ defined as $R_\varepsilon = (\left[\begin{smallmatrix}\bar{A}_{11} & 0\\ 0 & \bar{A}_{22}\end{smallmatrix}\right], \left[\begin{smallmatrix}\bar{B}_{11}\\ \varepsilon B_{21}\end{smallmatrix}\right], [\,\bar{C}_{11}\;\varepsilon C_{12}\,])$, where $(\bar{A}_{22},B_{21},C_{12})$ is minimal. It is known that $R_\varepsilon$ is minimal, iff $\bar{A}_{11}$ and $\bar{A}_{22}$ have no common eigenvalues [23]. For now, we assume that $\bar{A}_{11}$ and $\bar{A}_{22}$ have no common eigenvalue. The observability Gramian of $R_\varepsilon$ can be written as $W^{(\varepsilon)}_{\mathrm{o},k} = \left[\begin{smallmatrix}o(1) & o(\varepsilon)\\ o(\varepsilon) & o(\varepsilon^2)\end{smallmatrix}\right]$, where $o(\varepsilon)$ is a matrix function of appropriate size and $\lim_{\varepsilon\to0}\|\frac{o(\varepsilon)}{\varepsilon}\|_F = \eta<\infty$. Also $W^{(\varepsilon)}_{\mathrm{c},k}$ has the same form. Consider the unique p.d-balancing $R^{\mathrm{bl},k}_\varepsilon = \sqrt{S_\varepsilon}\circ R_\varepsilon$, where $S_\varepsilon$ solves the balancing equation $S_\varepsilon W^{(\varepsilon)}_{\mathrm{o},k}S_\varepsilon = W^{(\varepsilon)}_{\mathrm{c},k}$ (see Section II-C). $S_\varepsilon$ depends on smoothly $\varepsilon>0$. From the balancing equation, we see that the determinant of $S_\varepsilon$ must be of order $o(1)$. Since $S_\varepsilon$ is symmetric this can only happen if $S_\varepsilon$ is of the form $\left[\begin{smallmatrix}o(1) & o(\varepsilon)\\ o(\varepsilon) & o(1)\end{smallmatrix}\right]$. Since $S_\varepsilon^{-1}$ solves $W^{(\varepsilon)}_{\mathrm{o},k} = S_\varepsilon^{-1}W^{(\varepsilon)}_{\mathrm{c},k}S_\varepsilon^{-1}$, it also has the same form. Hence, $P_\varepsilon = \sqrt{S_\varepsilon}$ and $P_\varepsilon^{-1} = \sqrt{S_\varepsilon^{-1}}$ remain bounded, which means that (possibly after passing to a subsequence) $\lim_{\varepsilon\to0}P_\varepsilon = \bar{P}$ for some $\bar{P}\in GL(n)$. Hence, the limit realization $\bar{R}^{\mathrm{bl},k} = \bar{P}\circ\bar{R}$ is of minimal order $r$ at the boundary of $\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,bl},k}_{m,n,p}$, which proves the claim. If $\bar{A}_{11}$ and $\bar{A}_{22}$ have common eigenvalues, we replace $\bar{A}_{22}$ with $\bar{A}_{22}+\varepsilon I_n$, and the rest of the proof will essentially remain the same. ■

This result justifies the $^{-}$ in $\overline{\overline{\Sigma}}^{\mathrm{min}}_{m,n,p}$, i.e., its every element is related to an element in $\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,bl},k}_{m,n,p}$ by a state-space change of basis. By extending the idea in this proof, next result shows that the $\bar{A}_{22}$ in Propositions 3 and 4 cannot be arbitrary:

*Proposition 11:* For any $R$ in $\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,bl},k}_{m,n,p}$ (resp. $\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,a,bl}}_{m,n,p}$), $P\circ R$ belongs to $\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,bl},k}_{m,n,p}$ (resp. $\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,a,bl}}_{m,n,p}$) iff $P\in O(n)$.

*Proof:* We give the proof only for $\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,bl},k}_{m,n,p}$, and the proof for $\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,a}}_{m,n,p}$ is essentially the same. For $R\in\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,bl},k}_{m,n,p}$ the result follows from the bundle reduction theorem. For $R$ on the boundary of $\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,bl},k}_{m,n,p}$. We continue the proof of Proposition 10. Denote the Gramians of $R^{\mathrm{bl},k}_\varepsilon$ by $W^{\mathrm{bl},(\varepsilon)}_{\mathrm{o},k}$ and $W^{\mathrm{bl},(\varepsilon)}_{\mathrm{c},k}$. We have $W^{\mathrm{bl},(\varepsilon)}_{\mathrm{o},k} = W^{\mathrm{bl},(\varepsilon)}_{\mathrm{c},k}$ and both are again of the form $\left[\begin{smallmatrix}o(1) & o(\varepsilon)\\ o(\varepsilon) & o(\varepsilon^2)\end{smallmatrix}\right]$. Let us consider $P_\varepsilon$ for which $P_\varepsilon^\top W^{\mathrm{bl},(\varepsilon)}_{\mathrm{o},k}P_\varepsilon = P_\varepsilon^{-1}W^{\mathrm{bl},(\varepsilon)}_{\mathrm{c},k}P_\varepsilon^{-\top}$; with $P_\varepsilon P_\varepsilon^\top = S_\varepsilon$ we get the balancing equation $S_\varepsilon W^{\mathrm{bl},(\varepsilon)}_{\mathrm{o},k}S_\varepsilon = W^{\mathrm{bl},(\varepsilon)}_{\mathrm{c},k}$. The key point is that for $\forall\varepsilon>0$, $S_\varepsilon = I_n$ and since $S_\varepsilon$ depends smoothly on $\varepsilon$, in the limit also we have $S_{\varepsilon=0} = I_n$, which means $P_0\in O(n)$. ■

An important consequence (used to prove Theorem 24) is:

*Proposition 12:* $\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,bl},k}_{m,n,p}/O(n)$ (resp. $\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,a,bl}}_{m,n,p}/O(n)$) and $\overline{\Sigma}^{\mathrm{min}}_{m,n,p}$ (resp. $\overline{\Sigma}^{\mathrm{min,a}}_{m,n,p}$) are equal as sets.

*Proof:* A system $M\in\overline{\Sigma}^{\mathrm{min}}_{m,n,p}$ has a realization in $\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,bl},k}_{m,n,p}$, thus $\overline{\Sigma}^{\mathrm{min}}_{m,n,p}\subset\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,bl},k}_{m,n,p}/O(n)$. If there is a realization $R\in\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,bl},k}_{m,n,p}$ and $P\in GL(n)$ for which $P\circ R\in\overline{\overline{\mathcal{O}\Sigma}}^{\mathrm{min,bl},k}_{m,n,p}$, but $P\circ R$ is not in the $O(n)$-orbit of $R$, then the inclusion could be strict. But that cannot happen by Proposition 11. The case of a.s. systems is similar. ■

Now, we state the Helmke-Moore Lemma:

*Lemma 13 (Helmke-Moore):* (i) The $GL(n)$-orbit of $R\in\widetilde{\mathcal{L}}_{m,n,p}$ (resp. $R\in\widetilde{\mathcal{L}}^{\mathrm{a}}_{m,n,p}$), under the similarity action (3), is a closed subset of $\widetilde{\mathcal{L}}_{m,n,p}$ (resp. $\widetilde{\mathcal{L}}^{\mathrm{a}}_{m,n,p}$) iff it is a s-balanced realization with a Kalman canonical structure in which $A_{22}$ is complex diagonalizable; (ii) Let $R = (\left[\begin{smallmatrix}A_{11} & 0\\ 0 & A_{22}\end{smallmatrix}\right], \left[\begin{smallmatrix}B_{11}\\ 0\end{smallmatrix}\right], [\,C_{11}\; 0\,])$ be a s-balanced realization of minimal order $r$ with $A_{22}$ not (complex) diagonalizable, then there exists another s-balanced realization of minimal order $r$, $\bar{R} = (\left[\begin{smallmatrix}\bar{A}_{11} & 0\\ 0 & \bar{A}_{22}\end{smallmatrix}\right], \left[\begin{smallmatrix}\bar{B}_{11}\\ 0\end{smallmatrix}\right], [\,\bar{C}_{11}\; 0\,])$, where $\bar{A}_{22}$ is complex diagonalizable, $P_{11}\circ(A_{11},B_{11},C_{11}) = (\bar{A}_{11},\bar{B}_{11},\bar{C}_{11})$ for some $P_{11}\in GL(r)$, $A_{22}$ and $\bar{A}_{22}$ have the same eigenvalues (but are not similar), and the $GL(n)$-orbit of $\bar{R}$ belongs to the boundary of the $GL(n)$-orbit of $R$; moreover, the boundary of the $GL(n)$-orbit of $R$ is comprised of the $GL(n)$-orbits of such realizations $\bar{R}$.

*Proof:* See [19, p. 212-5]. The proof therein, due to the intended application, is for $R$ being complex-valued, however, the same proof applies to the real case, as well. ∎

For later reference it is useful to define:

*Definition 14:* (Diagonalizable s-balanced realizations) We call a realization $R \in \widetilde{\mathcal{L}}_{m,n,p}$ as in Lemma 13 a diagonalizable s-balanced realization, and the corresponding system a diagonalizable s-balanced system. We denote the space of diagonalizable s-balanced realizations and its a.s. subspace by $\overset{\circ}{\overline{\widetilde{\Sigma}}}{}^{\min}_{m,n,p}$ and $\overset{\circ}{\overline{\widetilde{\Sigma}}}{}^{\min,\mathrm{a}}_{m,n,p}$, respectively. The corresponding system spaces are, respectively, denoted as $\overset{\circ}{\overline{\Sigma}}{}^{\min}_{m,n,p} \triangleq \overset{\circ}{\overline{\widetilde{\Sigma}}}{}^{\min}_{m,n,p}/GL(n)$ and $\overset{\circ}{\overline{\Sigma}}{}^{\min,\mathrm{a}}_{m,n,p} = \overset{\circ}{\overline{\widetilde{\Sigma}}}{}^{\min,\mathrm{a}}_{m,n,p}/GL(n)$. The diagonalizable subsets of $\overline{\Sigma}_{m,n,p}$, $\overline{\widetilde{\Sigma}}_{m,n,p}$, and $\overline{\widetilde{\mathcal{O}\Sigma}}_{m,n,p}$ are denoted by $\overset{\circ}{\overline{\widetilde{\Sigma}}}_{m,n,p}$, $\overset{\circ}{\overline{\Sigma}}_{m,n,p}$, and $\overset{\circ}{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}$, respectively.

Using the Helmke-Moore Lemma, in Theorem 23, we show that diagonalizable s-balanced systems $\overset{\circ}{\overline{\Sigma}}{}^{\min}_{m,n,p}$ and $\overset{\circ}{\overline{\Sigma}}{}^{\min,\mathrm{a}}_{m,n,p}$ form metrizable subspaces of $\mathcal{L}_{m,n,p}$ and $\mathcal{L}^{\mathrm{a}}_{m,n,p}$. The *closedness* of the orbits in these ($GL(n)$-quotient) spaces is the key to their metrizability. In general, although distinct $GL(n)$-orbits do not intersect, the *closure* of an orbit may intersect another orbit. If two distinct $GL(n)$-orbits are already *closed*—as subsets of $\widetilde{\mathcal{L}}_{m,n,p}$ or $\widetilde{\mathcal{L}}^{\mathrm{a}}_{m,n,p}$, then obviously one cannot intersect the closure of the other one. This is the case e.g., for the orbits of minimal realizations and diagonalizable s-balanced realizations. In contrast, in part (ii) of the Lemma, the respective non-diagonalizable s-balanced system $[R]$ and the diagonalizable system $[\bar{R}]$ are not *separable* by distinct open neighborhoods, which means that $\overline{\Sigma}^{\min}_{m,n,p}$ is not Hausdorff. This is explains the source of non-Hausdorffness and non-metrizability of $\widetilde{\mathcal{L}}_{m,n,p}/GL(n)$.

## IV. Model Order Reduction in the Alignment Distance: Problem Formulation

We can formulate model order reduction based on the alignment distance in, at least, two ways, depending on whether we use the boundary or the elevated boundary of the manifold of minimal balanced realizations. The first approach is more mathematically elegant and natural, whereas the second approach, which is natural in its own way, results in a computationally simpler optimization, but is expressible in terms of what we call the "one-sided alignment distance." Both approaches enjoy more or less the same properties (e.g., both will give the same feedback robustness result, see Theorem 30). Ultimately, for practical considerations, Definition 15—the second approach—is our choice for later developments and what we call model order reduction in the alignment distance.

### *A. Exact Extension of the Alignment Distance to $\overset{\circ}{\overline{\Sigma}}_{m,n,p}$*

Ideally, because of Proposition 10, one wants to extend of the alignment distance (7) to the space of s-balanced systems $\overline{\Sigma}_{m,n,p}$. However, $\overline{\Sigma}_{m,n,p}$ itself is non-Hausdorff and non-metrizable; but the set of diagonalizable s-balanced systems, $\overset{\circ}{\overline{\Sigma}}_{m,n,p}$ (recall Definition 14), is metrizable and Hausdorff (see Theorem 23). Take any $M_1, M_2 \in \overset{\circ}{\overline{\Sigma}}_{m,n,p}$ with realizations $R_1, R_2 \in \overset{\circ}{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}$. Here, $\overset{\circ}{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}$ is the closure of $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ with only diagonalizable realizations on its boundary. Now define the (exact) extension of the alignment distance, which we denote by $\bar{d}_F$, as: $\bar{d}_F(M_1, M_2) = \min_Q \tilde{d}_F(Q \circ R_1, R_2)$. We also call $\bar{d}_F$ the extended alignment distance subordinate to $\overset{\circ}{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}$. It follows (e.g., from Theorem 3 in [3]) that $\bar{d}_F$ is a distance on $\overset{\circ}{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}/O(n)$ which matches its natural quotient topology; recall from Proposition 12 that this quotient (as a set) is nothing but $\overset{\circ}{\overline{\Sigma}}_{m,n,p}$. In Theorem 24, we show that this topology matches the natural topology of $\overset{\circ}{\overline{\Sigma}}_{m,n,p}$ (as a $GL(n)$-quotient). Next, we define model order reduction in the *extended* alignment distance as: $\inf_{\bar{M}} \bar{d}_F(M, \bar{M})$, where $\bar{M}$ is a diagonalizable system of minimal $r < n$ in $\overset{\circ}{\overline{\Sigma}}_{m,n,p}$.

### *B. Model Order Reduction in the Alignment Distance*

Consider the quotient (system) spaces $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}/O(n)$ and $\overline{\Sigma}_{m,n,p}$. They match in their interiors $\Sigma_{m,n,p}$. However, on its boundary, $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}/O(n)$ includes *distinct* "systems," as $O(n)$-orbits, which belong to the *same* $GL(n)$-orbit; the reason is essentially the elevated boundary (see Propositions 7 and 12). Having this in mind, we can define the alignment distance on $(\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}/O(n))$ subordinate to $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}$ in an obvious way: $d_F(M_1, M_2) = \min_Q \tilde{d}_F(Q \circ R_1, R_2)$, where $R_i \in \overline{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}$ is a realization of $M_i \in \overline{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}/O(n)$ $(i = 1, 2)$. Although, outside of $\Sigma_{m,n,p}$, $d_F$ is not a distance between $GL(n)$-orbits, we will show that model order reduction based on it can be related to a distance-like measure on the space of systems as $GL(n)$-orbits (specifically $\overset{\circ}{\overline{\Sigma}}_{m,n,p}$):

*Definition 15 (Model Reduction in the Alignment Distance):* Consider $M \in \Sigma_{m,n,p}$ and let $R \in \widetilde{\mathcal{O}\Sigma}_{m,n,p}$ be a balanced realization of $M$. Then the $r^{th}$-order model order reduction in the alignment distance (subordinate to the elevated closure $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}$) is defined as

$$\inf_{\bar{M}} d_F(M, \bar{M}) = \inf_{Q \in O(n), \bar{R}} \tilde{d}_F(Q \circ R, \bar{R}), \qquad (17)$$

where $\bar{R}$ is a balanced Kalman standard realization of minimal order $r$ at the elevated boundary of $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ (see Definitions 2 and 5) and $\bar{M}$ the system realized by $\bar{R}$. We call such a $\bar{R}$ a feasible realization.

This definition is independent of the choice of any specific realization $R \in \widetilde{\mathcal{O}\Sigma}_{m,n,p}$ of $M$. If $\bar{R}$ is a solution to (17), then, as a notational convention and for convenience, we will write the achieved quantity in (17) as $d_F(M, \bar{M})$, where now $\bar{M}$ is the system in $\overline{\Sigma}_{m,n,p}$ realized by $\bar{R}$, i.e., it is a system as a $GL(n)$-orbit. We will be mindful, however, that $d_F(M, \bar{M})$ is calculated for a specific realization of $\bar{M} \in \overline{\Sigma}_{m,n,p}$. A subtle and important point is that since the search for $\bar{M}$ over the entire of $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}/O(n)$, all the $GL(n)$-orbits in $\overline{\Sigma}_{m,n,p}$ will be covered. Definition 15 (or specifically its right hand side) is natural in the following sense: the alignment distance is defined by aligning realizations on the manifold of balanced minimal realizations (i.e., where the controllability and observability Gramians are full rank and equal), this definition is based on extending the notion of alignment to the space of balanced realizations, where the Gramians are equal but can be rank-deficient. Moreover, this definition allows to achieve a *smaller* reduction error, as it is a *relaxed* version of model order reduction in the exact extension of the alignment distance subordinate to $\mathring{\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}}_{m,n,p}$.

*Remark 16 (Analogy with rank-r Matrix Approximation):* At this stage, perhaps it is interesting to recall a formulation of the best rank-$r$ approximation of matrix $X \in \mathbb{R}^{n\times n}$ via the SVD that resembles our model order reduction problem: $\min_{U,V\in O(n),\bar{X}} \|U^\top XV - \bar{X}\|_F$, where $\bar{X}$ is diagonal, of the form $\bar{X} = \left[\begin{smallmatrix}\bar{X}_1 & 0\\ 0 & 0\end{smallmatrix}\right]$ with $\bar{X}_1 \succ 0$ being $r \times r$. Notice that the simplest or canonical structure that can be achieved via the transformation $U^\top XV$ is the diagonal form with positive entries (the SVD theorem). A more commonly used formulation is $\min_{U,V\in O(n),\bar{X}} \|X - U\bar{X}V^\top\|_F$, which thanks to $O(n)$-invariance of $\|\cdot\|_F$ is equivalent to the previous one.

*Remark 17:* Eising [24] has introduced the notion of a "distance to uncontrollability," which it is *not* a distance between *systems*, rather it is designed to measure how far a *specific* realization of a system is from uncontrollability. In fact, since by state-space changes of basis, one can bring a given realization as close to uncontrollability as desired, the notion of "the distance of a system to uncontrollability" is meaningless. There are, however, apparent similarities between our formulation of model order reduction and "distance to uncontrollability." In fact, the Frobenius norm based distance (6) between realizations appears in that context too, and is often referred to as the Eising distance.

The set of balanced Kalman standard realizations of minimal $r$ (i.e., the feasible set for $\bar{R}$ in the minimization problem (17)) is *not* a closed set; thus, although we conjecture otherwise, a-priori a feasible realization $\bar{R}$ achieving the infimum may not exist. However, this is not a severe problem for model order applications, as we have the following proposition:

*Proposition 18 (A-priori existence of solutions):* (i) In the case of $(\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}, \Sigma^{\text{min}}_{m,n,p})$ the problem in (17) always has a solution in the following sense: there exists a balanced Kalman standard realization $\bar{R}$ of minimal order at most $r$ on the elevated boundary of $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$ achieving the infimum; (ii) Likewise, in the case of $(\widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,n,p}, \Sigma^{\text{min,a}}_{m,n,p})$ a realization achieving the infimum always exists in the same sense except that such a realization is either asymptotically stable or it is only stable (i.e., it has pole(s) on the unit circle).

*Proof:* For (i), note that the set of balanced Kalman standard realizations of minimal order not larger than $r$ is the closure of the set of balanced Kalman standard realizations of minimal order $r$ in $\widetilde{\mathcal{L}}_{m,n,p}$ (see Definition 2 and Proposition 6). Also note that if $\bar{R} = (\bar{A}, \bar{B}, \bar{C})$ is a feasible realization for which one of $\|\bar{A}\|_F$, $\|\bar{B}\|_F$, or $\|\bar{C}\|_F$ is larger than $2(\|A\|_F + \|B\|_F + \|C\|_F)$, then $\tilde{d}_F(R, \bar{R}) > \tilde{d}_F(R, 0)$. This means that the closure of the feasible set for the minimization problem (17) can be considered as a bounded, hence a compact set ($O(n)$ is compact). Thus a solution achieving the infimum exits in the sense stated. For (ii) the situation is similar in terms of boundedness of the feasible set; but in this case the closure of the set of feasible realizations in $\widetilde{\mathcal{L}}_{m,n,p}$ includes balanced Kalman standard realizations with poles on and inside the unit circle. The statement is a consequence of this fact. ∎

Although (17) is not expressed in terms of a distance function in $\mathring{\overline{\Sigma}}_{m,n,p}$ (i.e., systems as $GL(n)$-orbits), a slightly different form of "alignment distance" can be useful for that:

*Definition 19 (One-sided Alignment Distance):* Let $M_1, M_2 \in \mathring{\overline{\Sigma}}_{m,n,p}$. We call

$$\overleftarrow{d}_F(M_1,M_2) = \begin{cases} \min\limits_{Q,P} \tilde{d}_F(Q\circ R_1, P\circ R_2) & M_2 \text{ of minimal order } r<n,\\ \min\limits_{Q} \tilde{d}_F(Q\circ R_1, R_2) & \text{otherwise}\end{cases} \tag{18}$$

the one-sided alignment distance between $M_1$ and $M_2$ subordinate to the closure $\overline{\widetilde{\mathcal{O}\Sigma}}_{m,n,p}$. In the above, $R_1$ is a balanced Kalman standard realization of $M_1$ in $\overline{\widetilde{\mathcal{O}\Sigma}}_{m,n,p}$, $R_2$ is a balanced Kalman standard realization of $M_2$ in $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}$, $Q \in O(n)$, and $P = \left[\begin{smallmatrix} I_r & 0\\ 0 & P_{22}\end{smallmatrix}\right] \in GL(n)$.

Clearly, in $\overleftarrow{d}_F(\cdot,\cdot)$ the order of arguments matters. Note that the right-hand side of (18), indeed, depends only $M_1$ and $M_2$, and not a specific choice of the realizations. However, while $R_2$ can be in the elevated closure, $R_1$ must be in the closure. Since the $GL(n)$-orbit of $R_2 \in \widetilde{\mathcal{L}}_{m,n,p}$ is closed (by Lemma 13), as $P$ or $P^{-1}$ becomes unbounded, $P \circ R$ must become unbounded; hence "min" is used in (18). Although $\overleftarrow{d}_F$ is *not* a distance on the entire of $\mathring{\overline{\Sigma}}_{m,n,p}$ we have:

*Proposition 20:* The following hold on $\mathring{\overline{\Sigma}}_{m,n,p}$: (i) $\overleftarrow{d}_F$ matches the alignment distance $d_F$ (7) on $\Sigma_{m,n,p}$; (ii) $\overleftarrow{d}_F(M_1, M_2) \geq 0$ with equality iff $M_1 = M_2$; (iii) If $M \in \Sigma_{m,n,p}$ is a third (minimal) system, then the triangle inequality holds: $\overleftarrow{d}_F(M_1,M_2) \leq \overleftarrow{d}_F(M_1,M) + \overleftarrow{d}_F(M,M_2)$; (iv) $M \mapsto \overleftarrow{d}_F(M, M_0)$ is continuous for fixed $M_0 \in \mathring{\overline{\Sigma}}_{m,n,p}$.

*Proof:* (i) is obvious. (ii) for minimal systems follows from (i), for $M_1 \in \mathring{\overline{\Sigma}}_{m,n,p}$ of minimal order $r < n$, let $R_1, R_1'$ be two balanced Kalman realizations, by Proposition 7, there are $Q \in O(n)$ and $P = \left[\begin{smallmatrix} I_r & 0\\ 0 & P_{22}\end{smallmatrix}\right] \in GL(n)$ such that $\tilde{d}_F(Q \circ R_1, P \circ R_1') = 0$. Conversely, let $\inf_{Q,P} \tilde{d}_F(Q\circ R_1, P\circ R_2) = 0$ for diagonalizable balanced Kalman standard realizations $R_1$ and $R_2$. By Lemma 13, $GL(n) \circ R_2$ is closed hence, $\tilde{d}_F(Q' \circ R_1, P' \circ R_2) = 0$ for a $Q \in O(n)$ and $P' \in GL(n)$, i.e., $R_1$ and $R_2$ realize the same system. (iii) follows from $O(n)$-invariance of $\tilde{d}_F$. (iv) follows from standard properties of quotient maps and that the minimum of a function continuously depending on a parameter is continuous. ∎

Now, we can formulate a model order reduction problem as

$$\inf\nolimits_{\bar{M}} \overleftarrow{d}_F(M, \bar{M}), \tag{19}$$

where $\bar{M}$ is a system of order $r < n$ in $\mathring{\overline{\Sigma}}_{m,n,p}$. Our original problem (17) is a slight—but practical—relaxation of this problem, because in (17) no diagonalizablity constraint is imposed on the $\bar{A}_{22}$-matrix of $\bar{R}$. Since diagonalizable $\bar{A}_{22}$-matrices are dense in $\mathbb{R}^{n-r\times n-r}$, solving (17) is indeed more natural, and does not affect the achievable minimum cost.

### C. More on the Existence of Asymptotically Stable Solutions

Proposition 18 does not rule out the possibility that an a.s. system may have a reduced order approximation that is only stable (i.e., may have poles on the unit circle). The source of this is that the optimization involved is on an open set. As far as issues such as robustness of stability under feedback are concerned, the requirement that the reduced order approximation of an a.s. must be a.s. is not essential. In most input-output based model reduction formulations, if the original system is a.s., the solution will be a.s.; the reason is that the cost function involved (i.e., a norm of the difference of system impulse response and the solution impulse response) will blow up if the solution is not a.s. As seen next, in certain special cases, we can improve Proposition 18. Further improvements may be possible, but they seem challenging.

*Proposition 21:* If $R = (A, B, C) \in \widetilde{\mathcal{O}\Sigma}^{\text{min,a}}_{m,n,p}$ and $A$ is symmetric, then any solution $\bar{R}$ to (17) is a.s.

*Proof:* First we note that for any $R \in \widetilde{\mathcal{O}\Sigma}^{\text{min,a}}_{m,n,p}$ with $A$ symmetric the top (or bottom) sub-realization is balanced and a.s. To see this let $R_{11} = (A_{11}, B_{11}, C_{11})$ be the top sub-realization of order $r$. Since $A$ and $A_{11}$ both are symmetric their spectral radii is equal to their 2-norms. Thus we have $\|A_{11}\|_2 \leq \|A\|_2 < 1$. The same holds for $A_{22}$. Also note that from blandness (see (2)) we must have $BB^\top = C^\top C$ and in particular $B_{11}B_{11}^\top = C_{11}^\top C_{11}$. This implies that $R_{11} = (A_{11}, B_{11}, C_{11})$ is balanced. To see the main result let $Q \in O(n)$ and $\bar{R}$ (with possibly poles on the unit circle) solve (17). Then by applying the above results to $Q \circ R$ which is balanced and a.s., we see that for $\bar{R}$ we must have $\bar{R}_{11} = (Q \circ R)_{11}$ and $\bar{A}_{22} = (Q^\top AQ)_{22}$; thus $\bar{R}$ is a.s. ∎

Another example is when the target system is of order $r = 1$ (for a proof see Appendix A):

*Proposition 22:* If $r = 1$ and $R \in \widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,n,p}$, then any $\bar{R} = (\bar{A}, \bar{B}, \bar{C})$ achieving the infimum in (17) is a.s.

## V. More on the Metrization of the Kalman Decomposition

In this section, we briefly study the extension of the noncompact alignment distance to $\mathring{\bar{\Sigma}}_{m,n,p}$ (Theorem 23), and the metrization of the (standard) Kalman decomposition for model order reduction. Results in this section are primarily of theoretical interest.

### A. Extension of the Noncompact Alignment Distance

The manifold $\tilde{\Sigma}_{m,n,p}$ can be equipped with a $GL(n)$-invariant Riemannian metric $\tilde{g}_R$ and the corresponding distance $\tilde{d}$ (see [3, Section VI. B] for examples). Notice again that we mean distances consistent with the natural topology induced from the Euclidean topology of $\tilde{\mathcal{L}}_{m,n,p}$. We assume that $\tilde{d}$ is an *incomplete* or a *finite-escape-time* distance, by which we mean that given any $\bar{R}$ at the boundary of $\tilde{\Sigma}_{m,n,p}$ and $\{R_i\}_i$ a sequence of interior points converging to $\bar{R}$—in the ambient distance, i.e., $\tilde{d}_F(R_i, \bar{R}) \to 0$, then for any interior point $R \in \tilde{\Sigma}_{m,n,p}$ (which can be connected to the $R_i$'s) the sequence $\{\tilde{d}(R, R_i)\}_i$ remains bounded. It can be seen that if the associated Riemannian metric $\tilde{g}_R$ is such that $\tilde{g}_{R_i}$ remains bounded (as a quadratic form) as $R_i$ approaches $\bar{R}$, then $\tilde{d}$ will be incomplete. The Krishnaprasad-Martin distance $\tilde{d}^{\text{KM}}$ is an example of such a distance (see [25] and [3, Section VI. B]). Next, we note that $\tilde{\Sigma}_{m,n,p}$ is an open subset of the space of s-balanced realizations $\bar{\tilde{\Sigma}}_{m,n,p}$ to which $\tilde{d}$ can be extended by the next standard procedure: Every non-minimal s-balanced realization $\bar{R}$ can be considered as the limit of a sequence of minimal realizations $\{R_i\}_i$ converging in the natural ambient Euclidean distance $\tilde{d}_F$. It is easy to see that by setting $\bar{\tilde{d}}(R_1, R_2) = \lim_i \tilde{d}(R_1^{(i)}, R_2^{(i)})$, where $\{R_j^{(i)}\}_i$ ($j = 1, 2$) is a sequence in $\tilde{\Sigma}_{m,n,p}$ converging to $R_j \in \bar{\tilde{\Sigma}}_{m,n,p}$, $(\bar{\tilde{\Sigma}}_{m,n,p}, \bar{\tilde{d}})$ becomes a metric space with distance $\bar{\tilde{d}}$ being $GL(n)$-invariant and matching the Euclidean topology of $\bar{\tilde{\Sigma}}_{m,n,p}$.

Now we turn to the extension of the noncompact alignment distance from $\Sigma_{m,n,p}$ to $\bar{\Sigma}_{m,n,p}$. Starting with, $\bar{\tilde{d}}$ in order to have a group action induced *positive-definite* distance $\bar{d}$, one needs to have the orbits being a closed subset of $\bar{\tilde{\Sigma}}_{m,n,p}$. Since otherwise, for two realizations $R_1, R_2$ belonging to two distinct $GL(n)$-orbits, the closure of the orbit $GL(n) \circ R_1$ intersects the orbit $GL(n) \circ R_2$, and we get $\inf_{P \in GL(n)} \bar{\tilde{d}}(P \circ R_1, R_2) = 0$, i.e., $\bar{d}([R_1], [R_2]) = 0$. Thus we need to pass to the subspace of diagonalizable s-balanced realizations $\mathring{\bar{\tilde{\Sigma}}}_{m,n,p}$. This leads to the extension of the noncompact alignment distance to the set of diagonalizable s-balanced systems $\mathring{\bar{\Sigma}}_{m,n,p}$:

$$\bar{d}(M_1, M_2) = \inf_{P \in GL(n)} \bar{\tilde{d}}(P \circ R_1, R_2), \qquad (20)$$

where $R_i \in \mathring{\bar{\tilde{\Sigma}}}_{m,n,p}$ is any realization of $M_i \in \mathring{\bar{\Sigma}}_{m,n,p}$ ($i = 1, 2$). See [3, Theorem 3] for a proof that $\bar{d}$ is a distance. It also follows from [3, Theorem 3] that the extended noncompact alignment distance induces the natural quotient topology, hence we have:

*Theorem 23:* The spaces of diagonalizable s-balanced systems $\mathring{\bar{\Sigma}}^{\text{min}}_{m,n,p}$ and $\mathring{\bar{\Sigma}}^{\text{min,a}}_{m,n,p}$ are metrizable (topological) subspaces of $\mathcal{L}_{m,n,p}$.

The next theorem generalizes the theory of the alignment distance to $\mathring{\bar{\Sigma}}_{m,n,p}$ in terms of the induced topology (cf. [3, Theorem 19]).

*Theorem 24:* The system spaces $\mathring{\bar{\Sigma}}^{\text{min}}_{m,n,p}$ (resp. $\mathring{\bar{\Sigma}}^{\text{min,a}}_{m,n,p}$) and $\overline{\widetilde{\mathcal{O}\mathring{\Sigma}}}^{\text{min,bl},k}_{m,n,p}/O(n)$ (resp. $\overline{\widetilde{\mathcal{O}\mathring{\Sigma}}}^{\text{min,a}}_{m,n,p}/O(n)$) are homeomorphic. In particular, the extended alignment distance $\bar{d}_F$ (defined in Section IV-A) induces the natural quotient topology on $\mathring{\bar{\Sigma}}^{\text{min}}_{m,n,p}$ and $\mathring{\bar{\Sigma}}^{\text{min,a}}_{m,n,p}$.

*Proof:* Since both cases are similar, we only prove the claim for $\mathring{\bar{\Sigma}}^{\text{min}}_{m,n,p}$ and $\overline{\widetilde{\mathcal{O}\mathring{\Sigma}}}^{\text{min,bl},k}_{m,n,p}/O(n)$. The fact the two spaces are equal as sets follows from Proposition 12. We need to show that natural quotient topologies are the same. $\overline{\widetilde{\mathcal{O}\mathring{\Sigma}}}^{\text{min,bl},k}_{m,n,p}/O(n)$ becomes a metric space using the extension of alignment distance $\bar{d}_F$ (associated with $\tilde{d}_F$) and $\mathring{\bar{\Sigma}}^{\text{min}}_{m,n,p}$ also is metrized using the non-compact alignment distance $\bar{d}$. $\tilde{d}_F$ and $\bar{\tilde{d}}$ induce the same topology on $\overline{\widetilde{\mathcal{O}\mathring{\Sigma}}}^{\text{min,bl},k}_{m,n,p} \subset \mathring{\bar{\tilde{\Sigma}}}^{\text{min}}_{m,n,p}$, and the respective projection maps of $\mathring{\bar{\tilde{\Sigma}}}^{\text{min}}_{m,n,p}$ and $\overline{\widetilde{\mathcal{O}\mathring{\Sigma}}}^{\text{min,bl},k}_{m,n,p}/O(n)$ map open balls to open balls of the same radii (see proof of [3, Theorem 3]). It follows that any open ball in one topology contains an open ball in the other topology. Thus the topologies match. ∎

Finally, we turn to model order reduction and metrization of the standard Kalman decomposition using the noncompact alignment distance:

*Definition 25:* Let $M \in \Sigma_{m,n,p}$ and $\bar{d}$ be the noncompact alignment distance associated with the extension of $GL(n)$-invariant $\tilde{d}$ to $\bar{\bar{\Sigma}}_{m,n,p}$. Then we define the $r^{th}$ order model reduction in the non-compact alignment distance as

$$\inf_{\bar{M}} \bar{d}(M, \bar{M}) = \inf_{P \in GL(n), \bar{R}} \bar{\tilde{d}}(P \circ R, \bar{R}) \quad (21)$$

where $\bar{M} \in \mathring{\bar{\bar{\Sigma}}}_{m,n,p}$ is a diagonalizable s-balanced system of minimal order $r < n$ at the boundary of $\Sigma_{m,n,p}$ and $\bar{R}$ is a balanced Kalman standard realization of $\bar{M}$ on the boundary of $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ of minimal order $r$.

In theory, $\bar{R}$ can be any minimal order $r$ realization and does not need to be a balanced realization on the boundary of $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$. However, since $\tilde{d}$ is $GL(n)$-invariant we need to fix the allowable change of coordinates in $\bar{R}$ to only orthogonal changes, to avoid potential unbounded solutions. Further study of problem (21) is left to a later work. We also mention that computationally building $\bar{\tilde{d}}$ can be very challenging, let alone solving problem (21).

### B. Non-Distance Based Model Order Reduction

The basic idea of comparing a realization under change of coordinates with another realization does not need a $GL(n)$-invariant distance. For example, if we consider $\tilde{d}_F$ in (6), then for any systems $M_1, M_2 \in \mathring{\tilde{\Sigma}}_{m,n,p}$ with realizations $R_1, R_2 \in \mathring{\bar{\bar{\Sigma}}}_{m,n,p}$ we can define a *divergence* as $D_F(M_1, R_1) = \min_{P \in GL(n)} \tilde{d}_F(P \circ R_1, R_2)$. This definition is not independent of the choice of $R_1$ unless we restrict $R_1$ to balanced realizations $\overline{\overline{\widetilde{\mathcal{O}\Sigma}}}_{m,n,p}$ (due to $O(n)$-invariance of $\tilde{d}_F$). We can show that $D_F(M_1, R_2) \geq 0$ with equality only if $R_1$ is a realization of $M_1$ (this follows from closedness of orbits). Now, model order reduction can be defined as

$$\inf_{P \in GL(n), \bar{R}} \tilde{d}_F(P \circ R, \bar{R}), \quad (22)$$

where $R \in \widetilde{\Sigma}_{m,n,p}$ is a (minimal) realization of $M \in \Sigma_{m,n,p}$ and $\bar{R}$ is a balanced Kalman standard realization of minimal order $r$ on the boundary of $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ or its elevated boundary. Interestingly, rather similar formulations have appeared in the context of grey-box system identification [26], [27]. Algorithmically, solving (22) is not much different from our problem (17), when $\bar{R}$ is on the elevated boundary of $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$.

## VI. An Alternating Minimization Algorithm: Align, Truncate, & Project (ATP)

In this section, we derive an efficient algorithm for solving the model order reduction problem (17) using alternating minimization between $Q$ and $\bar{R}$. The algorithm is called Align, Truncate, Project (ATP). Let $Q$ and $\bar{R} = \left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & \bar{A}_{22} \end{bmatrix}, \begin{bmatrix} \bar{B}_{11} & 0 \end{bmatrix}, \begin{bmatrix} \bar{C}_{11} \\ 0 \end{bmatrix}\right) \in \overline{\overline{\widetilde{\mathcal{O}\Sigma}}}{}^{\text{min,bl},k}_{m,n,p}$ solve (17). If $Q$ is fixed in $\min_{\bar{R}} \tilde{d}_F(Q \circ R, \bar{R})$, where $\bar{R}$ is in the form Proposition 6, then we must have $\bar{A}_{22} = (Q^\top A Q)_{22}$. This is an important *decoupling* that happens thanks to the use of the elevated boundary in Definition 15 as opposed to the (actual) boundary, and the reason for our choice of Definition 15. Now, the top sub-realization of $\bar{R}$, $\bar{R}_{11} \in \widetilde{\mathcal{O}\Sigma}{}^{\text{min,bl},k}_{m,r,p}$, solves

$$\min_{\bar{R}_{11} \in \widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,r,p}} \tilde{d}_F((Q \circ R)_{11}, \bar{R}_{11}). \quad (23)$$

The above problem is *projection* (in the (Euclidean) $\tilde{d}_F$ distance) of the truncated part or top sub-realization of $Q \circ R$ (namely, $(Q \circ R)_{11}$) onto the space $\widetilde{\mathcal{O}\Sigma}{}^{\text{min,bl},k}_{m,r,p}$. The iterative ATP algorithm now is:

1: **function** APT$(M, r)$
2: `choose a balanced realization` $R \in \widetilde{\mathcal{O}\Sigma}{}^{\text{min,bl},k}_{m,n,p}$ `of` $M$ `and an initial guess` $\bar{R}^0 \in \widetilde{\mathcal{O}\Sigma}{}^{\text{min,bl},k}_{m,n,p}$ `of minimal order` $r$
3: **repeat**
4: `solve` $\min_Q \tilde{d}_F^2(Q \circ R, \bar{R}^i)$ `to find` $Q^i$ ▷ Alignment Step
5: *`truncate`* $Q^i \circ R$ `to get` $\hat{\bar{R}}^{i+1}_{11} = (Q^i \circ R)_{11}$ `and` $\bar{A}^{i+1}_{22} = (Q^{i\top} A Q^i)_{22}$ `in` $\bar{R}^{i+1}$ ▷ Truncation Step
6: *`project`* $\hat{\bar{R}}^{i+1}_{11}$ `onto` $\widetilde{\mathcal{O}\Sigma}{}^{\text{min,bl},k}_{m,r,p}$ `to get` $\bar{R}^{i+1}_{11}$ `in` $\bar{R}^{i+1}$ ▷ Projection Step
7: **until** `convergence`
8: **return** $\bar{R}^{i+1}$
9: **end function**

Notice that in step 2, $R$ can be a sorted d-balanced realization of $M$ and $\bar{R}^0$ can be its $r^{th}$ order truncation. Such a realization is of minimal $r$, generically; more specifically, if $\lambda_r > \lambda_{r+1}$ the strong sub-realization of order $r$ (i.e., the truncated realization) is minimal of order $r$, but if $\lambda_r = \lambda_{r+1}$ it may be non-minimal [28]. However, even in this case it can be made minimal by an orthogonal change basis.

The main computational challenge is the step of projection (23). In discrete-time, sub-realizations of a balanced realization (even if d-balanced) are, in general, are not balanced. In certain cases, however, no projection might be needed, e.g., if the $A$-matrix in the balanced realization $R$ is symmetric (recall Proposition 21). In practice, the first step of alignment might give a good enough approximate solution (see Section IX and Figure 2); or as an approximation one might simply re-balance the truncated realization using p.d-balancing (see Section II-C).

### A. Projection via Riemannian Gradient Descent

The manifolds $\widetilde{\mathcal{O}\Sigma}{}^{\text{min,bl},k}_{m,r,p}$ and $\widetilde{\mathcal{O}\Sigma}{}^{\text{min,a,bl}}_{m,r,p}$ are submanifolds of $\mathcal{L}_{m,r,p}$ (which is the same as $\mathbb{R}^{r^2+rm+rp}$). We equip these manifolds with the natural Riemannian metric induced from the ambient Euclidean space. Given a realization $\hat{R} \in \widetilde{\mathcal{L}}^{\text{min}}_{m,n,p}$, a projection of $\hat{R}$ on $\widetilde{\mathcal{O}\Sigma}{}^{\text{min,bl},k}_{m,r,p}$ is defined as a minimizer of the function $f : \widetilde{\mathcal{O}\Sigma}{}^{\text{min,bl},k}_{m,r,p} \to \mathbb{R}$

$$f(R; \hat{R}) = \tilde{d}_F^2(R, \hat{R}), \quad (24)$$

i.e., a point on $\widetilde{\mathcal{O}\Sigma}{}^{\text{min,bl},k}_{m,r,p}$ which is closest to $\hat{R}$ measured in Euclidean distance (see (6)). A similar definition applies to $\widetilde{\mathcal{O}\Sigma}{}^{\text{min,a,bl}}_{m,r,p}$ with $\hat{R} \in \widetilde{\mathcal{L}}^{\text{min,a}}_{m,n,p}$. In the sequel, unless otherwise stated we assume that $\widetilde{\mathcal{O}\Sigma}{}^{\text{min,bl},k}_{m,r,p}$ includes the case $k = \infty$, and thus we treat both cases almost similarly. Notice that, in general, $\hat{R}$ need not be minimal to define its projection on $\widetilde{\mathcal{O}\Sigma}{}^{\text{min,bl},k}_{m,r,p}$; however, we make this assumption since it brings about certain important benefits, which will become clear shortly. This is not a major limitation because minimal realizations are generic (or dense) in $\widetilde{\mathcal{L}}_{m,r,p}$.

To describe the algorithm it is useful to introduce the operator $\text{vec}(X)$ which stacks the columns of matrix $X \in \mathbb{R}^{m \times n}$

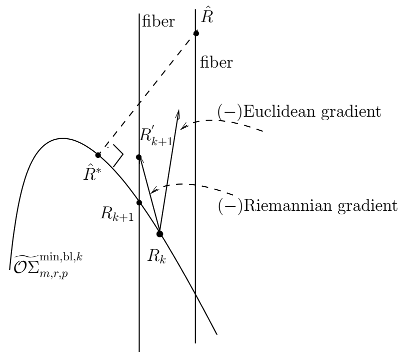


Fig. 1. Riemannian gradient descent for projection of $\hat{R}$ onto the manifold of balanced realizations.

sequentially (from the first column to the $n^{th}$) to yield a vector of size $nm$. The inverse of this operator is denoted by $\mathrm{vec}^{-1}$. For aesthetics often we may write $\overrightarrow{X}$ instead of $\mathrm{vec}(X)$, i.e., $\overrightarrow{X} \equiv \mathrm{vec}(X)$. For a realization $R = (A, B, C) \in \widetilde{\mathcal{L}}_{m,n,p}$, we define $\mathrm{vec}(R) \equiv \overrightarrow{R} \triangleq \left[\mathrm{vec}(A)^\top, \mathrm{vec}(B)^\top, \mathrm{vec}(C)^\top\right]^\top$, and $\mathrm{vec}(\cdot)$ induces a natural isomorphism between $\widetilde{\mathcal{L}}_{m,r,p}$ and $\mathbb{R}^{r^2+rm+rp}$. Define the function $h : \mathbb{R}^{r^2+rm+rp} \to \mathbb{R}^{r^2}$ as

$$h(\overrightarrow{R}) = \overrightarrow{W}_{\mathrm{c},k} - \overrightarrow{W}_{\mathrm{o},k}, \tag{25}$$

where $W_{\mathrm{c},k}$ $W_{\mathrm{o},k}$ are respectively the controllability and observability Gramians of $R$ of order $k \geq r$. Notice that $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p} = \mathrm{vec}^{-1}(h^{-1}(0_{n^2\times 1})) \cap \widetilde{\Sigma}^{\min}_{m,r,p}$ and for $k = \infty$, $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a},\mathrm{bl}}_{m,r,p} = \mathrm{vec}^{-1}(h^{-1}(0_{n^2\times 1})) \cap \widetilde{\Sigma}^{\min,\mathrm{a}}_{m,r,p}$.

Minimization of $f$ in (24) is an example where the cost function is extremely simple but the constraint set is a complicated manifold. The Euclidean gradient of $f$ at $R \in \widetilde{\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ is noting but $\frac{1}{2}(\overrightarrow{R} - \overrightarrow{\hat{R}})$. We simply equip $\widetilde{\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ with the Riemannian metric induced from the ambient Euclidean space $\mathbb{R}^{r^2+rm+rp}$. This implies that the Riemannian gradient of $f$ at $R \in \widetilde{\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ is found by *orthogonal* projection of the Euclidean gradient onto the *tangent space* of $\widetilde{\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ at $R$. This orthogonal projection matrix $\Pi_R$ can be derived explicitly from (25), the defining equations of the manifold $\widetilde{\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$. The related calculations can be found in Appendix B.

Figure 1 shows the main steps of the algorithm. We initialize by setting $R_0 = \sqrt{\nu(\hat{R})^{-1}} \circ \hat{R}$ as the p.d-balanced version of $\hat{R}$, where $\nu : \widetilde{\Sigma}^{\min}_{m,r,p} \to \mathcal{S}(r)$ is the bundle reduction map described in Section II-C. Notice that this transformation is simply moving $\hat{R}$ along its respective fiber into $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$. At each step $k$, we move along the Riemannian descent direction by step-size $\mu$ and then do p.d-balancing to get back to the manifold; if the cost function is not reduced enough, then we need to reduce the step-size, and repeat the p.d-balancing step. One could check along the way for asymptotic stability (if needed). The algorithm based on Armijo's steps-size selection rule is described as the following steps (see [29, Ch. 4] and [30, pp. 29-31] for more on Armijo's rule):

1: **function** PROJECT_ON_BALANCED($R$) ▷ $R \in \Sigma^{\min}_{m,r,p}$
2: `Choose` $R_0 \in \widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ `and` $\alpha \in (0, 1)$.
3: **repeat**
4: `set` $\mu = 1$ `and find` $\Pi_{R_i}$ `the orthogonal projection onto the tangent space of` $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ `at` $R_i$
5: `set` $\mathrm{grad} f_i = 1/2\Pi_{R_i}(\overrightarrow{R}_i - \overrightarrow{\hat{R}})$
6: `set` $\overrightarrow{R}'_{i+1} = \overrightarrow{R}_i - \mu\, \mathrm{grad} f_i$ `and` $R'_{i+1} = \mathrm{vec}^{-1}(\overrightarrow{R}'_{i+1})$
7: `if` $R'_{i+1} \notin \Sigma^{\min}_{m,r,p}$ `then set` $\mu \leftarrow \mu/2$ **go to** 6
8: `find p.d-balancing transformation` $S \in \mathcal{S}(r)$ `such that` $R''_{i+1} = \sqrt{S} \circ R'_i \in \widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$
9: `if` $f(R''_{i+1}) > f(R_i) - \alpha\mu\|\mathrm{grad} f_i\|^2$ `then set` $\mu \leftarrow \mu/2$ `and` **go to** 6 `else set` $R_{i+1} = R''_{i+1}$
10: **until** `convergence`
11: **return** $R_{i+1}$
12: **end function**

The parameter $\alpha$ (usually set in $[10^{-5}, 10^{-1}]$) ensures "enough reduction" in each update, and is needed to rule out pedagogical examples of non-convergence. In practice, we have observed convergence with only checking the strict decrease condition i.e., verifying $f(R''_{i+1}) < f(R_i)$ in step 9. Note that for $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a},\mathrm{bl}}_{m,r,p}$ (i.e., $k = \infty$), in step 6, the condition $R'_{i+1} \notin \widetilde{\Sigma}^{\min,\mathrm{a}}_{m,r,p}$ needs to be ruled, and in step 9, the projection matrix $\Pi_{R_i}$ associated with $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{a},\mathrm{bl}}_{m,r,p}$ has to be used. Next, we have this convergence result:

*Proposition 26:* Any accumulation point of the above algorithm in $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ is a critical point of $f$ (24) on $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ (including $k = \infty$).

*Proof:* The result follows from a general convergence result, Theorem 4.3.1 in [29, p. 65]. To apply that theorem, we need to show that the p.d-balancing transformation $R \mapsto \sqrt{\nu(R)^{-1}} \circ R$ is a so-called retraction (see [29, p. 55] for the definition). An explicit way exits to verify this (see [29, Proposition 4.1.2, p. 57]): if we find an open subset of $\mathbb{R}^{r^2+rm+rp}$, $\mathcal{E}^*$, and a diffeomorphism $\phi : \widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p} \times \mathcal{S}(r) \to \mathcal{E}^*$, such that $\phi(R, I_r) = R$ for every $R \in \widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$; then, $\phi_1^{-1}$, the first component of $\phi^{-1}$ is a retraction. For that, we simply choose $\phi(R, S) = S \circ R$ for any $R \in \widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ and $S \in \mathcal{S}(r)$ ($\phi_1^{-1}(R) = \sqrt{\nu(R)^{-1}} \circ R$). That $\phi$ is a diffeomorphism simply follows from smoothness of $\circ$ and $\nu$. Notice that $\mathcal{E}^* = \widetilde{\Sigma}^{\min}_{m,r,p}$ (or $\widetilde{\Sigma}^{\min,\mathrm{a}}_{m,r,p}$ when $k = \infty$). ■

Often in practice, $\hat{R}$ is close enough to $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ so that $f$ has a unique global minimizer, to which the algorithm will converge if started close enough.

## VII. A-PRIORI BOUNDS AND COMPARISON WITH D-BALANCED TRUNCATION

Although d-balanced truncation is not based on any optimality criterion, interestingly, there is a well-known a-priori bound on the $L^\infty$ norm (in frequency domain) of the approximation error for a.s. systems [5], [6]. The upper bound is $2\sum_{i=r+1}^{n} \lambda_i$, i.e., twice the sum of the $n - r$ residual (smallest) Hankel singular values, which resembles a similar bound in the rank-$r$ matrix approximation problem. Here, we derive a simple and essentially similar a-priori error bound for model reduction in the alignment distance for a.s. systems:

*Proposition 27:* Let $\lambda_1 \geq \cdots \geq \lambda_r \geq \cdots \geq \lambda_n > 0$ be the Hankel singular values of $M \in \Sigma^{\min,\mathrm{a}}_{m,n,p}$. Let $\bar{M}$ be a best $r^{th}$ order a.s. approximation of $M$ in the alignment distance solving problem (17), then we have

$$d_F^2(M,\bar{M}) \le 2(\sum_{i=r+1}^{n}\lambda_i)(1+\frac{1}{\lambda_r}) + (1-\|A_{22}\|_2)^{-1}\Big(\frac{1}{\lambda_r}(\sum_{i=r+1}^{n}\lambda_i)^2(2+\frac{1}{\lambda_r})\Big), \quad (26)$$

where $A_{22}$ is taken from the weak sub-realization $R_{22} = (A_{22},B_{21},C_{12})$ of order $n-r$ of any sorted d-balanced realization $R$ of $M$.

*Proof:* The result follows simply from recognizing that the alignment distance between $\bar{M}$ and $M$ is not larger than $d_F(R,\bar{R})$, where $R = (A,B,C)$ is a sorted d-balanced realization of $M$ and $\bar{R}$ is any d-balanced realization of order $r$ on the boundary. Notice that $R_{11} = (A_{11},B_{11},C_{11})$, the strong sub-realization of order $r$ in $R$, is not d-balanced or even balanced. However, it follows from a result in [6], that if $R_{11}$ is modified as follows, then it will be d-balanced:

$$\begin{aligned}\bar{A}_{11} &= A_{11} + A_{12}(I_{n-r}-A_{22})^{-1}A_{21} \\ \bar{B}_{11} &= B_{11} + A_{12}(I_{n-r}-A_{22})^{-1}B_{21} \\ \bar{C}_{11} &= C_{11} + C_{12}(I_{n-r}-A_{22})^{-1}A_{21}.\end{aligned} \quad (27)$$

By choosing $\bar{R} = (\left[\begin{smallmatrix}\bar{A}_{11} & 0\\ 0 & A_{22}\end{smallmatrix}\right], [\,\bar{B}_{11}\ 0\,], \left[\begin{smallmatrix}\bar{C}_{11}\\ 0\end{smallmatrix}\right])$, we have:

$$\begin{aligned}d_F^2(M,\bar{M}) &\le d_F^2(R,\bar{R}) = \|A_{12}\|_F^2 + \|A_{21}\|_F^2 \\ &+ \|B_{21}\|_F^2 + \|C_{12}\|_F^2 + \|A_{12}(I_{n-r}-A_{22})^{-1}A_{21}\|_F^2 + \\ &\|A_{12}(I_{n-r}-A_{22})^{-1}B_{21}\|_F^2 + \|C_{12}(I_{n-r}-A_{22})^{-1}A_{21}\|_F^2.\end{aligned} \quad (28)$$

Now let $\Lambda = \left[\begin{smallmatrix}\Lambda_1 & 0\\ 0 & \Lambda_2\end{smallmatrix}\right]$ be the d-balanced Gramian of $M$ (singular values put in decreasing order). Obviously (from the Lyapunov equations (2)) $B_{21}B_{21}^\top \preceq \Lambda_2$, $C_{12}C_{12}^\top \preceq \Lambda_2$, $A_{21}\Lambda_1A_{21}^\top \preceq \Lambda_2$, and $A_{12}^\top\Lambda_1A_{12} \preceq \Lambda_2$. From the first two we have $\|B_{21}\|_F^2 \le \sum_{i=r+1}^n\lambda_i$ and $\|C_{12}\|_F^2 \le \|\Lambda_2\|_F$. From the last two we have $\lambda_rA_{21}A_{21}^\top \preceq \Lambda_2$ and $\lambda_rA_{12}A_{12}^\top \preceq \Lambda_2$, and hence $\|A_{21}\|_F^2 \le \frac{1}{\lambda_r}\sum_{i=r+1}^n\lambda_i$ and $\|A_{12}\|_F^2 \le \frac{1}{\lambda_r}\sum_{i=r+1}^n\lambda_i$. Each of the last three terms in (28) also can be bounded as: $\|A_{12}(I_{n-r}-A_{22})^{-1}A_{21}\|_F^2 \le \|A_{12}\|_F^4\|(I_{n-r}-A_{22})^{-1}\|_2^2$, $\|A_{12}(I_{n-r}-A_{22})^{-1}B_{21}\|_F^2 \le \|A_{12}\|_F^2\|B_{21}\|_F^2\|(I_{n-r}-A_{22})^{-1}\|_2^2$, and $\|C_{12}(I_{n-r}-A_{22})^{-1}A_{21}\|_F^2 \le \|C_{12}\|_F^2\|A_{21}\|_F^2\|(I_{n-r}-A_{22})^{-1}\|_2^2$. The final result follows from these inequalities and that $\|(I_{n-r}-A_{22})^{-1}\|_2 \le (1-\|A_{22}\|_2)^{-1}$. ■

This bound can be very conservative—and most likely can be improved—as it is simply based on the evaluation of the cost function at a feasible point. The upper bound in (26) is interpreted as follows: the first term is due to truncation and the second term is due to projection of the truncated top sub-realization (onto $\widetilde{\Sigma}^{\text{min,a,bl}}_{m,r,p}$), which may not be balanced. The bound can be improved easily when $r \le m,p$:

*Proposition 28:* If $r \le \min\{m,p\}$, then the second term in the upper bound in (26) can be replaced by $2\frac{\lambda_{r+1}}{\lambda_r}\sum_{i=r+1}^n\lambda_i$.

*Proof:* The key point is that if $A$ is d-balanced, we can build a d-balanced realization $R'_{11} = (A_{11},B'_{11},C'_{11})$ from sub-realization $R_{11} = (A_{11},B_{11},C_{11})$. To see this, note that from (2) we have $\Lambda_1 = B_{11}B_{11}^\top + A_{11}\Lambda_1A_{11}^\top + A_{12}\Lambda_2A_{12}^\top$, where $\Lambda_1$ is diagonal. Then since $r \le m$ we can find $B'_{11}$ such that $B'_{11}B'^\top_{11} = B_{11}B_{11}^\top + A_{12}\Lambda_2A_{12}^\top$. In particular, it is easy to show that we can choose $B'_{11}$ such that $\|B_{11}-B'_{11}\|_F^2 \le \|A_{12}\Lambda_2A_{12}^\top\|_F$. To bound this, note that $\|A_{12}\Lambda_2A_{12}^\top\|_F \le \lambda_{r+1}\|A_{12}A_{12}^\top\|_F \le \lambda_{r+1}\|A_{12}\|_F^2$, which can be bounded by $\frac{\lambda_{r+1}}{\lambda_r}\sum_{i=r+1}^n\lambda_i$ (see the proof of Proposition 27). Similarly, $C'_{11}$ can be found and a bound can derived, which will be added to this bound. ■

### A. Connections with D-balanced Truncation

Next, recalling the discussion in Section I-A, we see how model order reduction in the alignment distance can be considered as an enhanced version of d-balanced truncation. The above proofs suggest that in certain cases d-balanced truncation can be considered as an approximate solution to model reduction in the alignment distance. However, the realization alignment built in problem (17) may render the two significantly different, as seen next.

*Example 29:* Consider the d-balanced realization $R = (A,B,C)$: $A = \left[\begin{smallmatrix}0.9999 & -0.0010\\ -0.0010 & 0.9487\end{smallmatrix}\right]$, $B = \left[\begin{smallmatrix}0.1026\\ 0.9997\end{smallmatrix}\right]$, $C = [\,0.1026\ 0.9997\,]$, where $W_o = W_c = \Lambda = \left[\begin{smallmatrix}88.7345 & 0\\ 0 & 9.9931\end{smallmatrix}\right]$. Since $A$ is symmetric alignment distance reduction does not need a projection step. If we use Moore's truncation of the d-balanced realization as an approximation we get the reduced order system with realization $\bar{R}_T = (\left[\begin{smallmatrix}0.9999 & 0\\ 0 & 0.9487\end{smallmatrix}\right], \left[\begin{smallmatrix}0.1026\\ 0\end{smallmatrix}\right], [\,0.1026\ 0\,])$ ($(\bar{R}_\text{T})_{11}$ being the minimal first order solution) and (alignment distance) error of $d_{\text{dbl}} = 1.2783$, whereas with the alignment distance based reduction we get the first order system $\bar{R}_{\text{ATP}} = (\left[\begin{smallmatrix}0.9490 & 0\\ 0 & 0.9996\end{smallmatrix}\right], \left[\begin{smallmatrix}1.0050\\ 0\end{smallmatrix}\right], [\,1.0050\ 0\,])$ and error $d_{\text{ATP}} = 0.0059$, which is significantly lower than $d_{\text{dbl}}$.

Here, although the Hankel singular value of the strong sub-realization is much larger than that of the weak sub-realization ($88.7 \gg 9.99$), the norms of $B_{21}$ and $C_{12}$ are much larger than those of $B_{11}$ and $C_{11}$ ($.99 \gg .1$). Thus overall the truncated d-balanced realization is not a good solution to the alignment problem and the realization alignment reduces the error significantly. In analogy with the continuous-time case (see proof of Proposition 3 and [8]), we call the norms of rows of $B_{21}$ and $C_{12}^\top$ balanced gains. Interestingly, as a shortcoming of d-balanced truncation, it has been argued that d-balanced truncation is *blind* to the balanced gains (in continuous-time) and may result in poor $L^2$ norm errors [8]. In our example, the $\ell^2$ error norms are $d_{\text{bl},\ell^2} = 3.1576$ and $d_{\text{ATP},\ell^2} = 0.5631$, indicating that the alignment distance model reduction gives better $\ell^2$ error in this example. However, since our original criterion is different from $\ell^2$ this situation does not hold in general. Recall that although a zero (or small) alignment distance between two systems implies zero (or small) distance between their impulse responses, the alignment distance itself is not computed based on direct comparison of the impulse responses. Finding an explicit relation between input-output based distances and the alignment distance is, indeed, a relevant and interesting question.

## VIII. Robustness of Internal Stability under Feedback in the Alignment Distance

Robustness of stability under feedback is essentially a question about the topology induced by a distance used to compare a system and its perturbations [31]. Distances such as the $L^2$, $L^\infty$, and Hankel norm based distances are not suitable in that regard [31], [32], since intuitively, these distances are defined for a.s. systems, and the distance between any unstable system and a.s. system in such distances is *infinity*. Instead, distances such a the *gap* metric are most suitable in a topologically exact sense [33]. We also recall that model order reduction in the gap metric is computationally challenging. The alignment distance provides an *immediate* and natural formulation of robustness of internal stability. The nature of the result, however, is limited or different compared with the gap metric and operator-theoretic methods, as here the universe of perturbations is limited to systems up to order $n$.

Consider a (possibly unstable) s-balanced system $M \in \overline{\Sigma}^{\min}_{m,n,p}$ and the closed-loop system around it with output feedback via a constant gain $K \in \mathbb{R}^{m\times p}$. The state-space equations of the closed-loop system can be written as

$$\begin{aligned} x_t &= Ax_{t-1} + B(u_t - Ky_{t-1}) \\ y_t &= Cx_t, \end{aligned} \tag{29}$$

where without loss of generality we assume that the realization $R = (A, B, C)$ is a realization in the reduced bundle of $k$-balanced realizations $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p}$ (for some $k \geq n$, see (5)).

*Theorem 30 (Robustness of Feedback Stability on $\mathring{\overline{\Sigma}}_{m,n,p}$):* Internal stability under constant gain output feedback is a robust property on the space of diagonalizable s-balanced systems $\mathring{\overline{\Sigma}}_{m,n,p}$ with respect to the exact extended alignment distance $\bar{d}_F$ subordinate to $\overline{\widetilde{\mathcal{O}\Sigma}}^{\min,\mathrm{bl},k}_{m,n,p}$, i.e., if $M \in \mathring{\overline{\Sigma}}_{m,n,p}$ is internally a.s. under constant gain $K$, for small enough $\epsilon > 0$, every $\bar{M} \in \mathring{\overline{\Sigma}}_{m,n,p}$ with $\bar{d}_F(M, \bar{M}) < \epsilon$, is also internally stable under feedback gain $K$. The same holds for the one-sided alignment distance $\overleftarrow{d}_F$ (18).

*Proof:* Let $M \in \mathring{\overline{\Sigma}}_{m,n,p}$ be a system for which the closed-loop system (29) is a.s., i.e., $\rho(A - BKC) < 1$, $\rho(X)$ being the spectral radius of matrix $X$. Notice that this relation is independent of any specific realization $R$ of $M$. Thus, without loss of generality, we can assume that the realization $R$ in (29) is in $\overline{\widetilde{\mathcal{O}\Sigma}}^{\min,\mathrm{bl},k}_{m,n,p}$. Let $\bar{M} \in \overline{\Sigma}^{\min}_{m,n,p}$ be another system with a realization $\bar{R} \in \overline{\mathring{\widetilde{\mathcal{O}\Sigma}}}^{\min,\mathrm{bl},k}_{m,n,p}$. If $\bar{d}_F(M, \bar{M}) < \epsilon$, then there exists $Q \in O(n)$ such that $\|Q^\top AQ - \bar{A}\|_F < \epsilon$, $\|Q^\top B - \bar{B}\|_F < \epsilon$, and $\|CQ - \bar{C}\|_F < \epsilon$; thus for small enough $\epsilon$, $Q^\top(A - BKC)Q$ and $\bar{A} - \bar{B}K\bar{C}$ can be made close enough. Then, it follows that, because of continuous dependence of the eigenvalues of a matrix on its entries, for small enough $\epsilon$, $\rho(Q^\top(\bar{A} - \bar{B}K\bar{C})Q) = \rho(\bar{A} - \bar{B}K\bar{C}) < 1$. A similar argument applies to the one-sided alignment distance. Notice that in this case, if $\bar{M}$ is a non-minimal system of minimal order $r$ then $\|Q^\top AQ - P^{-1}\bar{A}P\|_F < \epsilon$, $\|Q^\top B - P^{-1}\bar{B}\|_F < \epsilon$, and $\|CQ - \bar{C}P\|_F < \epsilon$ with $P = \left[\begin{smallmatrix} I_r & 0 \\ 0 & P_{22}\end{smallmatrix}\right] \in GL(n)$. So for small $\epsilon$, $P^{-1}(\bar{A} - \bar{B}K\bar{C})P$ will be close to $Q^\top(A - BKC)Q$; hence, $\rho(\bar{A} - \bar{B}K\bar{C}) < 1$. ■

A more workable statement (with the same proof) in relation with our model order reduction (17) is:

*Proposition 31:* Let $M \in \Sigma^{\min}_{m,n,p}$. Let $\bar{M}$ be a solution to (17) (in the sense of Proposition 18) with $d_F(M, \bar{M}) < \epsilon$. If $\bar{M}$ is internally a.s. under feedback gain $K$, then for small enough $\epsilon$, $M$ and any another system $M' \in \Sigma^{\min}_{m,n,p}$ with $d_F(M', \bar{M}) < \epsilon$ are internally a.s. under feedback gain $K$.

We should be cautious that in the above result $\bar{M}$ is a system of order $n$ and minimal order $r$. Thus it has a balanced Kalman canonical realization of the form (9), and internal stability of the closed loop system requires that the non-minimal sub-realization namely $(\bar{A}_{22}, 0, 0)$ be *internally* a.s. The implication is that if the non-minimal sub-realization is unstable, then robustness *cannot* be guaranteed. We leave quantitative and deeper results on robustness to future works.

## IX. Simulations

In this section we apply the ATP algorithm to an unstable MIMO system of order $n = 5$ and output-input dimension $(p, m) = (2, 2)$ to obtain a system of order $r = 3$. Consider a system $M$ with a d-balanced realization $R = (A, B, C)$:

$$A=\begin{bmatrix} -0.9214 & -0.0176 & 0.4130 & -0.1806 & -0.0241 \\ 0.0904 & 0.9624 & 0.8531 & -0.2160 & -0.0874 \\ 0.4050 & -0.9475 & 0.6156 & 0.1830 & -0.0708 \\ 0.2292 & -0.3691 & -0.1726 & 0.5934 & 0.4062 \\ 0.0390 & -0.1247 & 0.0399 & 0.5056 & 0.3770 \end{bmatrix}, B=\begin{bmatrix} 0.0661 & 2.2774 \\ 0.2831 & 1.8997 \\ -0.1828 & -0.3285 \\ 0.1285 & 1.2140 \\ 0.2215 & -0.1981 \end{bmatrix},$$

$$C=\begin{bmatrix} -1.6695 & 1.6789 & 0.3094 & 0.7627 & 0.1260 \\ -1.5226 & 0.9225 & -0.5052 & 0.9628 & -0.3751 \end{bmatrix}.$$

The system has poles $p_1, p_2 = 0.8485 \pm 0.8486\mathbf{i}$, $p_3 = -0.9800$, $p_4 = 0.9172$, and $p_5 = -0.0072$, where $p_1$ and $p_2$ are unstable with $|p_1| = |p_2| = 1.2$. The singular values of the system are $\lambda_1 = 25.9078, \lambda_2 = 21.7456, \lambda_3 = 12.1154, \lambda_4 = 2.7332, \lambda_5 = 0.4586$. We run the ATP algorithm with initial solution as the d-balanced truncated realization and the gradient projection algorithm in Section VI-A in which the projection operator onto the tangent space of $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ with $k = n = 5$ is used. In the implementation we use the algorithm in [18] to compute the alignment distance (the alignment step). Figure 2 shows the squared of the error $d_F^2(M, \bar{M}_k)$ in reduction in terms of iteration index $k$ (each align-truncate-project step is called one iteration). Here, the alignment distance is subordinate to the reduced bundle $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,n,p}$ with $k = n = 5$. The first point in the graph is the alignment distance error in simply using d-balanced truncation. In this case, the reduction in the error beyond the initial d-balanced truncation is not significant (although still tangible). The output of the ATP algorithm i.e., the final reduced order ($r = 3$) (balanced) realization is:

$$\bar{A}_{11} = \begin{bmatrix} -0.9509 & 0.1079 & 0.3762 \\ -0.1892 & 0.6584 & -0.6729 \\ 0.3489 & 0.7297 & 0.5540 \end{bmatrix}, \bar{B}_{11} = \begin{bmatrix} -0.0664 & 2.2651 \\ -0.3070 & -2.2463 \\ -0.1599 & -0.3762 \end{bmatrix}$$

$$\bar{C}_{11} = \begin{bmatrix} -1.6816 & -1.8453 & 0.2962 \\ -1.5258 & -1.2563 & -0.6109 \end{bmatrix}, \bar{A}_{22} = \begin{bmatrix} 0.7742 & -0.4958 \\ -0.5057 & 0.5366 \end{bmatrix}.$$

The eigenvalues of $\bar{A}_{11}$ are $p'_1 = -1.0349$ $p'_2, p'_3 = 0.6481 \pm 0.7166\mathbf{i}$ ($|p'_1| = |p'_2| = 0.9662$) and the eigenvalues of $\bar{A}_{22}$ are $p'_4 = 1.1700$ and $p'_5 = 0.1408$. The final approximation error or squared distance to minimality $d_F^2(M, \bar{M}) = 1.4048$. Notice that $\bar{A}_{22}$ has a unstable pole. Thus in this case, robustness based on Theorem 30 is out of question. Also while $\bar{A}_{11}$ has one unstable pole, its complex poles of are are stable. From an engineering point of view (in certain circumstances) this might be undesirable. To amplify the effect of unstable poles, we could try to use a different value for $k$, the order of the Gramians. For example, if we use $k = 2n = 10$, then we get $\bar{A}_{11}$ with poles $p''_1, p''_2 = 0.7415 \pm 0.7685\mathbf{i}$ and $p''_3 = -0.9850$, where $|p_1| = |p_2| = 1.0679$ and $\bar{A}_{22}$ with poles $p'_4 = 0.8772$ and $p'_5 = 0.2008$. In this case, $\bar{A}_{22}$ is a.s., thus there is possibility that based on Theorem 30 or Proposition 31, by stabilizing $\bar{M}$, $M$ also is stabilized. Notice that by changing $k$, we need to use a different ATP algorithm, as the manifold $\widetilde{\mathcal{O}\Sigma}^{\min,\mathrm{bl},k}_{m,r,p}$ and the projection operator $\Pi_R$ depends on $k$.

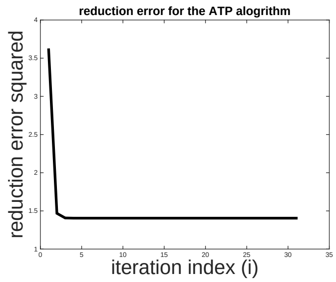


Fig. 2. The performance of the ATP algorithm in an example.

## X. Conclusions

In this paper, we showed how the alignment distance can be used to formulate the problem of model order reduction, thereby we showed how "minimal realization theory" and be quantitatively related to the problem of model order reduction. This formulation can be interpreted as an enhanced version of the popular d-balanced truncation method. As a byproduct, we showed that the set of diagonalizable s-balanced systems of fixed order is metrizable. We also developed an efficient algorithm for model order reduction, and established robustness of feedback stability under the alignment distance. A better understanding of model order reduction including better a-priori bounds, properties of the solution, and improved algorithms are among future possible research directions. We only studied discrete-time deterministic systems, other classes of systems especially stochastic systems also need to be studied. Additionally, the effectiveness of our methods in engineering applications has to be examined.

## Appendix A
## Proof of Proposition 22

*Proof:* Let $\bar{R} = (\left[\begin{smallmatrix}\bar{a}_{11} & 0\\ 0 & \bar{A}_{22}\end{smallmatrix}\right], \left[\begin{smallmatrix}\bar{B}_{11}\\ 0\end{smallmatrix}\right], \left[\begin{smallmatrix}\bar{C}_{11} & 0\end{smallmatrix}\right])$ be a solution which achieves the minimum in (17) ($\bar{a}_{11} \in \mathbb{R}, \bar{B}_{11} \in \mathbb{R}^{1\times m}$ and $\bar{C}_{11} \in \mathbb{R}^{p\times 1}$). Note that $\bar{R}_{11} = (\bar{a}_{11}, \bar{B}_{11}, \bar{C}_{11}) \in \widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,1,p}$ and we assume that $\bar{a}_{11} = 1$ or $\bar{a}_{11} = -1$ otherwise $\bar{R}_{11}$ will be a.s. We show that another realization $(\bar{a}, \bar{B}_{11}, \bar{C}_{11})$ with $|a| < 1$ achieves a lower cost in (17) than $\bar{R}_{11}$. As a first step we have the next simple lemma:

*Lemma 32:* Any realization $\bar{R} = (\bar{a}, \bar{B}, \bar{C})$ at the boundary of $\widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,1,p}$ in $\tilde{\mathcal{L}}_{m,1,p}$ (i.e., any realization which is stable but not a.s. and is a limit of a sequence of balanced a.s realizations) is characterized by: $\bar{a} = \pm 1$ and $\bar{B}\bar{B}^\top = \bar{C}^\top\bar{C}$.

*Proof:* Any such realization $\bar{R}$ is the limit of a sequence of a.s. balanced realizations $R_i = (a_i, B_i, C_i)$ for which the equality of controllability and observability Gramians implies that $\bar{B}_i\bar{B}_i^\top = \bar{C}_i^\top\bar{C}_i$. This in turn implies both relations above. Also conversely any realization with satisfying those relations is a limit of a sequence of a.s. stable balanced realizations. ■

Now let $Q \in O(n)$ be an orthogonal matrix that achieves the minimum in (17). Since $\|A\|_2 \leq 1$ (see [28]) and $A$ is a.s., then by Lemma 33 (bellow) all the diagonal entries of $Q^\top AQ$ are strictly smaller than 1 in absolute value. In particular, we have $|(Q^\top AQ)_{11}| < 1$. Thus if we replace $\bar{R}_{11}$ with the sub-realization $\bar{R}'_{11} = ((Q^\top AQ)_{11}, \bar{B}_{11}, \bar{C}_{11})$ which is an a.s. balanced realization, then it achieves a lower distance in (17) than what $\bar{R}$ achieves, which contradicts optimality of $\bar{R}$. (Note that by Lemma 32, $\bar{B}_{11}\bar{B}_{11}^\top = C_{11}^\top C_{11}$ which guarantees that $\bar{R}'_{11}$ is balanced). This completes the proof.

*Lemma 33:* If $\|A\|_2 \leq 1$ and $|\lambda_{\max}(A)| < 1$, then we have $|A_{ii}| < 1$ for $1 \leq i \leq n$.

*Proof:* Recall that $|A_{ii}| \leq \|A\|_2 \leq 1$. We show that $|A_{ii}| = 1$ leads to a contradiction. Let $i = 1$ for convenience. Consider the first coordinate vector $e_1$. Since $\|Ae_1\| \leq 1$ and $|A_{11}| = 1$ we have $A_{j1} = 0$ for $1 < j \leq n$. Similarly since $\|A^\top e_1\| \leq 1$ we must have $A_{1j} = 0$ for $1 < j \leq n$. Therefore, all elements in the first row and column of $A$ except for the diagonal entry $A_{11}$ are zero. Hence, $A_{11}$ (with $|A_{11}| = 1$) is an eigenvalue of $A$ which is a contradiction. ■

■

## Appendix B
## Projection On Tangent Spaces of Manifolds of Balanced Realizations

The main object of interest in our calculations is the derivative of $h$ (see (25)) at every point $R$, denoted as $\frac{\mathrm{d}\,h}{\mathrm{d}\,\vec{R}}$, which is a matrix of dimension $r^2 \times (r^2 + rm + rp)$:

$$\frac{\mathrm{d}h}{\mathrm{d}\,\vec{R}} = [\frac{\mathrm{d}\vec{W}_{\mathrm{c},k}}{\mathrm{d}\,\vec{A}} - \frac{\mathrm{d}\vec{W}_{\mathrm{o},k}}{\mathrm{d}\,\vec{A}}, \frac{\mathrm{d}\vec{W}_{\mathrm{c},k}}{\mathrm{d}\,\vec{B}}, -\frac{\mathrm{d}\vec{W}_{\mathrm{o},k}}{\mathrm{d}\,\vec{C}}] \tag{30}$$

Notice that due to symmetries in the Gramians $W_{\mathrm{c},k}$ and $W_{\mathrm{o},k}$, there are at most $\frac{r(r+1)}{2}$ independent constraints. Indeed, there are exactly $\frac{r(r+1)}{2}$ constraints, and the rank of $\frac{\mathrm{d}\,h}{\mathrm{d}\,\vec{R}}$ is $\frac{r(r+1)}{2}$. Thus the manifold of balanced realizations is of dimension $\ell = \frac{1}{2}r(r-1) + rm + rp$. Let $\frac{\mathrm{d}\,h}{\mathrm{d}\,\vec{R}} = U\left[\begin{smallmatrix}D_\ell & 0\\ 0 & 0\end{smallmatrix}\right]V^\top$, where $D_\ell \succ 0$, be the full SVD of $\frac{\mathrm{d}\,h}{\mathrm{d}\,\vec{R}}$. Note that $V$ is square and of dimension $r^2+rm+rp$. The last $\ell$ columns of $V$ denoted by the matrix $V_R$ form a basis for the tangent space of $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,r,p}$ or $\widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,r,p}$ at $R$ (depending on the value of $k$, see (25)). We denote the orthogonal projection operator as $\Pi_R = V_R V_R^\top$. For $\widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,r,p}$, $\frac{\mathrm{d}\,h}{\mathrm{d}\,\vec{R}}$ can be found from (32) below. For $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,r,p}$, $\frac{\mathrm{d}\,h}{\mathrm{d}\,\vec{R}}$ can be found from (33), (34), and (35) below.

We establish some notations first. Let $A, B \in \mathbb{R}^{r\times r}$. Then we have $\mathrm{vec}(B \otimes A) = K_{\mathrm{L}}\mathrm{vec}(B)$ and $\mathrm{vec}(A \otimes B) = K_{\mathrm{R}}\mathrm{vec}(B)$, where

$$K_{\mathrm{L}} = I_r \otimes \begin{bmatrix} I_r\otimes A(:,1)\\ \vdots\\ I_r\otimes A(:,r)\end{bmatrix} \text{ and } K_{\mathrm{R}} = \begin{bmatrix} I_r\otimes\big(A(:,1)\otimes I_r\big)\\ \vdots\\ I_r\otimes\big(A(:,r)\otimes I_r\big)\end{bmatrix}. \tag{31}$$

In the above $A(:,i)$ is the $i^{th}$ column of $A$ and $\otimes$ denotes the Kronecker product. By $H_{mn}$ we denote the $mn \times mn$ commutation matrix, i.e., $H_{mn}$ solves $\mathrm{vec}(X)^\top = H_{mn}\mathrm{vec}(X)$ for $X \in \mathbb{R}^{m\times n}$.

Using simple algebra and properties of $\mathrm{vec}(\cdot)$ one can show that for $R = (A, B, C) \in \widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,r,p}$:

$$\frac{\mathrm{d}\vec{W}_{\mathrm{c}}}{\mathrm{d}\,\vec{A}} = \Big(\mathrm{vec}(BB^\top)^\top \otimes I_{r^2}\Big)\Big((I_{r^2} - A^\top \otimes A^\top)^{-1} \otimes (I_{r^2} - A \otimes A)^{-1}\Big)\Big(K_{\mathrm{L}} + K_{\mathrm{R}}\Big) \tag{32a}$$

$$\frac{\mathrm{d}\vec{W}_{\mathrm{o}}}{\mathrm{d}\,\vec{A}} = \Big(\mathrm{vec}(C^\top C)^\top \otimes I_{r^2}\Big)\Big((I_{r^2} - A \otimes A)^{-1} \otimes (I_{r^2} - A^\top \otimes A^\top)^{-1}\Big)\Big(K_{\mathrm{L}} + K_{\mathrm{R}}\Big)H_{nn} \tag{32b}$$

$$\frac{\mathrm{d}\vec{W}_{\mathrm{c}}}{\mathrm{d}\,\vec{B}} = B \otimes I_r + (I_r \otimes B)H_{rm} \tag{32c}$$

$$\frac{\mathrm{d}\vec{W}_{\mathrm{o}}}{\mathrm{d}\,\vec{C}} = (C^\top \otimes I_r)H_{pr} + I_r \otimes C^\top. \tag{32d}$$

For the case of $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,r,p}$ ($k < \infty$) one can derive recursive equations to calculate $\frac{d\overrightarrow{W}_{c,k}}{d\overrightarrow{A}}$ as follows:

$$\frac{d\overrightarrow{W}_{c,k}}{d\overrightarrow{A}} = \sum_{i=1}^{k} \frac{d\overrightarrow{\mathcal{W}}_{c,i}}{d\overrightarrow{A}} \tag{33a}$$

$$\mathcal{W}_{c,i} = A^{i-1}BB^\top(A^\top)^{i-1} = A\mathcal{W}_{c,i-1}A^\top \tag{33b}$$

$$\frac{d\overrightarrow{\mathcal{W}}_{c,i}}{d\overrightarrow{A}} = (A\mathcal{W}_{c,i-1} \otimes I_r) + (A \otimes A)\frac{d\overrightarrow{\mathcal{W}}_{c,i-1}}{d\overrightarrow{A}} + (I_r \otimes A\mathcal{W}_{c,i-1})H_{rr} \quad (i \geq 2) \tag{33c}$$

where $\mathcal{W}_{c,1} = BB^\top$ and $\frac{d\overrightarrow{\mathcal{W}}_{c,1}}{d\overrightarrow{A}} = 0 \in \mathbb{R}^{r^2 \times r^2}$. Similarly for $\frac{d\overrightarrow{W}_{o,k}}{d\overrightarrow{A}}$ one can show that:

$$\frac{d\overrightarrow{W}_{o,k}}{d\overrightarrow{A}} = \sum_{i=1}^{k} \frac{d\overrightarrow{\mathcal{W}_{o,i}}}{d\overrightarrow{A}} \tag{34a}$$

$$\mathcal{W}_{o,i} = (A^\top)^{i-1}C^\top CA^{i-1} = A^\top\mathcal{W}_{o,i-1}A \tag{34b}$$

$$\frac{d\overrightarrow{\mathcal{W}}_{o,i}}{d\overrightarrow{A}} = (A^\top\mathcal{W}_{o,i-1} \otimes I_r)H_{rr} + (A^\top \otimes A^\top)\frac{d\overrightarrow{\mathcal{W}}_{o,i-1}}{d\overrightarrow{A}} + (I_r \otimes A^\top\mathcal{W}_{c,i-1}) \; (i \geq 2) \tag{34c}$$

where $\mathcal{W}_{o,1} = C^\top C$ and $\frac{d\overrightarrow{\mathcal{W}}_{o,1}}{d\overrightarrow{A}} = 0 \in \mathbb{R}^{r^2 \times r^2}$. Finally for $\frac{d\overrightarrow{W}_{c,k}}{d\overrightarrow{B}}$ and $\frac{d\overrightarrow{W}_{o,k}}{d\overrightarrow{C}}$ similar recursive equations are:

$$\frac{d\overrightarrow{W}_{c,k}}{d\overrightarrow{B}} = \sum_{i=1}^{k} \frac{d\overrightarrow{\mathcal{W}}_{c,i}}{d\overrightarrow{B}} \tag{35a}$$

$$\frac{d\overrightarrow{\mathcal{W}}_{c,1}}{d\overrightarrow{B}} = (B \otimes I_r) + (I_r \otimes B)H_{rm} \tag{35b}$$

$$\frac{d\overrightarrow{\mathcal{W}}_{c,i}}{d\overrightarrow{B}} = (A \otimes A)\frac{d\overrightarrow{\mathcal{W}}_{c,i-1}}{d\overrightarrow{B}}(i \geq 2) \tag{35c}$$

$$\frac{d\overrightarrow{W}_{o,k}}{d\overrightarrow{C}} = \sum_{i=1}^{k} \frac{d\overrightarrow{\mathcal{W}}_{o,i}}{d\overrightarrow{C}} \tag{35d}$$

$$\frac{d\overrightarrow{\mathcal{W}}_{c,1}}{d\overrightarrow{C}} = (I_r \otimes C^\top) + (C^\top \otimes I_r)H_{pr} \tag{35e}$$

$$\frac{d\overrightarrow{\mathcal{W}}_{o,i}}{d\overrightarrow{C}} = (A^\top \otimes A^\top)\frac{d\overrightarrow{\mathcal{W}}_{o,i-1}}{d\overrightarrow{C}}(i \geq 2). \tag{35f}$$

## ACKNOWLEDGMENT

The authors thank Laurent Younes for helpful comments and discussions.

## REFERENCES

[1] R. E. Kalman, "Canonical structure of linear dynamical systems," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 48, no. 4, p. 596, 1962.
[2] B. C. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Transactions On Automatic Control*, vol. 26, pp. 17–32, 1981.
[3] B. Afsari and R. Vidal, "Bundle reduction and the alignment distance on spaces of state-space LTI systems," to appear in IEEE Transactions to Automatic Control, available at http://vision.jhu.edu/publications.htm, 2015.
[4] ——, "The alignment distance on spaces of linear dynamical systems," in *IEEE Conference on Decision and Control*, 2013, pp. 1162–1167.
[5] D. F. Enns, "Model reduction with balanced realizations: An error bound and a frequency weighted generalization," in *Decision and Control, 1984. The 23rd IEEE Conference on*, vol. 23. IEEE, 1984, pp. 127–132.
[6] U. M. Al-Saggaf and G. F. Franklin, "Model reduction via balanced realizations: an extension and frequency weighting techniques," *Automatic Control, IEEE Transactions on*, vol. 33, no. 7, pp. 687–692, 1988.
[7] M. Hazewinkel and R. E. Kalman, "On invariants, canonical forms, and moduli for linear constant, finite-dimensional dynamical systems," in *Lecture Notes in Econ.-Mathematical System Theory*. New York: Springer-Veflag, 1976, vol. 131, pp. 48–60.
[8] P. Kabamba, "Balanced gains and their significance for model reduction," *Automatic Control, IEEE Transactions on*, vol. 30, no. 7, pp. 690–693, 1985.
[9] B. Afsari and R. Vidal, "Group action induced distances on spaces of high-dimensional linear stochastic processes," in *Geometric Science of Information*, ser. LNCS, vol. 8085, 2013, pp. 425–432.
[10] R. E. Kalman, "Algebraic geometric description of the class of linear systems of constant dimension," in *8th Ann. Princeton Conf. on Information Sciences and Systems*, Princeton, N.J., March 1974.
[11] P. S. Krishnaprasad, "Geometry of minimal systems and the identification problem," Ph.D. dissertation, Harvard University, 1977.
[12] D. F. Delchamps, "Global structure of families of multivariable linear systems with an application to identification," *Mathematical Systems Theory*, vol. 18, pp. 329–380, 1985.
[13] A. S. Khadr and C. F. Martin, "On the $G\ell(n)$ action on linear systems: The orbit closure problem," *Algebraic and Geometric Methods in Linear System Theory, Amer. Math. Soc, Providence*, 1980.
[14] B. Afsari and R. Vidal, "Model order reduction for discrete-time LTI systems using the alignment distance," in *American Control Conference*, 2015, pp. 2868–2875.
[15] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry Volume I*, ser. Wiley Classics Library Edition. John Wiley & Sons, 1963.
[16] U. Helmke, "Balanced realizations for linear systems: a variational approach," *SIAM Journal on Control and Optimization*, vol. 31, no. 1, pp. 1–15, January 1993.
[17] E. I. Verriest and W. S. Gray, "A geometric approach to the minimum sensitivity design problem," *SIAM journal on control and optimization*, vol. 33, no. 3, pp. 863–881, 1995.
[18] N. D. Jimenez, B. Afsari, and R. Vidal, "Fast Jacobi-type algorithm for computing distances between linear dynamical systems," in *European Control Conference*, 2013, pp. 3682 – 3687.
[19] U. Helmke and J. B. Moore, *Optimization and dynamical systems*. Springer-Verlag, 1994.
[20] R. J. Ober, "Balanced realizations: canonical form, parametrization, model reduction," *International Journal of Control*, vol. 46, no. 2, pp. 643–670, 1987.
[21] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their $L^\infty$ error bounds," *International Journal of Control*, vol. 39, no. 6, pp. 1115–1193, 1984.
[22] U. Helmke, "A global parametrization of asymptotically stable linear systems," *Systems & control letters*, vol. 13, no. 5, pp. 383–389, 1989.
[23] T. Kailath, *Linear Systems*. Prentice Hall, 1980.
[24] R. Eising, "The distance between a system and the set of uncontrollable systems," in *Mathematical Theory of Networks and Systems*. Springer, 1984, pp. 303–314.
[25] P. S. Krishnaprasad and C. F. Martin, "On families of systems and deformations," *International Journal of Control*, vol. 38, no. 5, pp. 1055–l079, 1983.
[26] P. Parrilo and L. Ljung, "Initialization of physical parameter estimates," in *Proceedings of the 13th IFAC Symposium on System Identification*, 2003, pp. 1524–1529.
[27] O. Prot, G. Mercère *et al.*, "Initialization of gradient-based optimization algorithms for the identification of structured state-space models," in *18th World Congress of International Federation of Automatic Control. IFAC*, vol. 18, no. 1, 2011, pp. 10 782–10 787.
[28] L. Pernebo and L. Silverman, "Model reduction via balanced state space representations," *IEEE Transactions on Automatic Control*, vol. 27, no. 2, pp. 382–387, April 1982.
[29] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ: Princeton University Press, 2008.
[30] D. P. Bertsekas, *Nonlinear programming*, 2nd ed. Athena Scientific, 1999.
[31] B. Francis, "On robustness of the stability of feedback systems," *Automatic Control, IEEE Transactions on*, vol. 25, no. 4, pp. 817–818, 1980.
[32] E. Jonckheere, M. Safonov, and L. Silverman, "Topology induced by the Hankel norm in the space of transfer matrices," in *Decision and Control including the Symposium on Adaptive Processes, 1981 20th IEEE Conference on*, vol. 20. IEEE, 1981, pp. 118–119.
[33] A. El-Sakkary, "The gap metric: Robustness of stabilization of feedback systems," *IEEE Transactions on Automatic Control*, vol. 30, no. 3, pp. 240–247, 1985.