Bundle Reduction and the Alignment Distance on Spaces of State-Space LTI Systems

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

***Abstract*—This paper introduces a large class of differential-geometric distances between finite-dimensional linear dynamical systems, collectively called the alignment distance. Contrary to the existing distances, the alignment distance is based on the state-space description of dynamical systems, and is defined on manifolds of systems of fixed order and fixed input-output dimension under a matrix rank constraint (e.g., minimality, controllability, or observability). While the quotient topology and principal fiber bundle structure associated with such manifolds have been known since the early days of modern control theory, distances natural to this structure have not been studied. The starting point for defining such a distance is to identify a linear system of order n with its equivalence class of state-space realizations, all related by the so-called similarity action, i.e., state-space change of basis under $GL(n)$, the Lie group of nonsingular $n \times n$ matrices. The main idea of the alignment distance is to first find the best state-space change of basis that brings a realization of a system "as close as possible" to a realization of another system (the alignment step), and then compare the aligned realizations. A direct implementation of this idea, due to noncompactness of $GL(n)$, is complicated. However, using the notion of "reduction of the structure group" of a principal bundle, we show that the change of basis can be restricted to an orthogonal change of basis, provided one uses realizations in a reduced subbundle. This key observation brings about significant computational benefits. As a technical contribution (possibly of independent interest), we show that several forms of realization balancing available in the control literature have differential-geometric significance, and are, indeed, examples of reducing the structure group from $GL(n)$ to its subgroup of orthogonal matrices $O(n)$. The alignment distance can be defined for stable and unstable systems, discrete or continuous-time, and stochastic systems.**

***Index Terms*—Linear dynamical systems, State-space description, manifolds, order, principal fiber bundles, balanced realizations, distances, quotient topology, Riemannian manifolds.**

I. Introduction

THE notion of a *distance* between linear dynamical systems is fundamental in control theory. A distance appears directly or indirectly in many basic control problems. The notable examples are the problems of robust control and model order reduction. Closely related to the notion of a distance is the notion of a *space of systems* and its geometrization.

The starting point in studying spaces of linear dynamical systems and their geometries is the basic question of *mathematical description* of systems.[1] Since the early days of modern control theory two competing —but not exactly equivalent—formulations for describing physical dynamical systems have existed: First, what Kalman in his seminal paper [1] calls "the old approach," namely, the *input-output*, *operator theoretic*, *transfer function*, or *external* description of a dynamical system; and, second, what he calls "the new approach," which is the *state-space*, *differential equation* (or difference equation), or *internal* description of a dynamical system. However, when it comes to distances between linear dynamical systems, almost the entire literature is exclusively devoted to distances based on the input-output description, recall e.g., the L^p, H_∞, Hankel-norm, and Gap distances [3], [4]. Perhaps one reason for this monopoly is that the geometry associated with the state-space description is *nonlinear* and complicated —compared with the rather simple linear geometry of the spaces of transfer functions.

In this paper, we introduce the (*realization*) *alignment distance*, which is defined based on the state-space description of linear systems. Recall that a linear system of *order* n has an equivalence class of state-space *realizations*, all related via the so-called *similarity action*, i.e., change of basis in the state space by the Lie *group* of nonsingular $n\times n$ matrices, $GL(n)$. Formally, one says that $GL(n)$ *acts* on the space of realizations of order n and fixed input-output dimension, and the space of systems of order n is the *quotient* space under this action (with a natural quotient *topology*). The basic idea is to find the "best" change of basis that brings given realizations of two systems "as close as possible," hence *aligning* them, and then comparing the aligned realizations. There are theoretical and computational challenges in materializing this idea, primarily stemming from *noncompactness* of $GL(n)$. For example, the mentioned quotient space is *non-metrizable*; however, it does have nice subsets, e.g., the *manifolds* of minimal, controllable, or observable systems, which descend from respective realization manifolds. Such a realization-system manifold pair forms a $GL(n)$-*principal fiber bundle*. As we will show, on these bundles, the action of $GL(n)$ can be *reduced* (in an exact *differential-geometric* sense) to the action of its subgroup of orthogonal matrices, $O(n)$. This helps us to convert our basic idea into a computationally-friendly distance that matches the quotient topology of the system manifolds.

A. A Informal Tour of our Results and the Alignment Distance

It is useful to have an example of the alignment distance early in the paper to give the reader a better sense of the developments to come and perhaps entice questions. We intentionally omit some details. Consider the manifold of asymptotically stable (a.s.), minimal, discrete-time, linear time

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.

[1] To avoid any confusion, unless otherwise indicated, by a "system" we exclusively mean a (finite-dimensional LTI) "dynamical system" as in [1], also called a "state-space system" [2].

invariant (LTI), deterministic systems of order n and input-output size (m, p), denoted by $\Sigma^{\text{min,a}}_{m,n,p}$. Take two systems $M_1, M_2 \in \Sigma^{\text{min,a}}_{m,n,p}$ with their respective minimal state-space realizations $R_i = (A_i, B_i, C_i)$ (see the state-space equation (4)). First, we transform R_i to a *balanced* realization R_i^{bl}, i.e., one for which the controllability and observability Gramians are equal, see (15). Such realizations always exist but are unique only up to an orthogonal state-space change of basis. We call this step *bundle reduction* or *standardization*. The key point is that while the full *internal symmetry group* is $GL(n)$, by the reduction step it is reduced to $O(n)$; this reduction does have an exact differential-geometric meaning (see Definition 7). The alignment distance d_F associated with the Frobenius norm $\|\cdot\|_F$ and the reduced subbundle of *balanced* realizations (see § II-D, § IV, and § V) is defined as

$$d_F^2(M_1, M_2) = \min_{Q \in O(n)} \|Q^\top A_1^{\text{bl}} Q - A_2^{\text{bl}}\|_F^2 + \|Q^\top B_1^{\text{bl}} - B_2^{\text{bl}}\|_F^2 + \|C_1^{\text{bl}} Q - C_2^{\text{bl}}\|_F^2, \tag{1}$$

where $^\top$ denotes the matrix transpose. This minimization is called the *realization alignment* problem; it is a *non-convex* problem for which, in general, no closed-form solution is known, but for moderate n it can be solved quite efficiently (see e.g., [5]). The bulk of this paper is devoted to showing (by differential-geometric tools) that (1) and a large family of such definitions are, in fact, distances that match the natural quotient topology of the respective system spaces (here, $\Sigma^{\text{min,a}}_{m,n,p}$). The term "the alignment distance" will be used in a generic sense, but it is understood that there is a family of such distances. As is obvious from this example, an alignment distance is associated with a *reduced realization subbundle* and a distance between realizations (see Definition 18).

B. A Historical Perspective

Kalman's work [6] is perhaps the first to examine the algebraic-geometric properties of the quotient spaces of minimal, observable, and controllable systems of fixed order and input-output size. It became clear that, interestingly, the notions of *observability* and *controllability* are relevant in guaranteeing that the quotient space has certain nice *topological* properties (such as being a *quasi-projective variety* in the language of algebraic-geometry and a *smooth manifold* in the language of differential geometry, see § III). The works by Hazewinkel and Kalman [7]–[10], Brockett [11], Clark [12], Krishnaprasad [13], Byrnes and Hunt [14], Delchamps [15], [16], and others furthered our understanding of the topological properties of the space of fixed order and size systems. There have been two parallel but analogous paths to analyze these quotient system spaces: the algebraic-geometric path and the differential-geometric one. Our choice, here, is the differential-geometric path, because it is more easily amenable to defining distances and computations. We should stress again that in (almost) all such works in the 1970s and 1980s no attempt was made to define distances naturally associated with the quotient structure. The only exception we have found is the works by Krishnaprasad and Martin [13], [17], where a *Riemannian metric* that can be related to the alignment distance was proposed (see § V-B1).

C. Broader Context, Significance, and Possible Applications

The spaces on which the alignment distance can be defined are the *manifolds* of systems of *fixed order* and fixed size together with appropriate matrix *rank* constraints. The rank constrains essentially ensure the manifold structure. Examples include the manifolds of minimal systems of order n and input-output size (m, p) denoted by $\Sigma^{\text{min}}_{m,n,p}$, or $\Sigma^{\text{co}}_{m,n,p}$ and $\Sigma^{\text{ob}}_{m,n,p}$ which are the manifolds of controllable and observable systems of order n and size (m, p), respectively. On the other hand, the input-output based distances are naturally defined on the *infinite dimensional linear* spaces of transfer functions or operators of fixed input-output size (m, p). By identifying a minimal system in $\Sigma^{\text{min}}_{m,n,p}$ with its transfer function matrix of McMillan degree n, an input-output distance may be used to define an induced distance on $\Sigma^{\text{min}}_{m,n,p}$. Such an induced distance is an *extrinsic* distance on $\Sigma^{\text{min}}_{m,n,p}$. The alignment distance also can be either *intrinsic* (e.g., Riemannian) or *extrinsic* (such as d_F in (1)). An extrinsic alignment distance, however, is computationally much cheaper to calculate. An important distinction with an extrinsic distance induced by an input-output distance on $\Sigma^{\text{min}}_{m,n,p}$ is that, an extrinsic alignment distance is induced from a *finite dimensional ambient space*. Another interesting point is that input-output based distances *cannot* be defined on spaces of non-minimal systems such as $\Sigma^{\text{co}}_{m,n,p}$ and $\Sigma^{\text{ob}}_{m,n,p}$, and they cannot account for the effect of the initial state of systems (at least directly).

An important question is: In what applications do the *spaces* or *families* of LTI systems of fixed order and fixed input-output size appear naturally? An immediate example is the problem of system identification, where as soon as one fixes the order of the system to be identified, the problem will be essentially a search on the manifold $\Sigma^{\text{min}}_{m,n,p}$ (see [18], [19]). In the Linear Parameter Varying (LPV) modeling of nonlinear and time-varying systems, one basically has a parameterized *family* of LTI dynamical systems of fixed order and size to model a nonlinear or time-varying system [20]. Here, one may need to *interpolate* between two systems associated with two parameter values by generating new systems of the *same* order associated with the in-between parameter values. Similarly in the framework of switched linear systems one has a *curve* (parameterized by time) on the space of systems of fixed order and size to model time-varying systems. In the framework of *multiple model control* also one might have a set systems in $\Sigma^{\text{min}}_{m,n,p}$. As a less traditional application, we also add that in problems related to modeling and classification of video sequences of human actions (or more generally, high-dimensional time-series data), MIMO LTI models of the same order and size have been used to model video sequences (see [21] and reference therein). An interesting problem, here, is the problem of *averaging* a very large number of systems of the same size and order for classification purposes. Simple Euclidean averaging of the respective transfer functions, obviously, gives a system with a *huge* order, whereas using the alignment distance (by minimizing the sum-of-the-squares-of-distances) one can get an *average* or *representative* system of the *same* order naturally (see [21], [22]). This is expected, since the alignment distance is a distance defined on the manifolds of systems of fixed order and size.

A more general framework could be to consider the space of systems of fixed input-output size but order not larger than a fixed number. A prominent example is the problem of *model order reduction*, which could be roughly phrased as: Find a system on the *boundary* of $\Sigma^{\min}_{m,n,p}$ closest (in appropriate sense) to a given system in $\Sigma^{\min}_{m,n,p}$. Note that the boundary points correspond to systems of minimal order smaller than n. As our recent work [23] shows the alignment distance can be naturally extended to formulate and solve this problem. In the current paper, however, we will be concerned only with the manifolds of systems of fixed order and size.

Our goal, in this paper, is to introduce the alignment distance as a new tool to the toolbox of control engineers with emphasis on its novelty and natural properties. Such aspects can be summarized as follows: The main feature of the alignment distance is that it is a family of distances based on the state-space description of linear systems; as such, it is a *topologically* and *methodologically* natural distance (see Remark 23). In applications where one has a set of systems of the same order and size, and the order is an important feature, the alignment distance can be potentially useful. From the example above, it is clear that, despite some differential-geometric jargon, the definition of the alignment distance is quite intuitive and even elementary. This contrasts with a distance such as the gap metric, whose starting point (the graph of a functional) is not intuitive or elementary. The alignment distance is a general framework that can be seamlessly defined for SISO/MIMO, stable/unstable, and deterministic/stochastic systems (see [24] and § VI-B). Moreover, since it is based on the state-space description, the alignment distance can potentially be extended to time-varying, LPV and switched linear system (see [25], [26]), or nonlinear systems. Finally, as a reader familiar with the *statistical analysis of shapes* [27], [28] would notice, the mathematical constructions presented in this paper are, partly, inspired by those in that field; and such a connection can be enriching to both fields.

D. Scope, Contributions, and Outline of the Paper

This paper is wholly devoted to theoretical foundations of the alignment distance. The basic idea of the alignment distance was introduced in [21], [22], [24]. This paper serves as a complete form of [22], in particular, containing proofs, detailed discussions, as well as new results. Mathematical preliminaries including group action induced distances and principal fiber bundles are introduced in § II. We try to be self-contained, but some basic familiarity with general topology, manifolds, and Lie groups is assumed. In § III, the principal fiber bundle structure of the manifolds of systems and their associated topology are reviewed. Crucial in defining the alignment distance is the (differential-geometric) step of standardization or reduction of the structure group. The most important technical results of the paper are in § IV, in which we show that several classes of realization normalization and balancing in the control literature are examples of reduction of the structure group. In § V, the alignment distance is defined and it is shown that its induced topology matches the natural quotient topology of the system manifolds. Certain discussions in relation to other problems in the literature are also included. In § VI, extensions to include the initial state and stochastic systems are considered, and § VII concludes the paper. Due to space limitation, the reader is referred to our earlier papers ([21]–[23]) and forthcoming ones for exemplary applications, e.g., in a work under preparation the robustness of internal stability under feedback in the alignment distance and its implication for model order reduction are studied.

II. Preliminaries: Group Action Induced Distances, Principal Fiber Bundles and Reduction of Structure Group

In this section, we shall review some preliminaries on group actions on topological spaces, group action induced distances, smooth principal fiber bundles, and reduction of the structure group. General references on the topics are [28]–[32].

A. Actions of Topological Groups and the Quotient Topology

The following abstract definition of a *group action* and the *orbit* or *quotient* space will be used concretely in the case of state-space realizations of systems. For completeness, certain topological facts are included in the definition.

Definition 1: Let G be a topological group and $\widetilde{\Sigma}$ a topological space. We say that G acts on $\widetilde{\Sigma}$ (from the right) if there is a continuous function $\Phi : \widetilde{\Sigma} \times G \to \widetilde{\Sigma}$ such that for $\forall R \in \widetilde{\Sigma}$ and $\forall P_1, P_2 \in G$ the following hold:

$$\begin{aligned} \Phi(R, P_1 P_2) &= \Phi(\Phi(R, P_1), P_2) \\ \Phi(R, id) &= R, \end{aligned} \quad (2)$$

where id is the identity element of G. For convenience, we denote the action by $\Phi(R, g) = g \circ R$, and may refer to $\circ$ as the action.[2] We denote the G-orbit (or the equivalence class) of $R \in \widetilde{\Sigma}$ by $[R] = \{g \circ R | \forall g \in G\}$. We denote the quotient set of the action (i.e., the set comprised of all the orbits) by $\Sigma = \widetilde{\Sigma}/G = \{[R] | R \in \widetilde{\Sigma}\}$. We may also call $\widetilde{\Sigma}$ a G-space or the top space and Σ the base or bottom space. The quotient (projection) function $\pi : \widetilde{\Sigma} \to \Sigma$ is defined as $\pi(R) = [R]$. The quotient space Σ has a unique (natural) topology in which a set $U \subset \Sigma$ is open if and only if $\pi^{-1}(U)$ is open in $\widetilde{\Sigma}$. This topology is called the quotient topology. This is the finest topology on Σ in which π is a continuous map.

The group action Φ (or $\circ$) induces an *equivalence relation* on $\widetilde{\Sigma}$, and the quotient space Σ is the set of all equivalence classes. The important topological question is what properties Σ will have or inherit from $\widetilde{\Sigma}$. In general, Σ may be a wild space, e.g., even though $\widetilde{\Sigma}$ is a metric space, hence Hausdorff, Σ may be non-Hausdorff, hence non-metrizable (i.e., its topology cannot be generated by any distance). Our main case of interest would be when $\widetilde{\Sigma}$ is a smooth manifold, G is a matrix Lie group, and Φ is a smooth action which is in addition *free* and *proper* (see §II-C for the definitions), in which case Σ will be a smooth manifold. In our setting, $\widetilde{\Sigma}$ and G both can be considered as submanifolds of certain Euclidean spaces and hence the smoothness of Φ is studied in the same way as in multivariable calculus.

[2] To be precise, and in accordance with most mathematical texts, since $\circ$ is a right action, we should write $R \circ g$. However, for convenience and aesthetics we choose $g \circ R$.

B. Metric Aspect: Group Action Induced Distances

Comparing the elements of a metric space under a group action is the basis of *pattern theory* and *statistical analysis of shapes*, and has been extensively used in image processing for registration purposes and other applications [27], [28]. The starting point is a *group-invariant* distance:

Definition 2: A distance $\tilde{d}_{\widetilde{\Sigma}}$ on a G-space $\widetilde{\Sigma}$ is called G-invariant if $\tilde{d}_{\widetilde{\Sigma}}(P \circ R_1, P \circ R_2) = \tilde{d}_{\widetilde{\Sigma}}(R_1, R_2)$ for every $P \in G$ and $R_1, R_2 \in \widetilde{\Sigma}$.

In this case G is a subgroup of the group of *isometries* of $\widetilde{\Sigma}$. Our goal is to compare elements in the quotient space (i.e., the equivalence classes). We express the general construction of a *group action induced* distance on the quotient space $\Sigma = \widetilde{\Sigma}/G$. Such a distance also is sometimes called a *quotient or orbital distance*, but we use the terminology borrowed from [28, Ch. 12]. We gather some important (known) results, hard to find in a single reference. For completeness, we give a proof in Appendix (see e.g., [28, Ch. 12], [33] for more details):

Theorem 3: Let $\widetilde{\Sigma}$ be a G-space with G-invariant distance function $\tilde{d}_{\widetilde{\Sigma}}$. Then the following hold:

1) For any $R_1, R_2 \in \widetilde{\Sigma}$, the quantity $\inf_{P \in G} \tilde{d}_{\widetilde{\Sigma}}(P \circ R_1, R_2)$ depends only on M_1 and M_2, the equivalence classes of R_1 and R_2. Define $d_\Sigma(M_1, M_2) = \inf_{P \in G} \tilde{d}_{\widetilde{\Sigma}}(P \circ R_1, R_2)$. Then d_Σ is a pseudo-distance on the quotient $\widetilde{\Sigma}/G$, i.e., it is symmetric, obeys the triangle inequality, it is positive semi-definite, but may not be positive definite.
2) If additionally the G-orbit of any $R \in \widetilde{\Sigma}$ is a closed set in $\widetilde{\Sigma}$ (with respect to the topology induced by $\tilde{d}_{\widetilde{\Sigma}}$), then $d_\Sigma(\cdot, \cdot)$ is a distance on the quotient space $\Sigma = \widetilde{\Sigma}/G$ and (Σ, d_Σ) will be a metric space whose metric induced topology coincides with its natural quotient topology.
3) If additionally G is compact or every closed and bounded set in $(\widetilde{\Sigma}, \tilde{d}_{\widetilde{\Sigma}})$ is compact, then the infimum is achieved and we write:

$$d_\Sigma(M_1, M_2) = \min_{P \in G} \tilde{d}_{\widetilde{\Sigma}}(P \circ R_1, R_2). \tag{3}$$

The topologically important fact is that a group action induced distance is *not* an arbitrary distance on the quotient space Σ (considered as a *set*), rather it is a distance that matches the unique natural quotient topology of Σ.

Part 3 with compact G is of main interest to us. The reason is that for noncompact G, guaranteeing the existence or construction of a group invariant distance $\tilde{d}_{\widetilde{\Sigma}}$ is quite complicated (see [33] and § V-B1), and additionally the closedness of the G-orbits in $\tilde{d}_{\widetilde{\Sigma}}$ needs extra provisions. These facts call for a form of *reduction* of the action of a noncompact group to a compact subgroup, something that will finally lead to the alignment distance, see § II-D, § IV, and § V.

C. Smooth Quotient Spaces and Principal Fiber Bundles

The well-known *quotient manifold theorem* states that the quotient space of a Lie group G acting smoothly, freely, and properly on a smooth manifold is a smooth manifold of dimension equal to the difference of the dimensions of the original manifold and G (see e.g., [32], [30, pp.144-150]). The freeness and properness properties are defined as follows:

Definition 4 (Free Action): Action $\circ$ on $\widetilde{\Sigma}$ is free if for every $R \in \widetilde{\Sigma}$, $P \circ R = R$ implies $P = id$.

Definition 5 (Proper Action): The smooth action $\circ$ of a Lie group G on a smooth manifold $\mathcal{M}$ is called proper if the following holds: If $\{R_i\}_i$ is a convergent sequence in $\mathcal{M}$ and $\{P_i\}_i$ is a sequence in G such that $\{P_i \circ R_i\}_i$ converges in $\mathcal{M}$, then a subsequence of $\{P_i\}_i$ converges in G.

Intuitively, properness (which only matters in the case of noncompact groups) has to do with making the quotient space a Hausdorff space and metrizable. Freeness has to do with the smoothness of the quotient space (see [32]–[34]). The smooth structure on the quotient space is determined uniquely from the top space by the quotient-taking operation.

In the case we have a smooth, proper and free action, then the pair $(\widetilde{\Sigma}, \Sigma)$ is called to form a *principal fiber bundle* with *structure group* G. In some contexts the structure group G may also be called the *symmetry group*. A principal fiber bundle may be written as $(\widetilde{\Sigma}, \Sigma, G)$, and if the context is clear we may refer to it as $(\widetilde{\Sigma}, \Sigma)$ or even $\widetilde{\Sigma}$ (i.e., the group G and the base Σ are assumed to be fixed). The set $\pi^{-1}(M) \subset \widetilde{\Sigma}$ which naturally is a closed submanifold of $\widetilde{\Sigma}$ is called the *fiber* above M. A fiber bundle is essentially a *smooth parameterized* family of objects of a fixed kind such as a fixed Lie group and a fixed vector space, respectively, in the case of principal bundles and vector bundles.

It follows from the definitions that if $(\widetilde{\Sigma}, \Sigma)$ is a principal fiber bundle with structure group G, then $\widetilde{\Sigma}$ is *locally diffeomorphic* to the product $\Sigma \times G$; that is, for every $R \in \widetilde{\Sigma}$ one could find an open neighborhood $\tilde{U} \ni R$ diffeomorphic to $\pi(\tilde{U}) \times G$. This is called *local triviality* of principal fiber bundles. However, in general, $\widetilde{\Sigma}$ is not globally diffeomorphic to $\Sigma \times G$, i.e., it is not (globally) trivial. What stops a principal fiber bundle from being trivial is a phenomenon known as *twisting*, similar to the famous Möbius band. An important fact is that $(\widetilde{\Sigma}, \Sigma)$ is trivial if and only if it admits a *smooth global section*. A global section is a smooth function $s : \Sigma \to \widetilde{\Sigma}$ such that $\pi(s(M)) = M$ for every $M \in \Sigma$, i.e., it is a function from the base space to the bundle space which *smoothly* assigns to every equivalence class a representative in the class. (Although smoothness is stronger than continuity, in this context both are treated the same, since in most cases of interest non-trivial bundles even do not admit continuous sections.) In control theory, a smooth or continuous section is known as a (smooth or continuous) *canonical form* (see Remark 11).

Remark 6 (Distances Induced by Sections): Let $(\widetilde{\Sigma}, \Sigma)$ be a trivial bundle, and assume that we are given a global continuous section $s : \Sigma \to \widetilde{\Sigma}$ and a distance (not necessarily group-invariant) $\tilde{d}_{\widetilde{\Sigma}}$ on $\widetilde{\Sigma}$. One can define a distance on Σ as: $d_\Sigma(M_1, M_2) = \tilde{d}_{\widetilde{\Sigma}}(s(M_1), s(M_2))$ for every $M_1, M_2 \in \Sigma$. We call d_Σ the distance induced by section (or canonical form) s. Since s is continuous the topology induced by d_Σ on Σ coincides with its natural quotient topology. Note that if $\tilde{d}_{\widetilde{\Sigma}}$ is G-invariant, then d_Σ might yield unrealistically large distances between points on the base space (compared with the group action induced distance, which includes an alignment or optimal positioning step). If the bundle is non-trivial (as the bundle of MIMO minimal systems is, see Remark 11 for more details), then, at best, one can find a discontinuous section

$s' : \Sigma \to \widetilde{\Sigma}$. In this case, one may want to define a "distance" as $d_{\widetilde{\Sigma}}(s'(M_1), s'(M_2))$. However, such a "distance" is useless, since it does not induce the same topology as the quotient topology on Σ. For example, at a discontinuity point of s', M, the distance between M and points in an arbitrary small open neighborhood (in the quotient topology) around M, cannot be decreased no matter how small the neighborhood is chosen.

D. Reduction of the Structure Group to the Orthogonal Group

We first define the notion of *reduction* of the structure group:

Definition 7 (Reduced or Standardized Subbundle): Let $(\widetilde{\Sigma}, \Sigma)$ be a principal fiber bundle with structure group G and let OG be a Lie subgroup of G. Assume that there exists an embedded submanifold $\widetilde{\mathcal{O}\Sigma}$ of $\widetilde{\Sigma}$, on which OG acts via the restriction of the action G on $\widetilde{\Sigma}$. If $(\widetilde{\mathcal{O}\Sigma}, \Sigma)$ is a principal bundle with structure group OG, then we call $(\widetilde{\mathcal{O}\Sigma}, \Sigma)$ or $\widetilde{\mathcal{O}\Sigma}$ a reduced, standardized, or OG- subbundle of $(\widetilde{\Sigma}, \Sigma)$, and we say that the structure group of $\widetilde{\Sigma}$ is reduced (from G) to OG. We stress that $\widetilde{\mathcal{O}\Sigma}/OG$ is not only equal to Σ as sets but also diffeomorphic to it, and we write $\widetilde{\mathcal{O}\Sigma}/OG \overset{\text{diff}}{=} \Sigma$.

We start by noting that the term "reduction of the structure group" is the standard term in differential geometry for this process. The term "standardization" is borrowed from the literature on statistical analysis of shapes [27], where a similar step is often used by which often reduction to a $SO(3)$-subbundle is achieved ($SO(3)$ being the group of rotations in $\mathbb{R}^3$).

Intuitively, if a bundle is reducible, then we can consider a smaller subbundle with a smaller structure group and still get the same base space not only as sets but also topologically and in the sense diffeomorphism. It is easy to see that for $\widetilde{\Sigma}$ to be reducible the action of G restricted to $\widetilde{\mathcal{O}\Sigma}$ must be only through OG, that is $\forall R \in \widetilde{\mathcal{O}\Sigma}, P \circ R \in \widetilde{\mathcal{O}\Sigma} \Leftrightarrow P \in OG$.

In general, reduction to an arbitrary (or small) subgroup may be impossible due to *topological obstructions*. For example, only in a trivial bundle the structure group can be reduced to the trivial subgroup $\{id\}$. However, often it is possible to reduce the *noncompact* structure group of a principal bundle to its *maximal compact* subgroup.

Proposition 8: Let OG be a closed subgroup of G and G/OG diffeomorphic to a Euclidean space, then any G-bundle $\widetilde{\Sigma}$ is reducible to an OG-bundle $\widetilde{\mathcal{O}\Sigma}$. In particular, this holds when $G = GL(n)$ and $OG = O(n)$, the subgroup of orthogonal matrices.

Proof: For the proof of the first statement see [29, p. 57, and 59]. The second statement follows from the polar decomposition of matrices, which implies that the quotient $GL(n)/O(n)$ is (diffeomorphic to) $\mathcal{S}(n)$, the manifold of $n\times n$ positive definite matrices, which in turn is diffeomorphic to the Euclidean space of $n\times n$ symmetric matrices and in the matrix exponential map and its inverse. ■

Notice that $O(n)$ is a maximal compact subgroup of $GL(n)$. Although not of our direct interest, we should mention that the essence of the (omitted) proof of this result is that, under the assumption made, $\widetilde{\Sigma}/OG$ will be a vector bundle over Σ, which contrary to a principal bundle always admits a global section; and such a global section will be mapped back to a reduced bundle (see [29, p. 57, and 59] for more details). A variant of such an explicit construction of reduced subbundles will be given in Proposition 13. As a matter of terminology, in the rest of this paper by a "reduced subbundle," we exclusively mean an $O(n)$-subbundle, since it is our only case of interest.

Remark 9 (No canonical reduction): Proposition 8 establishes existence of a reduced subbundle, however, gives no indication of uniqueness. An important point is that there is *no* reduction or subbundle that can *mathematically* be considered as natural or canonical. However, as shown in § IV, in specific applications, there may exist subbundles that for some theoretical, physical, or practical reasons stand out.

III. The Principal Bundle Structure of Manifolds of Systems of Fixed Size and Order

Our discussion will be limited to discrete-time, deterministic systems with no direct input-output path. However, essentially the same theory (with obvious modifications) applies to continuous-time systems or systems with direct input-output path. Consider the following state-space equation describing a discrete-time LTI dynamical system of order n with m inputs and p outputs (size (m, p)):

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t \\ y_t &= Cx_t, \end{aligned} \tag{4}$$

where $u_t \in \mathbb{R}^m$, $x_t \in \mathbb{R}^n$, and $y_t \in \mathbb{R}^p$ are input, state and output vectors, respectively. The triplet $R = (A, B, C) \in \widetilde{\mathcal{L}}_{m,n,p} \triangleq \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}$ is called *a* (state-space) *realization* of the system and $\widetilde{\mathcal{L}}_{m,n,p}$ the Euclidean *realization space*. For convenience, we assume zero initial state (see § VI-A for the case of non-zero intial state). Recall that, assuming zero initial state, the state-space equation (4) has an *internal symmetry*; namely, a realization $R = (A, B, C)$ and any other realization of the form

$$P \circ R \triangleq (P^{-1}AP, P^{-1}B, CP), \tag{5}$$

where $P \in GL(n)$, have the same input-output behavior. The group $GL(n)$ is sometimes called the internal symmetry group of linear dynamical systems [35]. The transformation $\circ$ (sometimes called the *similarity action* or *transformation*) corresponds to a change of basis in the state space of the form $x \to P^{-1}x$. It is easy to verify that $\circ$ in (5) defines a *right* smooth action by $GL(n)$ on the space of realizations $\widetilde{\mathcal{L}}_{m,n,p}$ (in its standard Euclidean topology and smooth structure). The action $\circ$ induces an equivalence relation on $\widetilde{\mathcal{L}}_{m,n,p}$, where the equivalence class of a realization $R \in \widetilde{\mathcal{L}}_{m,n,p}$ is

$$M = [R] = \{P \circ R | P \in GL(n)\}. \tag{6}$$

We call M a *system*; that is, a system is an equivalence class of realizations under the action (5). We call this definition of systems the *state-space based* or *internal* definition of a system. We call the quotient $\mathcal{L}_{m,n,p} = \widetilde{\mathcal{L}}_{m,n,p}/GL(n)$ the *space of systems* of order n and input-output size (m, p). The action $\circ$ is neither free nor proper on the *entire* of $\widetilde{\mathcal{L}}_{m,n,p}$; but we can pass to appropriate submanifolds of $\widetilde{\mathcal{L}}_{m,n,p}$, where $\circ$ is free and proper, and get quotient manifolds.

A. Rank Conditions and the Principal Bundle Structure

We establish some notation first. For a realization $R = (A, B, C) \in \widetilde{\mathcal{L}}_{m,n,p}$, denote by $\mathcal{C}_k = [B, AB, \ldots, A^{k-1}B]$ and $\mathcal{O}_k = [C^\top, (CA)^\top, \ldots, (CA^{k-1})^\top]^\top$, respectively, its controllability and observability matrices of order k ($n \le k \le \infty$). These matrices are realization-dependent, and under the similarity action (5) they transform as: $\mathcal{C}_k \to P^{-1}\mathcal{C}_k$ and $\mathcal{O}_k \to \mathcal{O}_k P$. Let $\widetilde{\Sigma}^{\text{co}}_{m,n,p}$ and $\widetilde{\Sigma}^{\text{ob}}_{m,n,p}$ denote, respectively, the subsets of controllable and observable realizations in $\widetilde{\mathcal{L}}_{m,n,p}$. As a convention, we always denote a realization space by $\widetilde{}$ and the corresponding system space (i.e., quotient space) by dropping the $\widetilde{}$ (the symmetry group will be clear from the context). For example, the controllable and observable system spaces are denoted by $\Sigma^{\text{co}}_{m,n,p} \triangleq \widetilde{\Sigma}^{\text{co}}_{m,n,p}/GL(n)$ and $\Sigma^{\text{ob}}_{m,n,p} \triangleq \widetilde{\Sigma}^{\text{ob}}_{m,n,p}/GL(n)$. Let $\widetilde{\Sigma}^{\min}_{m,n,p} \triangleq \widetilde{\Sigma}^{\text{co}}_{m,n,p} \cap \widetilde{\Sigma}^{\text{ob}}_{m,n,p}$ and $\Sigma^{\min}_{m,n,p} \triangleq \widetilde{\Sigma}^{\min}_{m,n,p}/GL(n)$ denote the subspace of minimal realizations and the corresponding system space. We denote by $\widetilde{\Sigma}^{\text{tC}}_{m,n,p}$ the set of realizations R where $\text{rank}(C) = n$. Here, we obviously need to assume that $p \ge n$ (tC stands for "tall C"). Such realizations appear as realizations of the so-called *tall* transfer functions in modeling high-dimensional time-series in econometrics as well as in modelling video sequence data (see e.g., [21], [36] and references therein). We use the superscript $^{\text{a}}$ over a set to denote its subset of a.s. realizations, e.g., $\widetilde{\Sigma}^{\text{co,a}}_{m,n,p}$ denotes the subset of a.s. realizations in $\widetilde{\Sigma}^{\text{co}}_{m,n,p}$. Finally, we use the superscript $^{\text{mp}}$ to denote a respective subset of minimum-phase realizations. A minimum-phase realization is one whose transfer function matrix (a $p \times m$ matrix) is of full rank everywhere outside the unit circle in the complex plane. For example, $\widetilde{\Sigma}^{\text{min,a,mp}}_{m,n,p}$ denotes the set of minimal, a.s., and minimum-phase realizations in $\widetilde{\mathcal{L}}_{m,n,p}$.

All the above realization subspaces are, in fact, *open* subsets of $\widetilde{\mathcal{L}}_{m,n,p}$, hence its submanifolds of dimension $n^2+nm+np$. The basic reason is that they are defined essentially by inequality constraints on continuous functions (e.g., matrix determinants, see also [37]). The next theorem shows that their corresponding system spaces are smooth manifolds:

Theorem 10: Let $\widetilde{\Sigma}_{m,n,p}$ be one of the manifolds $\widetilde{\Sigma}^{\text{ob}}_{m,n,p}$, $\widetilde{\Sigma}^{\text{co}}_{m,n,p}$, $\widetilde{\Sigma}^{\min}_{m,n,p}$, $\widetilde{\Sigma}^{\text{tC}}_{m,n,p}$, $\widetilde{\Sigma}^{\text{tC,co}}_{m,n,p}$, their a.s. or minimum phase submanifolds, and let $\Sigma_{m,n,p}$ be the respective quotient (system) space under the $GL(n)$ action (5). Then the followings hold:

1) $GL(n)$ acts smoothly, properly and freely on $\widetilde{\Sigma}_{m,n,p}$;
2) $\Sigma_{m,n,p}$ is a smooth $n(m+p)$-dimensional manifold and the realization-system space pair $(\widetilde{\Sigma}_{m,n,p}, \Sigma_{m,n,p})$ forms a smooth $GL(n)$ principal fiber bundle. Here, the smooth topological structure is the natural one induced by $\mathcal{L}_{m,n,p}$.

Proof: We only prove the result for $\widetilde{\Sigma}^{\text{ob}}_{m,n,p}$, the rest are similar. In view of the quotient manifold theorem, we just need to prove the first part. To see that action is free, let $R \in \widetilde{\Sigma}^{\text{ob}}_{m,n,p}$, and $P \circ R = R$ for some $P \in GL(n)$. It follows that $\mathcal{O}P = \mathcal{O}$, where $\mathcal{O}$ is the observability matrix of order n. This implies freeness of the action, because we must have $P = I_n$ since $\mathcal{O}$ is full rank.

To see properness, assume that $\{R_i\}_i$ is a sequence in $\widetilde{\Sigma}^{\text{ob}}_{m,n,p}$ converging to $R \in \widetilde{\Sigma}^{\text{ob}}_{m,n,p}$, and $\{P_i \circ R_i\}_i$ with $P_i \in GL(n)$ is converging to $\bar{R} \in \widetilde{\Sigma}^{\text{ob}}_{m,n,p}$. We need to show that there is a subsequence of $\{P_i\}_i$ converging in $GL(n)$. Since $\widetilde{\Sigma}^{\text{ob}}_{m,n,p}$ is a submanifold of $\widetilde{\mathcal{L}}_{m,n,p}$, without loss of generality we can assume that all these convergences are in the standard Euclidean distance. Next, with some abuse of notation, let $\mathcal{O}$, $\mathcal{O}_i$, and $\bar{\mathcal{O}}$ denote the observability matrices of order n corresponding to realizations R, R_i, and $\bar{R}$, respectively. These assumptions on the realizations, in an obvious way, translate to corresponding ones about the observability matrices: $\{\mathcal{O}_i\}_i$ converges to $\mathcal{O}$ and $\{\mathcal{O}_i P_i\}_i$ converges $\bar{\mathcal{O}}$. Let $P_i = U_i \Lambda_i V_i^\top$ be an SVD of P_i. We have $\|\mathcal{O}_i U_i \Lambda_i V_i^\top - \bar{\mathcal{O}}\|_F = \|\mathcal{O}_i U_i \Lambda_i - \bar{\mathcal{O}} V_i\|_F \to 0$ as $i \to \infty$. Since $O(n)$ is compact, $\{(U_i, V_i)\}_i$ has a converging subsequence (which we re-index as $\{(U_i, V_i)\}_i$) with limit (U, V). Thus we can write $\|\mathcal{O}_i U_i \Lambda_i - \bar{\mathcal{O}} V\|_F \to 0$ and $\mathcal{O}_i U_i \to \mathcal{O} U$ as $i \to \infty$. Since $\mathcal{O}U$ is full-rank and bounded, for large i, the norm of every column of $\mathcal{O}_i U_i$ will be larger than ϵ and smaller than η, for some $\epsilon, \eta > 0$. Now, if any diagonal element of Λ_i tends to zero or infinity the corresponding column in $\mathcal{O}_i U_i \Lambda_i$ will tend to zero or infinity, which contradicts $\|\mathcal{O}_i U_i \Lambda_i - \bar{\mathcal{O}} V\|_F \to 0$ (recall that $\bar{\mathcal{O}} V$ is full-rank too, and, in particular, none of its columns is zero). Thus the diagonal elements of Λ_i must remain bounded and away from zero. Hence, $\{\Lambda_i\}_i$ must have a converging subsequence (re-indexed $\{\Lambda_i\}_i$) with limit Λ (nonsingular). Therefore, as $i \to \infty$, $U_i \Lambda_i V_i^\top \to U \Lambda V^\top \in GL(n)$, which means that a subsequence of $\{P_i\}_i$ converges in $GL(n)$. ■

The second part of this theorem (at least for controllable, observable, or minimal systems) is a standard result, although the proof (based on properness and freeness of the action) we gave here is hard to find in the literature.

The main requirement in Theorem 10 is the existence of at least one rank-n matrix that transforms via the action (5). Thus the result holds for other submanifolds not mentioned in the statement, e.g., $\widetilde{\Sigma}^{\text{tC,co,a,mp}}_{m,n,p}$. It is interesting to mention that $\widetilde{\Sigma}^{\text{min,a}}_{m,n,p}$ and $\widetilde{\Sigma}^{\min}_{m,n,p}$ are diffeomorphic, and both are diffeomorphic to their continuous-time counterparts [38], [39]. Since Theorem 10 and many results in the rest of the paper apply to numerous submanifolds of $\widetilde{\mathcal{L}}_{m,n,p}$, it is convenient to use $(\widetilde{\Sigma}_{m,n,p}, \Sigma_{m,n,p})$ to denote a generic realization-system $GL(n)$-principal fiber bundle or briefly a *realization bundle* (a terminology borrowed from [16]).

Remark 11 (On the Topology): The topology and parameterization of manifolds of linear dynamical have been studied rather extensively in the literature, especially motivated by the problem of system identification see e.g, [8], [9], [11], [12], [14], [16], [40]–[44]. An important negative result is that the realization bundles $\widetilde{\Sigma}^{\min}_{m,n,p}$, $\widetilde{\Sigma}^{\text{co}}_{m,n,p}$, and $\widetilde{\Sigma}^{\text{ob}}_{m,n,p}$ are non-trivial, hence do not admit global *continuous* canonical forms unless, respectively, $\min(m, p) = 1$, $m = 1$, and $p = 1$. A rigorous proof of this practically known fact ended the search for continuous global canonical forms for MIMO system identification in the 1970s (see [8], [9], [14] for proofs and discussions). For us, the most relevant fact is that almost all the system manifolds of interest have complicated non-Euclidean topologies and do *not* admit *global* parameterizations. Thus, in order to define distances that respect the topologies of the system manifolds and are based on *comparing realizations*, one has to resort to local *overlapping* canonical forms and *switching* between, which is quite complicated. In contrast, the alignment distance which is still based on comparing realizations, does not use overlapping canonical forms, i.e., does not fix a realization; rather the comparison is through an optimization that finds a closest realization of one system to a realization of another system.

IV. Examples of Reduction of Realization Bundles: Normalization and Balancing

In the control literature, to the best of our knowledge, the notion of reducing the structure group (in its exact sense with its differential geometric significance) has not received much attention with the only exception being the works of Delchamps [15], [16], [45]. Delchamps has used three specific reductions of the structure group (from $GL(n)$ to $O(n)$) in order to define Riemannian metrics on the so-called abstract *state bundle* and *connections* on the realization bundle to prove certain global properties of system identification algorithms. Such tools were primarily of theoretical interest as opposed to computational. As we show in this section, several notions of *realization* or *Gramian balancing* introduced in the literature (e.g., [46], [47], [48], [49]) are closely related to reduction of the structure group. We will distinguish between balancing and the more often used but restricted *diagonal balancing*, since, as will be seen shortly, the latter does not possess any useful differential geometric meaning.

A. Realization Bundle Reduction Maps

A relatively simple technical tool exists to verify if a given subset of a $GL(n)$-principal bundle is a reduced subbundle (i.e., an $O(n)$-subbundle). To this end, we follow [30, p.149] and first define the notion of an *equivariant map*:

Definition 12 (Equivariant Maps): Let X and Y be two manifolds on which the Lie group G acts from the right. Denote the actions by $\circ_X$ and $\circ_Y$, respectively. We say that the smooth map $f : X \to Y$ is an equivariant map if $f(P \circ_X R) = P \circ_Y f(R)$ for every $P \in G$ and $R \in X$.

Proposition 13: Let $(\widetilde{\Sigma}_{m,n,p}, \Sigma_{m,n,p})$ be a realization-system bundle (e.g., as in Theorem 10). Let $\mathcal{S}(n)$ denote the manifold of $n \times n$ positive definite matrices. Consider the natural smooth right action of $GL(n)$ on $\mathcal{S}(n)$ via $S \mapsto P^\top SP$ for $\forall P \in GL(n)$ and $S \in \mathcal{S}(n)$. Assume that there exists a smooth equivariant map $\nu : \widetilde{\Sigma}_{m,n,p} \to \mathcal{S}(n)$, that is for every $P \in GL(n)$ and $R \in \widetilde{\Sigma}_{m,n,p}$ one has

$$\nu(P \circ R) = P^\top \nu(R) P. \tag{7}$$

Then $\widetilde{\mathcal{O}\Sigma}_{m,n,p} = \nu^{-1}(I_n)$ is an $O(n)$-subbundle or reduced subbundle of $\widetilde{\Sigma}_{m,n,p}$ of dimension $\frac{n(n-1)}{2} + n(m+p)$. (Here, $\nu^{-1}(I_n)$ is the preimage of I_n under ν.)

Proof: First, note that given $R \in \widetilde{\Sigma}_{m,n,p}$, by choosing P equal to the inverse of the (unique) symmetric square root of $\nu(R) \in \mathcal{S}(n)$, we get $\nu(P \circ R) = I_n$ from (7); thus, any system $M \in \Sigma_{m,n,p}$ has a realization $R \in \widetilde{\Sigma}_{m,n,p}$ with $\nu(R) = I_n$. This means that the fiber over every system intersects $\nu^{-1}(I_n)$. Next, note that $R \in \nu^{-1}(I_n)$ and $P \circ R \in \nu^{-1}(I_n)$ if and only if $P \in O(n)$, which means that the action of $GL(n)$ is through $O(n)$. These two facts imply that the quotient $\widetilde{\mathcal{O}\Sigma}_{m,n,p}/O(n)$ is equal to $\Sigma_{m,n,p}$ as a set. Next, we show that $\widetilde{\mathcal{O}\Sigma}_{m,n,p} = \nu^{-1}(I_n)$ is an embedded submanifold of $\widetilde{\Sigma}_{m,n,p}$, which would imply that $\widetilde{\mathcal{O}\Sigma}_{m,n,p}/O(n)$ is smoothly embedded in $\Sigma_{m,n,p}$ as a smooth manifold, and therefore $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ is an $O(n)$-reduced subbundle of $\widetilde{\Sigma}_{m,n,p}$. That $\nu^{-1}(I_n)$ is an embedded submanifold, follows from the constant rank level set theorem [32, p. 182]; to apply that theorem it suffices to show that $\nu_{*|_R} : T_R\widetilde{\Sigma}_{m,n,p} \to T_{\nu(R)}\mathcal{S}(n)$, the derivative (tangent map) of ν at $R \in \widetilde{\Sigma}_{m,n,p}$, is of constant rank for every R. The key observation is that $\nu^{||}_{*|_R}$, the derivative of ν along the fiber, is of rank $\frac{n(n+1)}{2}$, which is the maximum possible and equal to the dimension of $\mathcal{S}(n)$, which in turn would imply that ν is a submersion and the dimension of $\nu^{-1}(I_n)$ is as claimed. To see that $\nu^{||}_{*|_R}$ is of rank $\frac{n(n+1)}{2}$, we identify the fiber at R with $GL(n)$, thus from $\nu(P \circ R) = P^\top \nu(R) P$, we can write $\nu^{||}_{*|_R}(X) = X^\top \nu(R) P + P^\top \nu(R) X$, where $X \in \mathbb{R}^{n\times n}$ is a tangent vector along the fiber at R. Clearly, the kernel of $\nu^{||}_{*|_R}$ is $\{X \in \mathbb{R}^{n\times n} | P^\top \nu(R) X = \text{skew symmetric}\}$ and is of dimension $\frac{n(n-1)}{2}$. Thus the rank of $\nu^{||}_{*|_R}$ is $\frac{n(n+1)}{2}$, which is the same as the dimension of $\mathcal{S}(n)$; and this allows us to apply the constant rank level set theorem.

■

The reader might wonder why $\mathcal{S}(n)$ should appear in this result. The answer is in the polar decomposition of matrices and Proposition 8 (also recall our discussion following that proposition). Indeed, the above proof can be considered as a half-way constructive proof of Proposition 8 (for $GL(n)$ and $O(n)$); "half-way" because we are not establishing the existence of ν. However, this is not an issue, because as to seen soon, the existence of numerous instances of ν comes immediately from control theory.

For ease of reference we define:

Definition 14 ((Realization Bundle) Reduction Map): We call the map ν as in Proposition 13 a realization bundle reduction map or simply a reduction map (on $\widetilde{\Sigma}_{m,n,p}$). We denote a generic reduced realization subbundle as $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$.

1) Geometric Mean of Bundle Reduction Maps

Proposition 13 gives an easy way to construct a reduced subbundle when a reduction map is available; however, it is silent on how to construct such a reduction map. Here, we introduce a method to construct a *new* reduction map from old ones, which is closely related to realization or Gramian balancing. First, note that if ν_1 and ν_2 are reduction maps, so is $\nu_1 + \nu_2$, as well as their average $\frac{1}{2}(\nu_1 + \nu_2)$. Their matrix product, however, is not necessarily a reduction map. But, interestingly, a form of their *geometric* mean or average (specifically their Riemannian mean) will be. We briefly describe the construction. For more complete exposition the reader is referred to [50]–[52], [53, Ch. XII]. The manifold $\mathcal{S}(n)$ is equipped with a natural Riemannian metric of non-positive curvature (see [53, Ch. XII] and [51] for details). In this metric, the geodesic $t \mapsto S(t)$ starting from $S(0) = S_1$ to $S(1) = S_2$ can be expressed as $S(t) = S_1^{1/2} e^{t \log(S_1^{-1/2} S_2 S_1^{-1/2})} S_1^{1/2}$ and the corresponding distance is $d_{\mathcal{S}(n)}(S_1, S_2) = \| \log(S_1^{-1/2} S_2 S_1^{-1/2}) \|_F$ [51], [53, Ch. XII], where e^X and $\log(X)$ denote the standard matrix exponential and logarithm of X, and $S^{\frac{1}{2}} \equiv \sqrt{S}$ is the unique (matrix) square root of $S \in \mathcal{S}(n)$. The *Riemannian center of mass*, *Riemannian mean or average*, or simply the *geometric mean* of a set of points $\{S_i\}_{i=1}^N \subset \mathcal{S}(n)$, denoted by $\bar{S}$, is defined as the global minimizer of $S \mapsto \sum_i d^2_{\mathcal{S}(n)}(S_i, S)$. It can be shown that the geometric mean is *unique* and depends *smoothly* on the points S_i [50]–[52]. The geometric mean has several interesting properties, of which the most relevant to us is *congruence invariance*: if $\bar{S}$ is the geometric mean of $\{S_i\}_{i=1}^N$, then $P^\top \bar{S} P$ is the mean of $\{P^\top S_i P\}_{i=1}^N$ for any $P \in GL(n)$ [51]. For two pints S_1 and S_2, the average is simply the *midpoint* on the geodesic connecting the two points, i.e., $\bar{S} = S(\frac{1}{2})$, and it is the unique solution to the equation: $\bar{S} S_1^{-1} \bar{S} = S_2$ (see e.g., [51]). For reasons to become clear soon, we call this equation a *balancing* equation. Notice that this equation is symmetric with respect to S_1 and S_2. For $\mathcal{S}(1) = \mathbb{R}^+$, we get $\bar{S} = \sqrt{S_1 S_2}$, which is the usual geometric mean of S_1 and S_2. We add that, alternatively, the uniqueness and smoothness of the solution to the balancing equation can be shown by noting that it is, indeed, a continuous-time algebraic Riccati equation for which uniqueness and smoothness results are available (see e.g., [45]).

The next proposition shows that the geometric mean of two reduction maps is a reduction map, and it will be the basis for bundle reduction based on *realization (or Gramian) balancing*:

Proposition 15 (Geometric Mean of Reduction Maps): Let $\nu_1, \nu_2 : \widetilde{\Sigma}_{m,n,p} \to \mathcal{S}(n)$ be two reduction maps. Then the following hold: (i) the (Riemannian) geometric mean of $\nu_1(R)$ and $\nu_2(R)$ (also called the balancing reduction map associated with ν_1 and ν_2), denoted by $R \mapsto \bar{\nu}(R)$, is a (smooth) reduction map, and is the unique solution of the balancing equation

$$\bar{\nu}(R)\nu_1(R)^{-1}\bar{\nu}(R) = \nu_2(R). \tag{8}$$

(ii) For every $R \in \widetilde{\Sigma}_{m,n,p}$, let $P \in GL(n)$ be a solution to

$$\nu_1(P \circ R)^{-1} = \nu_2(P \circ R). \tag{9}$$

Then P is of the form $P = \bar{\nu}(R)^{-1/2}\Theta$ with $\Theta \in O(n)$, and $P \circ R \in \bar{\nu}^{-1}(I_n)$. In particular, $P = \bar{\nu}(R)^{-1/2}$ is the positive definite balancing transformation.

Proof: (i) The fact that (8) has a unique solution $\bar{\nu}(R)$ in $\mathcal{S}(n)$ which depends smoothly on R follows from our preceding discussion about geometric means. We just need to show that $R \mapsto \bar{\nu}(R)$ is an equivariant map. But that immediately follows from the congruence invariance property of the geometric mean, because $\bar{\nu}(P \circ R)$ is nothing but the geometric mean of $P^\top \nu_1(R) P$ and $P^\top \nu_2(R) P$.

(ii) If P solves (9), then $(PP^\top)^{-1}\nu_1(R)^{-1}(PP^\top)^{-1} = \nu_2(R)$, hence $(PP^\top)^{-1} = \bar{\nu}(R)$, and the stated form of P follows immediately. Note that $R \in \bar{\nu}^{-1}(I_n)$ iff $\nu_1^{-1}(R) = \nu_2(R)$; thus if P solves (9), then $P \circ R \in \bar{\nu}^{-1}(I_n)$. ■

This proposition essentially holds true for the geometric mean of more than two reduction maps, except that the respective balancing equation would different.

B. Reduction via Normalization and Balancing of Gramians

Given a realization $R = (A, B, C) \in \widetilde{\mathcal{L}}_{m,n,p}$, let us denote the observability and controllability Gramians of order $k \geq n$ by $W_{\mathrm{o},k} = \mathcal{O}_k^\top \mathcal{O}_k$ and $W_{\mathrm{c},k} = \mathcal{C}_k \mathcal{C}_k^\top$, respectively. For $k = \infty$, we write $W_{\mathrm{c},\infty} \equiv W_{\mathrm{c}}$ and $W_{\mathrm{o},\infty} \equiv W_{\mathrm{o}}$, in which case we assume that A is a.s. Under the similarity action (5), these matrices transform as: $W_{\mathrm{c},k} \to P^{-1} W_{\mathrm{c},k} P^{-\top}$ and $W_{\mathrm{o},k} \to P^\top W_{\mathrm{o},k} P$. This means that the maps $R \mapsto \nu_{\mathrm{co},k}(R) = W_{\mathrm{c},k}^{-1}$ and $R \mapsto \nu_{\mathrm{ob},k}(R) = W_{\mathrm{o},k}$ are reduction maps (recall Definition 14). The realization bundle can be taken as various other related bundles such as $\widetilde{\Sigma}^{\mathrm{min}}_{m,n,p}$ or $\widetilde{\Sigma}^{\mathrm{min,a}}_{m,n,p}$ (if $k = \infty$). In the case of the manifold of tall systems $\widetilde{\Sigma}^{\mathrm{tC}}_{m,n,p}$, $R \mapsto \nu_C(R) = C^\top C$ is a reduction map.

The simplest form of reduction is what we call *normalization*, which refers to making a Gramian equal to identity. For example, the *input-normalized* realization subbundle is defined by constraining the controllability Gramian $W_{\mathrm{c},k}$ to be the identity matrix. Such a subbundle can be defined on the bundle of controllable realizations, i.e.,

$$\widetilde{\mathcal{O}\Sigma}^{\mathrm{co,in},k}_{m,n,p} = \{R \in \widetilde{\Sigma}^{\mathrm{co}}_{m,n,p} | W_{\mathrm{c},k} = I_n\}. \tag{10}$$

To see that $\widetilde{\mathcal{O}\Sigma}^{\mathrm{co,in},k}_{m,n,p}$ is a reduced subbundle of $\widetilde{\Sigma}^{\mathrm{co}}_{m,n,p}$, in view of Proposition 13, we just need to note that $\widetilde{\mathcal{O}\Sigma}^{\mathrm{co,in},k}_{m,n,p} = \nu_{\mathrm{co},k}^{-1}(I_n)$. It is useful to stress that by the very definition of a reduced subbundle the quotient space $\widetilde{\mathcal{O}\Sigma}^{\mathrm{co,in},k}_{m,n,p}/O(n)$ is diffeomorphic to $\Sigma^{\mathrm{co}}_{m,n,p}$. Obviously, the subbundle of input-normalized realizations can be defined on the realization bundle of controllable and a.s. realizations:

$$\widetilde{\mathcal{O}\Sigma}^{\mathrm{co,a,in}}_{m,n,p} = \{R \in \widetilde{\Sigma}^{\mathrm{co,a}}_{m,n,p} | W_{\mathrm{c}} = I_n\}. \tag{11}$$

Similarly, it can be defined for minimal realizations $\widetilde{\Sigma}^{\mathrm{min}}_{m,n,p}$ or a.s. minimal realizations $\widetilde{\Sigma}^{\mathrm{min,a}}_{m,n,p}$, in which case are denoted by $\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,in},k}_{m,n,p}$ and $\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,a,in}}_{m,n,p}$, respectively. In the case of controllable tall realizations $\widetilde{\Sigma}^{\mathrm{tC,co}}$ and $\widetilde{\Sigma}^{\mathrm{tC,co,a}}$, the subbundles $\widetilde{\mathcal{O}\Sigma}^{\mathrm{tC,co,in},k}_{m,n,p}$ and $\widetilde{\mathcal{O}\Sigma}^{\mathrm{tC,co,a,in}}_{m,n,p}$ can be defined in obvious ways. Subbundles of *output-normalized* realizations can be defined by restricting the observability Gramian to the identity matrix I_n. For example, we can define

$$\widetilde{\mathcal{O}\Sigma}^{\mathrm{co,on},k}_{m,n,p} = \{R \in \widetilde{\Sigma}^{\mathrm{ob}}_{m,n,p} | W_{\mathrm{o},k} = I_n\}, \tag{12}$$

which is a subbundle of the bundle of observable realizations $\widetilde{\Sigma}^{\mathrm{ob}}_{m,n,p}$. In the same manner, the reduced subbundles $\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,on},k}_{m,n,p}$ and $\widetilde{\mathcal{O}\Sigma}^{\mathrm{min,a,on}}_{m,n,p}$ can be defined, which are, respectively, subbundles of $\widetilde{\Sigma}^{\mathrm{min}}_{m,n,p}$ and $\widetilde{\Sigma}^{\mathrm{min,a}}_{m,n,p}$. Another form of output normalization in the case of the bundle of tall realizations $\widetilde{\Sigma}^{\mathrm{tC}}_{m,n,p}$ is

$$\widetilde{\mathcal{O}\Sigma}^{\mathrm{tC,nC}}_{m,n,p} = \{R \in \widetilde{\Sigma}^{\mathrm{tC}}_{m,n,p} | C^\top C = I_n\}, \tag{13}$$

for which we can use the reduction map ν_C to show that it is a reduced subbundle of $\widetilde{\Sigma}^{\mathrm{tC}}_{m,n,p}$. Similar to the case of observable or minimal realizations, subbundles such as $\widetilde{\mathcal{O}\Sigma}^{\mathrm{tC,a,nC}}_{m,n,p}$ or $\widetilde{\mathcal{O}\Sigma}^{\mathrm{tC,co,a,nC}}_{m,n,p}$ can be defined in obvious ways.

Once one has two reduction maps, then one can combine the two maps through the step of geometric averaging via the balancing equation as in Proposition 15 and (8) to get a new reduction map or reduced subbundle. An immediate example is the controllability and observability reduction maps $\nu_{\mathrm{co},k}$ and $\nu_{\mathrm{ob},k}$. In view of Proposition 15, let ν_{bl} be their geometric mean or as we called the balancing reduction map associated with $\nu_{\text{co},k}$ and $\nu_{\text{ob},k}$. Then consider $\nu_{\text{bl},k}^{-1}(I_n) \subset \widetilde{\Sigma}^{\min}_{m,n,p}$. For any realization $R \in \nu_{\text{bl},k}^{-1}(I_n)$, we have $\nu_{\text{ob},k}(R)^{-1} = \nu_{\text{co},k}(R)$, which implies that $W_{\text{c},k} = W_{\text{o},k}$, i.e., for such a realization the controllability and observability Gramians are equal. Thus, we define the reduced subbundle of k-balanced realizations as

$$\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p} = \{R \in \widetilde{\Sigma}^{\min}_{m,n,p} | W_{\text{c},k} = W_{\text{o},k}\}. \quad (14)$$

A similar argument shows that the subbundle of a.s. balanced minimal realizations

$$\widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,n,p} = \{R \in \widetilde{\Sigma}^{\text{min,a}}_{m,n,p} | W_{\text{c}} = W_{\text{o}}\}, \quad (15)$$

is a reduced subbundle of $\widetilde{\Sigma}^{\text{min,a}}_{m,n,p}$. From part (ii) of Proposition 15, for a given $R \in \widetilde{\Sigma}^{\min}_{m,n,p}$, the positive definite balancing transformation P is the inverse of the square root of the geometric mean of $W_{\text{o},k}$ and $W_{\text{c},k}^{-1}$, which can be expressed explicitly as mentioned in § IV-A1 (see also [54, p. 232] where the explicit formula is given but not in terms geometric means).

Balanced realizations have been of interest in the literature due to their desirable or physically-meaningful properties (see e.g., [46]–[48], [54]). In fact, Moore's main idea for using diagonally balanced realizations in [47], was the fact that the numerical condition numbers of the Gramians W_{c} and W_{o} can diverge significantly under the similarity action. Thus, it was argued that the best situation to gain information about minimality of a system (i.e., its level of simultaneous observability and controllability) is to consider a realization of the system for which both the Gramians are equal (see Proposition 9 in [47]). Moore's extra step of simultaneous diagonalization of the (equal) Gramians by an orthogonal matrix was mainly devised to facilitate model order reduction. The reader is referred to [23], where model order reduction in the alignment distance is shown to be an enhanced version of Moore's diagonally balanced truncation.

Interestingly, a large class of balancing transformations appear as optimal solutions to certain *variational* problems [46], [48], [54], [55]. For example, given a realization $R \in \widetilde{\Sigma}^{\min}_{m,n,p}$ consider the function $h_{\text{bl}} : GL(n) \to \mathbb{R}$ defined as

$$h_{\text{bl}}(P; R) = \text{tr}\big(P^{-1} W_{\text{c},k} P^{-\top} + P^\top W_{\text{o},k} P\big). \quad (16)$$

Notice that $h_{\text{bl}}(\Theta P; R) = h_{\text{bl}}(P; R)$ for every $\Theta \in O(n)$, i.e., h_{bl} is constant on $O(n)$. It can be shown that there is a P which is unique up to a right orthogonal factor and solves $\min_{P \in GL(n)} h_{\text{bl}}(R; P)$, and such a solution satisfies the balancing equation $P^{-1} W_{\text{c},k} P^{-\top} = P^\top W_{\text{o},k} P$ [46].

For the bundle of controllable tall realizations $\widetilde{\Sigma}^{\text{tC,co}}_{m,n,p}$, we can define the C$k$-balanced realization subbundle as

$$\widetilde{\mathcal{O}\Sigma}^{\text{tC,co,Cbl},k}_{m,n,p} = \{R \in \widetilde{\Sigma}^{\text{tC,co}}_{m,n,p} | C^\top C = W_{\text{c},k}\}, \quad (17)$$

and in the case of the bundle of a.s. controllable tall realizations $\widetilde{\Sigma}^{\text{tC,co,a}}_{m,n,p}$ we define

$$\widetilde{\mathcal{O}\Sigma}^{\text{tC,co,a,Cbl}}_{m,n,p} = \{R \in \widetilde{\Sigma}^{\text{tC,co,a}}_{m,n,p} | C^\top C = W_{\text{c}}\}. \quad (18)$$

Again, using Proposition 15 with ν_C and $\nu_{\text{co},k}$ (and ν_{co} in the case of $\widetilde{\Sigma}^{\text{tC,co,a}}_{m,n,p}$) we see that $\widetilde{\mathcal{O}\Sigma}^{\text{tC,co,Cbl},k}_{m,n,p}$ and $\widetilde{\mathcal{O}\Sigma}^{\text{tC,co,a,Cbl}}_{m,n,p}$ are reduced subbundles of $\widetilde{\Sigma}^{\text{tC,co}}_{m,n,p}$ and $\widetilde{\Sigma}^{\text{tC,a,co}}_{m,n,p}$, respectively.

We summarize the forgoing results:

Theorem 16: The following hold:

1) $\widetilde{\mathcal{O}\Sigma}^{\text{co,in},k}_{m,n,p}$ (resp. $\widetilde{\mathcal{O}\Sigma}^{\text{co,a,in}}_{m,n,p}$) is a standardized or reduced subbundle of $\widetilde{\Sigma}^{\text{co}}_{m,n,p}$ (resp. $\widetilde{\Sigma}^{\text{co,a}}_{m,n,p}$).
2) $\widetilde{\mathcal{O}\Sigma}^{\text{ob,on},k}_{m,n,p}$ (resp. $\widetilde{\mathcal{O}\Sigma}^{\text{ob,a,on}}_{m,n,p}$) is a standardized or reduced subbundle of $\widetilde{\Sigma}^{\text{ob}}_{m,n,p}$ (resp. $\widetilde{\Sigma}^{\text{ob,a}}_{m,n,p}$).
3) The following are standardized or reduced subbundles of $\widetilde{\Sigma}^{\min}_{m,n,p}$ (resp. $\widetilde{\Sigma}^{\text{min,a}}_{m,n,p}$): $\widetilde{\mathcal{O}\Sigma}^{\text{min,in},k}_{m,n,p}$, $\widetilde{\mathcal{O}\Sigma}^{\text{min,on},k}_{m,n,p}$, and $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$ (resp. $\widetilde{\mathcal{O}\Sigma}^{\text{min,a,on}}_{m,n,p}$, $\widetilde{\mathcal{O}\Sigma}^{\text{min,a,in}}_{m,n,p}$, and $\widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,n,p}$).
4) $\widetilde{\mathcal{O}\Sigma}^{\text{tC,nC}}_{m,n,p}$ (resp. $\widetilde{\mathcal{O}\Sigma}^{\text{tC,a,nC}}_{m,n,p}$) is a standardized or reduced subbundle of $\widetilde{\Sigma}^{\text{tC}}_{m,n,p}$ (resp. $\widetilde{\Sigma}^{\text{tC,a}}_{m,n,p}$); and $\widetilde{\mathcal{O}\Sigma}^{\text{tC,co,Cbl},k}_{m,n,p}$ (resp. $\widetilde{\mathcal{O}\Sigma}^{\text{tC,co,a,Cbl}}_{m,n,p}$) is a standardized or reduced subbundle of $\widetilde{\Sigma}^{\text{tC,co}}_{m,n,p}$ (resp. $\widetilde{\Sigma}^{\text{tC,co,a}}_{m,n,p}$).

We remind the reader of the adopted convention that, e.g., by referring to the (full) bundle $\widetilde{\Sigma}^{\text{co}}_{m,n,p}$ we mean the $GL(n)$-principal bundle $(\widetilde{\Sigma}^{\text{co}}_{m,n,p}, \Sigma^{\text{co}}_{m,n,p}, GL(n))$, i.e., the structure group is $GL(n)$ and the base space is $\Sigma^{\text{co}}_{m,n,p}$. Recall that an $\mathcal{O}$ would indicate an $O(n)$-subbundle.

We also add that in a more general context, equivariant assignment of a positive definite matrix to a realization R has appeared in the literature in relation to the algebraic Riccati equations or Lyapunov equations (see e.g., [49], [56]). In this context, balancing arises when there are two such (dual) equations given, which in our terminology gives rise to two reduction maps that can be used to define a balancing reduction map as prescribed by Proposition 15.

C. Another Example: Euclidean Norm Balancing

An interesting example of the variational formulation of balancing in the literature is called the Euclidean norm balancing [46] or in the language of [55] clustering, where a minimal realization $R = (A, B, C) \in \widetilde{\Sigma}^{\min}_{m,n,p}$ is transformed to a norm-balanced one by minimizing the function $h_{\text{nbl}} : GL(n) \to \mathbb{R}$

$$h_{\text{nbl}}(P; R) = \|P^{-1} A P\|_F^2 + \|P^{-1} B\|_F^2 + \|CP\|_F^2. \quad (19)$$

This function is also constant on $O(n)$. Also note that we are not assuming a.s. realizations. We show that this definition results in a reduced realization subbundle. Let $S = (PP^\top)^{-1} \in \mathcal{S}(n)$, then it can be shown that the first order optimality condition for h_{nbl} is equivalent to the equation

$$SAS^{-1} A^\top S + SBB^\top S = A^\top SA + C^\top C, \quad (20)$$

which has a unique solution $S \in \mathcal{S}(n)$, and this solution characterizes every global minimizer of h_{nbl} up to a right orthogonal factor (see [46] or [54, p. 221 and pp. 261-3]). Using this uniqueness result, it is easy to verify that the function $R \mapsto \nu_{\text{nbl}}(R) = (PP^\top)^{-1}$ is an equivariant function from $\widetilde{\Sigma}^{\min}_{m,n,p}$ to $\mathcal{S}(n)$. We just need to show that ν_{nbl} is a smooth function, hence a reduction map. To show this we use the derivations in [54, pp. 260-1], which have a slightly different context. Consider the smooth function $F : \widetilde{\Sigma}^{\min}_{m,n,p} \times \mathcal{S}(n) \to \mathbb{R}^{n \times n}$ defined as $F(R, S) = SAS^{-1} A^\top S + SBB^\top S - A^\top SA - C^\top C$. We need to show that the solution to $F(R, S) = 0$ depends smoothly on R. Resorting to the implicit function theorem, we need to show that the derivative of $S \mapsto F(R, S)$ is a full rank matrix at any solution S and its vicinity. It is easy to see that the derivative F_* at S (a linear map from the tangent space $T_S\mathcal{S}(n)$ to itself $\mathbb{R}^{n\times n}$) can be written as

$$F_* = SAS^{-1}A^\top \otimes I_n - SAS^{-1} \otimes SAS^{-1} + \\ I_n \otimes SAS^{-1}A^\top + SBB^\top \otimes I_n + I_n \otimes SBB^\top - A^\top \otimes A^\top, \quad (21)$$

where $\otimes$ is the Kronecker product. We can re-write F_* as

$$F_* = (\sqrt{S}\otimes\sqrt{S})\big(\tilde{A}\tilde{A}^\top \otimes I_n - \tilde{A}\otimes\tilde{A} + I_n\otimes\tilde{A}\tilde{A}^\top \quad (22) \\ + \tilde{B}\tilde{B}^\top \otimes I_n + I_n \otimes \tilde{B}\tilde{B}^\top - \tilde{A}^\top\otimes\tilde{A}^\top\big)(\sqrt{S^{-1}}\otimes\sqrt{S^{-1}}),$$

where $\tilde{R} = (\tilde{A},\tilde{B},\tilde{C}) = \sqrt{S^{-1}}\circ R$. Notice that (20) can be re-written as $\tilde{A}\tilde{A}^\top + \tilde{B}\tilde{B}^\top = \tilde{A}^\top\tilde{A} + \tilde{C}^\top\tilde{C}$. Thus the middle term in (22) can be written as

$$F_* = (\sqrt{S}\otimes\sqrt{S})\big(\tilde{A}^\top\tilde{A} \otimes I_n - \tilde{A}\otimes\tilde{A} + I_n\otimes\tilde{A}\tilde{A}^\top \quad (23) \\ + \tilde{C}^\top\tilde{C} \otimes I_n + I_n \otimes \tilde{B}\tilde{B}^\top - \tilde{A}^\top\otimes\tilde{A}^\top\big)(\sqrt{S^{-1}}\otimes\sqrt{S^{-1}}).$$

It is shown in [54, p. 260, Lemma 5.1] that the middle expression in the above is a matrix all whose eigenvalues are of positive real parts as long as $\tilde{R}$ is controllable or observable. This suffices to show that F_* is full rank in a neighborhood around a solution S. Finally, the desired smoothness of ν_{nbl} follows from the implicit function theorem. Thus we have:

Theorem 17: The subbundle of Euclidean norm balanced realizations defined as

$$\widetilde{\mathcal{O}\Sigma}^{\text{min,nbl}}_{m,n,p} = \{R \in \widetilde{\Sigma}^{\text{min}}_{m,n,p} | AA^\top + BB^\top = A^\top A + C^\top C\} \quad (24)$$

is a standardized or reduced subbundle of $\widetilde{\Sigma}^{\text{min}}_{m,n,p}$.

Notice that $\widetilde{\mathcal{O}\Sigma}^{\text{min,nbl}}_{m,n,p} = \nu^{-1}_{\text{nbl}}(I_n)$, and for any realization $R \in \widetilde{\mathcal{O}\Sigma}^{\text{min,nbl}}_{m,n,p}$ we have $\|B\|_F = \|C\|_F$, which justifies the name. We refer the reader to [54], [57] for algorithms to compute Euclidean norm balanced realizations.

V. The Alignemnt Distance

The system manifolds of interest are base spaces of $GL(n)$ realization bundles. The framework of group action induced distances described in §II-B requires a $GL(n)$-invariant distance on the respective realization bundle space to define a distance on the base system space. An important fact is that constructing $GL(n)$-invariant distances, due to noncompactness of $GL(n)$, is theoretically and numerically complicated. In contrast, $O(n)$-invariant (or unitarily-invariant) distances are abundant. Recall that among matrix norms many of them are $O(n)$-invariant, but none is $GL(n)$-invariant. Here, is where reducing the structure group of realization bundles becomes useful. The formal definition of the alignment distance is:

Definition 18 (The Alignment Distance): Let $(\widetilde{\Sigma}_{m,n,p}, \Sigma_{m,n,p})$ be a $GL(n)$ realization-system bundle. Let $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ be an $O(n)$-subbundle of $\widetilde{\Sigma}_{m,n,p}$ (e.g., any of the reduced or standardized subbundles of $\widetilde{\Sigma}_{m,n,p}$ in § IV). Also let $\tilde{d}_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}(\cdot,\cdot)$ be an $O(n)$-invariant distance on $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$. Given any two systems $M_1, M_2 \in \Sigma_{m,n,p}$ and their respective (standardized) realizations $R_1, R_2 \in \widetilde{\mathcal{O}\Sigma}_{m,n,p}$ define the group action induced distance on $\Sigma_{m,n,p}$ as

$$d_{\Sigma_{m,n,p},\widetilde{\mathcal{O}\Sigma}_{m,n,p}}(M_1,M_2) = \min_{Q\in O(n)} \tilde{d}_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}(Q\circ R_1, R_2). \quad (25)$$

We call $d_{\Sigma_{m,n,p},\widetilde{\mathcal{O}\Sigma}_{m,n,p}}(\cdot,\cdot)$ the alignment distance associated with, induced by, or subordinate to $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ and $\tilde{d}_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}(\cdot,\cdot)$. If the context is clear enough, then we simply may call $d_{\Sigma_{m,n,p},\widetilde{\mathcal{O}\Sigma}_{m,n,p}}(\cdot,\cdot)$ the alignment distance and may write it as $d_{\Sigma_{m,n,p}}(\cdot,\cdot)$ or simply d. The minimization problem is called the realization alignment problem or simply the alignment problem.

Figure 1 pictorially shows this definition. A $GL(n)$ realization bundle $\widetilde{\Sigma}_{m,n,p}$ and its reduced $O(n)$-subbundle $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ are depicted. Before further explanation, we should stress that the figure is not quite accurate, since the subbundle $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ has zero thickness compared with $\widetilde{\Sigma}_{m,n,p}$. Two systems $M_1, M_2 \in \Sigma_{m,n,p}$ and their total fibers of realizations are shown. The (standardized) realizations R_1 and R_2 of M_1 and M_2 in $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ are chosen and *aligned* according to (25) to find the alignment distance; the orthogonal matrix Q is *a* change of basis that aligns R_1 to R_2. Alignment is thought to be achieved when we have a *horizontal* line segment connecting $Q\circ R_1$ and R_2, the length of which is the alignment distance. The *sub-fibers* within $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ (e.g, $O(n)\circ R_1$ and $O(n)\circ R_2$) are shown as *parallel vertical* lines, implying that by moving any such horizontal line segment vertically, it traverses the entire respective sub-fibers. This also indicates that the alignment distance is independent of the choice of the realizations of M_1 and M_2 in $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$. The terms "vertical" and "horizontal" are borrowed from the Riemannian submersion case [58] (see also §V-B1), and used metaphorically, since in our setting we are not assuming any notion of angle. However, when using $\tilde{d}_F$ (see (29)), due to smoothness (of $\tilde{d}^2_F$), the line segment will be perpendicular to both the sub-fibers $O(n)\circ R_1$ and $O(n)\circ R_2$, where standard Euclidean inner product is used to define angles.

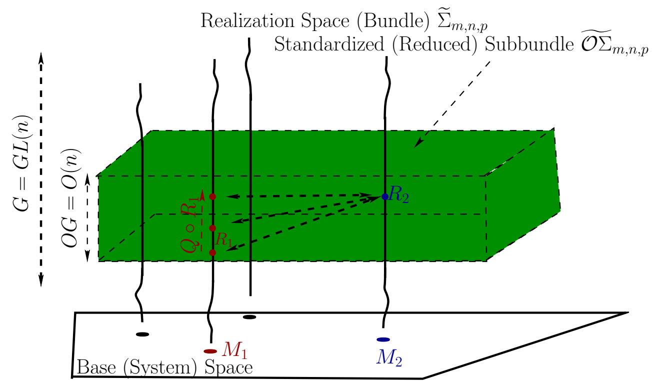

Fig. 1. Computing the alignment distance between to systems M_1 and M_2, see (25). The box area shows a reduced $O(n)$-subbundle, $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$. The full $GL(n)$-realization bundle $\widetilde{\Sigma}_{m,n,p}$ is noncompact (and unbounded). The realizations R_1 and R_2 are, respectively, any realizations of M_1 and M_2 in $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$, and Q is a best orthogonal change of basis that aligns R_1 to R_2.

Theorem 19 (Alignment Distance Topology): The topology induced by the alignment distance (25) on $\Sigma_{m,n,p}$ coincides with its natural quotient topology (independent of the choice of the reduced bundle $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ and the distance $\tilde{d}_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}(\cdot,\cdot)$).

Proof: Because of bundle reduction property we have $\Sigma_{m,n,p} = \widetilde{\Sigma}_{m,n,p}/GL(n) \overset{\text{diff}}{=} \widetilde{\mathcal{O}\Sigma}_{m,n,p}/O(n)$. Since $O(n)$ is compact it follows from part 3 of Theorem 3 that the alignment distance (25) induces the same topology on $\Sigma_{m,n,p}$ as its natural quotient topology (w.r.t. $GL(n)$). Obviously this is independent of the specific reduced subbundle $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ and the distance $\tilde{d}_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}$ as long it induces the manifold topology of $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$. ■

Remark 20 (Applicability of Extrinsic Distances): We remark that since $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ is an embedded submanifold of $\widetilde{\Sigma}_{m,n,p}$ and hence of $\widetilde{\mathcal{L}}_{m,n,p}$, the distance $\tilde{d}_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}$ can be any $O(n)$-invariant distance on the Euclidean space $\widetilde{\mathcal{L}}_{m,n,p}$ which induces the natural Euclidean topology or any distance on $\widetilde{\Sigma}_{m,n,p}$ which induces its manifold topology.

Remark 21 (Noncompact Alignment Distance): In principle, any group action induced distance on the full $GL(n)$-bundle $(\widetilde{\Sigma}_{m,n,p}, \Sigma_{m,n,p})$ can also be called an alignment distance. We call such a distance a *noncompact* alignment distance; and since in practice such distances are difficult to construct and calculate, we reserve the term "alignment" for distances as in Definition 18.

Remark 22 (On Zero Alignment Distance): On the manifold $\Sigma^{\text{min}}_{m,n,p}$, if the alignment distance between two (minimal) systems is zero, then the two systems are indistinguishable in the input-output sense, and their distance in any input-output based distance on $\Sigma^{\text{min}}_{m,n,p}$ also will be zero. Conversely, if two systems are indistinguishable in the input-output sense, then their alignment distance will be zero. However, the alignment distance can also be defined on system spaces such as $\Sigma^{\text{ob}}_{m,n,p}$, on which input-output based distances are meaningless, as such distances can only compare the minimal parts of two observable systems. Here, it is insightful to note that two observable systems that have indistinguishable input-output behaviors may not have zero alignment distance.

Remark 23 (On Naturalness of the Alignment Distance): The natural way of thinking about the alignment distance is to *start* with state-space systems and, at least temporarily, forget about the input-output description. Then, the set of systems of order n and size (m,p), $\mathcal{L}_{m,n,p}$, has a natural quotient topology inherited from $\widetilde{\mathcal{L}}_{m,n,p}$. The alignment distance is natural in the *topological* sense, because it induces this natural quotient topology on the subspaces of $\mathcal{L}_{m,n,p}$ which are manifolds. In the case of minimal systems $\Sigma^{\text{min}}_{m,n,p}$, one can establish a homeomorphism between $\Sigma^{\text{min}}_{m,n,p}$ and $\mathcal{H}_{m,n,p}$, the space of strictly proper transfer functions of McMillan degree n and dimension $p \times m$; thus, it happens that, in this case, input-output based distances defined on $\mathcal{H}_{m,n,p}$ are also topologically natural, as they induce the same topology. However, such a homeomorphism cannot be defined between spaces such as $\Sigma^{\text{ob}}_{m,n,p}$ and spaces of transfer functions; hence, input-output distances cannot be defined on $\Sigma^{\text{ob}}_{m,n,p}$ and other non-minimal system spaces. In our view—as a somewhat subjective statement—the alignment distance is natural in the *methodological* sense too, i.e., once one has a quotient space under a group action and knows Theorem 3, then, the most immediate plan to define a distance on it is to try find a group-invariant distance on the top space and descend it to the bottom space.

The alignment distance (but not the topology it induces) depends on both the chosen reduced subbundle $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ and the $O(n)$-invariant distance $\tilde{d}_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}(\cdot,\cdot)$. In that sense, the alignment distance is a general family of distances. Ultimately, the choice of the reduced subbundle and the distance $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ depends on the application. As for the choice of the $O(n)$-invariant distance on $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$, one can have either an *intrinsic* distance or an *extrinsic* one. An intrinsic distance is defined based on the length of curves defined on $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$, whereas an extrinsic distance is based on the curves in a space ambient to $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ e.g., the Euclidean space $\widetilde{\mathcal{L}}_{m,n,p}$. An example of an intrinsic distance is a Riemannian distance which can be complicated to calculate numerically (see § V-B1). An extrinsic distance, however, is much simpler to work with, because often solving the alignment problem reduces to a static optimization problem on the compact set $O(n)$.

From the computational point of view, perhaps the simplest distance on a realization space is the Frobenius norm based (extrinsic) distance $\tilde{d}_F(\cdot,\cdot)$ defined as

$$\tilde{d}^2_F(R_1,R_2) = \|A_1-A_2\|^2_F+\|B_1-B_2\|^2_F+\|C_1-C_2\|^2_F. \quad (26)$$

This is the natural Euclidean distance on $\widetilde{\mathcal{L}}_{m,n,p}$ and all its embedded submanifolds including all the reduced subbundles in § IV. One could have a weighted version of this distance too. In [5] a fast algorithm for solving the alignment problem in this distance is given. Obviously, many other distances are possible, which could prove useful in applications. For example, the following distance can be used on $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$:

$$\tilde{d}^2_{FL}(R_1,R_2) = \sum_{l=1}^{L} \|A_1^l - A_2^l\|^2_F + \|B_1-B_2\|^2_F + \|C_1-C_2\|^2_F \quad (27)$$

to encode more *behavioral* information. To be precise, $\tilde{d}_{FL}$ is not a distance on $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ as a subset of $\widetilde{\mathcal{L}}_{m,n,p}$, rather through the diffeomorphism $(A,B,C) \mapsto (A,\ldots,A^L,B,C)$. For a.s. realizations a variation of this distance is:

$$\tilde{d}^2_{FI}(R_1,R_2) = \|(I_n-A_1)^{-1} - (I_n-A_2)^{-1}\|^2_F + \|B_1-B_2\|^2_F + \|C_1-C_2\|^2_F \quad (28)$$

A. Example: closed-form solution for first order systems

The alignment problem is a non-convex problem and in general no closed-form solution for it is expected. However, for first order systems, it can be found readily, which gives some insight. Consider two systems $M_1, M_2 \in \Sigma^{\text{min,a}}_{m,1,p}$ with realizations $R_i = (a_i, b_i, c_i) \in \widetilde{\Sigma}^{\text{min,a}}_{m,1,p}$. We find the alignment distance associated with the subbundle of balanced realizations $\widetilde{\mathcal{O}\Sigma}^{\text{min,a,bl}}_{m,1,p}$. Now $R_i^{\text{bl}} = (a_i, \sqrt{\|c_i\|\|b_i\|^{-1}}b_i, \sqrt{\|b_i\|\|c_i\|^{-1}}c_i)$ is a balanced realization of M_i. The optimization is on a binary variable $Q \in \{\pm 1\} = O(1)$, and we get:

$$d^2_F(M_1,M_2) = (a_1-a_2)^2 + 2\|b_1\|\|c_1\| + 2\|b_2\|\|c_2\| - 2\big|\cos\measuredangle(b_1,b_2) + \cos\measuredangle(c_1,c_2)\big| \sqrt{\|b_1\|\|c_1\|\|b_2\|\|c_2\|}, \quad (29)$$

where $\cos\measuredangle(x_1,x_2)$ is the cosine of the angle between two vectors x_1 and x_2. With respect to $\Sigma^{\text{min}}_{m,1,p}$, which includes both stable and unstable systems, it is easy to see that the same expression holds for the alignment distance on $\Sigma^{\text{min}}_{m,1,p}$ associated with $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl,1}}_{m,1,p}$, the subbundle of 1-balanced realizations, as well as with $\widetilde{\mathcal{O}\Sigma}^{\text{min,nbl}}_{m,1,p}$, the subbundle of Euclidean norm balanced realizations (see § IV-C).

B. Discussions and Relation to the Existing Literature

Here, we discuss some interesting relations between the alignment distance and the existing control literature.

1) Relation to Riemannian Distances on $\Sigma^{\min,a}_{m,n,p}$

There are at least two classes of Riemannian distances on the manifold of minimal systems $\Sigma^{\min}_{m,n,p}$. The first one, due to Krishnaprasad and Martin [13], [17], is an internal distance (similar to the alignment distance), and the second one, due to Hanzon and Marcus [18], [59], is an external distance (based on input-output description of systems). Since the latter is not specifically related to our approach, we focus on the former. The starting point is exactly the $GL(n)$ realization-system principal fiber bundle $(\widetilde{\Sigma}^{\min}_{m,n,p}, \Sigma^{\min}_{m,n,p})$, and then a $GL(n)$-invariant Riemannian *metric* is defined on the realization bundle $\widetilde{\Sigma}^{\min}_{m,n,p}$, which in turn induces a Riemannian metric on $\Sigma^{\min}_{m,n,p}$. This is an example of a construction called *Riemannian submersion*, which is a general way of defining a Riemannian metric on a quotient space [58]. The specific metric defined in [17] is

$$\tilde{g}^{\mathrm{KM}}_R(\mathrm{d}R, \mathrm{d}R) = \mathrm{tr}(W_{\mathrm{o},n}\mathrm{d}AW_{\mathrm{c},n}\mathrm{d}A^\top) + \mathrm{tr}(\mathrm{d}B^\top W_{\mathrm{o},n}\mathrm{d}B) + \mathrm{tr}(\mathrm{d}CW_{\mathrm{c},n}\mathrm{d}C^\top), \quad (30)$$

where $\mathrm{d}R = (\mathrm{d}A, \mathrm{d}B, \mathrm{d}C)$ is a *horizontal* tangent vector in the tangent space $T_R\widetilde{\Sigma}^{\min}_{m,n,p}$. Here, $T_R\widetilde{\Sigma}^{\min}_{m,n,p}$ is decomposed into a *vertical* subspace (along the fiber at R) and its *complement* called the *horizontal* subspace. A general tangent vector $\mathrm{d}R$ can be written uniquely as $\mathrm{d}R = \mathrm{d}R^h + \mathrm{d}R^v$ in terms of its vertical and horizontal components. We define $\tilde{g}^{\mathrm{KM}}_R(\mathrm{d}R, \mathrm{d}R) = \tilde{g}^{\mathrm{KM}}_R(\mathrm{d}R^h, \mathrm{d}R^h) + \tilde{g}^{\mathrm{KM}}_R(\mathrm{d}R^v, \mathrm{d}R^v)$, where $\tilde{g}^{\mathrm{KM}}_R(\mathrm{d}R^h, \mathrm{d}R^h)$ is as in (30), and the exact form of $\tilde{g}^{\mathrm{KM}}_R(\mathrm{d}R^v, \mathrm{d}R^v)$ is immaterial for our purposes. The crucial point is that $\tilde{g}^{\mathrm{KM}}_R(\mathrm{d}R, \mathrm{d}R)$ is a positive definite quadratic form at each horizontal tangent space and is invariant along every fiber, i.e., $\tilde{g}^{\mathrm{KM}}_{P\circ R}(P \circ \mathrm{d}R, P \circ \mathrm{d}R) = \tilde{g}^{\mathrm{KM}}_R(\mathrm{d}R, \mathrm{d}R)$, for $\forall P \in GL(n)$ and any horizontal tangent vector $\mathrm{d}R$. Thus the Riemannian metric $\tilde{g}^{\mathrm{KM}}_R$ on the top (realization) space $\widetilde{\Sigma}^{\min}_{m,n,p}$ induces a Riemannian metric $g^{\mathrm{KM}}_{[R]}$ on the base (system) space $\Sigma^{\min}_{m,n,p}$. By this construction, we have a $GL(n)$-invariant Riemannian distance on $\widetilde{\Sigma}^{\min}_{m,n,p}$ denoted by $\tilde{d}^{\mathrm{KM}}$ and the corresponding Riemannian distance on $\Sigma^{\min}_{m,n,p}$ denoted by d^{KM}. It is easy to see that, indeed, d^{KM} is a group action induced distance associated with $\tilde{d}^{\mathrm{KM}}$, in the sense of Theorem 3, i.e., an example of the noncompact alignment distance (see Remark 21). Numerical computation of d^{KM} is, in general, difficult. In [17] it was merely used to give an alternative proof of the principal fiber bundle structure of $(\widetilde{\Sigma}^{\min}_{m,n,p}, \Sigma^{\min}_{m,n,p})$.

Next, we show that although the alignment distance associated with the Frobenius distance (26) is an extrinsic distance, it can be related to a Riemannian metric. First, note that the approach of Krishnaprasad-Martin in defining $GL(n)$-invariant metrics on the realization bundle can be extended in various ways, e.g., the following Riemannian metric is a simple variation that penalizes closeness to non-minimality by using the inverses of the Gramians :

$$\tilde{g}^{\mathrm{p}}_R(\mathrm{d}R, \mathrm{d}R) = \mathrm{tr}(W^{-1}_{\mathrm{c},n}\mathrm{d}A\, W^{-1}_{\mathrm{o},n}\, \mathrm{d}A^\top) + \mathrm{tr}(\mathrm{d}B^\top W^{-1}_{\mathrm{c},n}\mathrm{d}B) + \mathrm{tr}(\mathrm{d}C\, W^{-1}_{\mathrm{o},n}\mathrm{d}C^\top), \quad (31)$$

where $\mathrm{d}R$ is a hoirzontal tangent vector. A more interesting variant arises from using a bundle reduction map $\nu(R)$ (see Definition 14). Let us define a Riemannian metric on the (generic) realization bundle $\widetilde{\Sigma}_{m,n,p}$ as

$$\tilde{g}^{\nu}_R(\mathrm{d}R,\mathrm{d}R) = \mathrm{tr}\big(\nu(R)\,\mathrm{d}A\;\nu(R)^{-1}\mathrm{d}A^\top\big) + \mathrm{tr}(\mathrm{d}B^\top\nu(R)\,\mathrm{d}B) + \mathrm{tr}(\mathrm{d}C\;\nu(R)^{-1}\,\mathrm{d}C^\top), \quad (32)$$

where $R \in \widetilde{\Sigma}_{m,n,p}$. Consider the restriction of this metric to the submanifold $\widetilde{\mathcal{O}\Sigma}_{m,n,p} = \nu^{-1}(I_n)$, i.e., where $\nu(R) = I_n$:

$$\tilde{g}^{\mathrm{F}}_R(\mathrm{d}R, \mathrm{d}R) = \mathrm{tr}\big(\mathrm{d}A\,\mathrm{d}A^\top\big) + \mathrm{tr}(\mathrm{d}B^\top\mathrm{d}B) + \mathrm{tr}(\mathrm{d}C\,\mathrm{d}C^\top). \quad (33)$$

The key point is that now $R \in \widetilde{\mathcal{O}\Sigma}_{m,n,p}$. This Riemannian metric is only $O(n)$-invariant, which matches that fact that $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ is an $O(n)$-bundle. But more importantly, $\tilde{g}^{\mathrm{F}}_R$ is the infinitesimal version of the Frobenius distance $\tilde{d}_F$ (26), i.e., given $R \in \widetilde{\mathcal{O}\Sigma}_{m,n,p}$ and $\mathrm{d}R$ in the tangent space to $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ at R, $\tilde{g}^{\mathrm{F}}_R(\mathrm{d}R, \mathrm{d}R)$ in (33) is equal to $\lim_{\epsilon\downarrow 0} \tilde{d}^2_F(R, R + \epsilon\,\mathrm{d}R)$. Therefore, the interesting finding is that the alignment distance associated with the Frobenius distance $\tilde{d}_F$ and $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ approximates the Riemannian distance on $\Sigma_{m,n,p}$ associated with the Riemannian metric (33). For large distances they deviate, but for computing the Riemannian distance one has to solve the complicated geodesic equation, whereas for the alignment distance one only needs to solve the alignment problem (25). Interestingly, it is easy to see from calculations in §V-A that, on each connected component of $\Sigma^{\min,a}_{1,1,1}$ (but not between different components) the alignment distance subordinate to $\widetilde{\mathcal{O}\Sigma}^{\min,a,bl}_{1,1,1}$ and $\tilde{d}_F$ coincides with the Riemannian distance associated with (33) on $\widetilde{\mathcal{O}\Sigma}^{\min,a,bl}_{1,1,1}$ (see [11] on the components of $\Sigma^{\min,a}_{1,n,1}$).

2) On Comparing Realizations

At the core of the alignment distance is a distance between realizations. Although group action induced distances between systems have not appeared in literature, distances between realizations are not unprecedented. The scenario in which such distances have been used is more akin to parameterization of a system or realization than modeling the *time behavior* (as one does in the input-output formulation). The most notable example is the so-called *Eising distance* [60], [61], [54, p. 259]. In fact, the Eising distance is nothing but the Frobenius distance $\tilde{d}_F$ in (26). The distance to uncontrollability can be formulated as: $\inf_{\bar{R}} \tilde{d}_F(R, \bar{R})$, where $\bar{R} \in \overline{\widetilde{\Sigma}}^{\mathrm{ob}}_{m,n,p} \triangleq \widetilde{\mathcal{L}}_{m,n,p} \setminus \widetilde{\Sigma}^{\mathrm{ob}}_{m,n,p}$. i.e., find the closest uncontrollable realization $\bar{R}$ to a given controllable realization R. A 2-norm version of this problem also has been studied [60]. Although the distance to uncontrollability is a distance between *realizations*, in the literature it is unduly referred to as distance to uncontrollable systems [60]–[62]. A system-level version of this problem will not make much sense, since given $R \in \widetilde{\Sigma}^{\mathrm{co}}_{m,n,p}$ one can make $P \circ R$ as close to uncontrollability as one wishes by varying $P \in GL(n)$. In a more formal language, the closure of the $GL(n)$-orbit of R in $\widetilde{\mathcal{L}}_{m,n,p}$ contains uncontrollable realizations. In fact, this is really the indication that $\widetilde{\mathcal{L}}_{m,n,p}/GL(n)$ is a non-Hausdorff space. The reader is referred to [63] for determining the closure of the $GL(n)$-orbit of $R \in \widetilde{\Sigma}^{\mathrm{co}}_{m,n,p}$. Interestingly, if one tries to formulate the problem of distance to non-minimality at the system-level using the alignment distance, then one will arrive at the problem of model order reduction in the alignment distance. In a recent paper we pursued this approach [23].

As another example of the use of the distance $\tilde{d}_F$ between realizations, we already mentioned the variational formulation of the problem of balancing [46], [48], [54], [55]. Another example is in the so-called grey-box system identification [64].

VI. Extensions and Generalizations

The basic notion of the alignment distance can be extended in various ways. We provide examples regarding including the initial state in the distance and the alignment distance for stochastic systems (see [22] for some other examples).

A. Distances Accounting for the Initial State

Define the *realization-state* space as $\widetilde{\mathcal{LS}}_{m,n,p} = \widetilde{\mathcal{L}}_{m,n,p} \times \mathbb{R}^n$. Then $GL(n)$ acts on $\widetilde{\mathcal{LS}}_{m,n,p}$ as

$$P \circledcirc (R, x) = (P \circ R, P^{-1}x). \tag{34}$$

We think of the quotient $\mathcal{LS}_{m,n,p} = \widetilde{\mathcal{LS}}_{m,n,p}/GL(n)$ as a *system-state* space (or *system-with-state* as called in [18]). The meaning is that, in (4), for any realization $R \in \widetilde{\mathcal{L}}_{m,n,p}$ and the corresponding (unique) initial state $x_0 = x \in \mathbb{R}^n$ all realization-state pairs $P \circledcirc (R, x)$ ($P \in GL(n)$) are indistinguishable from an input-output point of view for $t \geq 0$. We denote a system-state as $\mathbf{M} = (M, \mathbf{x}) \in \mathcal{LS}_{m,n,p}$. By choosing realization-state submanifolds of the form $(\widetilde{\Sigma}_{m,n,p}, \mathbb{R}^n)$, where $\widetilde{\Sigma}_{m,n,p}$ is any of the realization bundles in § III and Theorem 10, all of our results and constructions from § IV and § V can be extended easily to yield the alignment distance on the system-state spaces. For example, take two system-space pairs $\mathbf{M}_i = (M_i, \mathbf{x}_i)$ ($i = 1, 2$), where $M_i \in \Sigma^{\text{min}}_{m,n,p}$. Then choose a corresponding realization-state pair (R_i, x_i) and convert the pair to a k-balanced realization-state pair $(R_i^{\text{bl},k}, x_i^{\text{bl},k}) = P \circledcirc (R_i, x_i)$, where P is a state-space change of basis such that $R^{\text{bl},k} = P \circ R_i \in \widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$. An alignment distance between $\mathbf{M}_1$ and $\mathbf{M}_2$ associated with the Frobenius distance and $\widetilde{\mathcal{O}\Sigma}^{\text{min,bl},k}_{m,n,p}$ can be defined as

$$d_F^2(\mathbf{M}_1, \mathbf{M}_2) = \min_{Q \in O(n)} \tilde{d}_F^2(Q \circ R_1^{\text{bl},k}, R_2^{\text{bl},k}) + \|Q^\top x_1^{\text{bl},k} - x_2^{\text{bl},k}\|^2. \tag{35}$$

B. Stochastic Systems

Let $\{u_t\}_{-\infty}^{+\infty}$ in (4) be an m-dimensional, stationary, Gaussian, *white* process with zero mean and *identity* covariance. Consider the space of stochastic realizations

$$\widetilde{\mathcal{SL}}_{m,n,p} = \{(A, B, C) \in \widetilde{\mathcal{L}}^{\text{mp,a}}_{m,n,p} | \text{rank}(B) = m\}, \tag{36}$$

where $\widetilde{\mathcal{L}}^{\text{mp,a}}_{m,n,p}$ is the submanifold of a.s. stable and minimum-phase realizations in $\widetilde{\mathcal{L}}_{m,n,p}$ (see § III-A). The main object of interest is not the output y_t itself rather its covariance sequence or power spectral density. In view of this, two types of symmetries appear: the internal symmetry as in the deterministic case and the symmetry at input (via group $O(m)$). Specifically, $GL(n) \times O(m)$ acts on $\widetilde{\mathcal{SL}}_{m,n,p}$ as

$$(P, \Theta) \bullet (A, B, C) = (P^{-1}AP, P^{-1}B\Theta, CP), \tag{37}$$

i.e., realizations R and $(P, \Theta) \bullet R$ generate the same power spectral density. Conversely, due to the minimum-phase assumption, if the power spectral density of the output of two realizations $R_1, R_2 \in \widetilde{\mathcal{SL}}_{m,n,p}$ are equal, then their transfer functions are equal up to an $O(m)$ right factor [42], [65, p. 201] and if the realizations are minimal then $R_1 = (P, \Theta) \bullet R_2$ for some P and Θ. The space of stochastic systems is defined as $\mathcal{SL}_{m,n,p} = \widetilde{\mathcal{SL}}_{m,n,p}/(GL(n) \times O(m))$. The condition $\text{rank}(B) = m$ ensures that the action of $O(m)$ is free. Thus, as in § III, by passing to suitable submanifolds of $\widetilde{\mathcal{SL}}_{m,n,p}$ such as observable, controllable, minimal, or realizations with tall C, etc., one can get principal fiber bundles. We denote a generic stochastic realization-system bundle by $(\widetilde{\mathcal{S}\Sigma}_{m,n,p}, \mathcal{S}\Sigma_{m,n,p})$. For example, $(\widetilde{\mathcal{S}\Sigma}^{\text{min}}_{m,n,p}, \mathcal{S}\Sigma^{\text{min}}_{m,n,p})$ denotes the bundle of minimal stochastic realization-systems. All the (appropriate) bundle reduction schemes in § IV or specific stochastic balancing methods (e.g., as in [56]) can be applied to such a pair to yield a reduced bundle $(\widetilde{\mathcal{OS}\Sigma}_{m,n,p}, \mathcal{S}\Sigma_{m,n,p})$. Given an $O(n)$-invariant distance $\tilde{d}_{\widetilde{\mathcal{OS}\Sigma}_{m,n,p}}$ (e.g., $\tilde{d}_F$ in (26)), the alignment distance on $\mathcal{S}\Sigma_{m,n,p}$ is defined as

$$d^2_{\mathcal{S}\Sigma_{m,n,p}}(M_1, M_2) = \min_{(Q,\Theta)} \tilde{d}^2_{\widetilde{\mathcal{OS}\Sigma}_{m,n,p}}((Q, \Theta) \bullet R_1, R_2), \tag{38}$$

where R_i ($i = 1, 2$) is any realization of M_i in $\widetilde{\mathcal{OS}\Sigma}_{m,n,p}$ and $(Q, \Theta) \in O(n) \times O(m)$.

An alternative approach is where the group $O(m)$ is quotiented out and not explicit in defining the distance. We call $\widetilde{\mathcal{DS}\Sigma}_{m,n,p} = \widetilde{\mathcal{S}\Sigma}_{m,n,p}/(\{I_n\} \times O(m))$ a (deterministic) *pseudo* realization space for $\widetilde{\mathcal{S}\Sigma}_{m,n,p}$. One can verify that

$$\widetilde{\mathcal{DS}\Sigma}_{m,n,p} \overset{\text{diff}}{=} \{(A, BB^\top, C) | (A, B, C) \in \widetilde{\mathcal{S}\Sigma}_{m,n,p}\}. \tag{39}$$

We call $R^D = (A, BB^\top, C)$ a *deterministic pseudo* realization of $R = (A, B, C)$. We denote the action of $GL(n)$ on $\widetilde{\mathcal{DSL}}_{m,n,p}$ by $\star$ and note that

$$P \star (A, BB^\top, C) = (P^{-1}AP, P^{-1}BB^\top P^{-\top}, CP). \tag{40}$$

One can verify that $\mathcal{S}\Sigma_{m,n,p} \overset{\text{diff}}{=} \widetilde{\mathcal{DS}\Sigma}_{m,n,p}/GL(n)$. A reduced deterministic pseudo realization subbundle $\widetilde{\mathcal{ODS}\Sigma}_{m,n,p}$ can be constructed, in an obvious way, from the corresponding stochastic reduced realization subbundle $\widetilde{\mathcal{OS}\Sigma}_{m,n,p}$. Having an $O(n)$-invariant distance on $\widetilde{\mathcal{ODS}\Sigma}_{m,n,p}$, we define the alignment distance on $\mathcal{S}\Sigma_{m,n,p}$ as

$$d^2_{\mathcal{S}\Sigma_{m,n,p}}(M_1, M_2) = \min_{Q \in O(n)} \tilde{d}^2_{\widetilde{\mathcal{ODS}\Sigma}_{m,n,p}}(Q \star R_1^D, R_2^D), \tag{41}$$

where R_i^D ($i = 1, 2$) is any pseudo realization of M_i in $\widetilde{\mathcal{ODS}\Sigma}_{m,n,p}$. Computing this distance requires only a *single* minimization (cf. (38)). A simple example for $\tilde{d}_{\widetilde{\mathcal{OS}\Sigma}_{m,n,p}}$ is $\tilde{d}_F$.

Interestingly, the impulse response of the deterministic system with realization $R^{\text{cov}} = (A, W_c C^\top, C)$ for $t \geq 0$ is equal to the output covariance sequence of the minimal stochastic system (4) [59]. There is also a one-to-one correspondence between pseudo realizations $(A, BB^\top, C)$ and (A, W_c, C) (via the Lyapunov equation $W_c = BB^\top + AW_cA^\top$); thus the set

$$\widetilde{\mathcal{WS}\Sigma}^{\text{min}}_{m,n,p} = \{R^{\text{wc}} = (A, W_c, C) | (A, B, C) \in \widetilde{\mathcal{SL}}^{\text{min}}_{m,n,p}\} \tag{42}$$

is diffeomorphic to $\widetilde{\mathcal{DS}\Sigma}^{\text{min}}_{m,n,p}$, where the action of $GL(n)$ on this set is the same as $\star$. Here, the reduced subbundle can be derived from the corresponding stochastic reduced realization subbundle $\widetilde{\mathcal{OS}\Sigma}_{m,n,p} \subset \widetilde{\mathcal{SL}}^{\min}_{m,n,p}$. Now, simply by replacing R^D with R^{cov} or with $R^{\text{wc}} = (A, W_{\text{c}}, C)$ in (41) we can construct another class of distances on $\mathcal{S}\Sigma^{\min}_{m,n,p}$.

Finally, we remark that the case of stochastic systems is an example of a more general situation where we have *internal*, *input*, and *output* symmetries on the space of realizations; namely a group $G = G_{\text{out}} \times G_{\text{int}} \times G_{\text{in}}$ acts on a realization space $\widetilde{\Sigma}_{m,n,p}$, where the groups G_{int}, G_{in}, and G_{out} act at the input, internally, and at the output, respectively. The basic idea of realization alignment can be applied here, too. In particular, if G is a noncompact Lie group and an appropriate realization-system bundle exists, then by passing to a useful reduced subbundle one can only deal with a compact symmetry group and define an alignment distance easily. Note that in a more general setting, any of the mentioned groups can be discrete, e.g., G_{out} can be a permutation group, in which case our approach applies again; however, if the discrete group is noncompact, then the situation is more complicated.

VII. Conclusion

The differential-geometric foundation of the alignment distance was described. The alignment distances is, indeed, a family of distances, which crucially depend on the chosen reduced subbundles (e.g., balanced realizations vs. input-normalized) and the distances thereon. The usefulness of the alignment distance in specific control applications could depend on the choice of these "parameters," and would require further research. We already alluded to some potential applications in § I-C. Search for more application-specific reduced subbundles and distances could be a possible research direction, together with computational algorithms. The Frobenius-norm based alignment distance—the only one we have used so far—is an extrinsic distance, but an interesting question is to what extend geometrical notions like "geodesic" can be defined and used. Whether the alignment distance can be related (explicitly or implicitly) to input-output distances is an interesting question. Extension of the alignment distance to other classes of system such as LPV systems is another immediate research direction.

Appendix
Proof of Theorem 3

Proof: The fact that $\inf_{P\in G} \tilde{d}_{\tilde{\Sigma}}(P\circ R_1, R_2)$ depends only on M_1 and M_2 is a direct consequence of G-invariance of $\tilde{d}$. That $d_\Sigma(M_1, M_2) \geq 0$ is obvious by the definition. That d_Σ is symmetric is because $\tilde{d}_{\tilde{\Sigma}}$ is symmetric and G-invariant. To see the triangle inequality for d_Σ notice that

$$\begin{aligned}&\inf_{P_1} \tilde{d}_{\tilde{\Sigma}}(P_1 \circ R_1, R_2) \leq \tilde{d}_{\tilde{\Sigma}}(P_1 \circ R_1, P_3 \circ R_3) + \qquad (43)\\ &\tilde{d}_{\tilde{\Sigma}}(P_3 \circ R_3, R_2) = \tilde{d}_{\tilde{\Sigma}}((P_1 P_3^{-1}) \circ R_1, R_3) + \tilde{d}_{\tilde{\Sigma}}(P_3 \circ R_3, R_2)\\ &\Rightarrow \inf_{P_1} \tilde{d}_{\tilde{\Sigma}}(P_1 \circ R_1, R_2) \leq \inf_{P} \tilde{d}_{\tilde{\Sigma}}(P \circ R_1, R_3) + \inf_{P_3} \tilde{d}_{\tilde{\Sigma}}(P_3 \circ R_3, R_2).\end{aligned}$$

In the above, $P = P_1 P_3^{-1}$ ($\circ$ is a right action), but clearly the minimization over $P \in G$ imposes no constraint on $P \in G$.

Next, we show statement 2. To see that d_Σ is positive definite when the G-orbits are closed, let $\{P_i\}_i$ be such that $\lim_i \tilde{d}_{\tilde{\Sigma}}(P_i \circ R_1, R_2) = 0$. Then $\{P_i \circ R_1\}_i$ converges to R_2. But since the orbit of R_1 closed, this implies that R_2 must belong to the orbit, i.e., $R_2 = \bar{P} \circ R_1$ for some $\bar{P} \in G$. Thus, $d_\Sigma(M_1, M_2) = 0$ implies that $M_1 = M_2$ or, $d_\Sigma(M_1, M_2) > 0$ unless $M_1 = M_2$.

Now, we show that the metric and quotient topologies on Σ coincide. Let $\pi : \widetilde{\Sigma} \to \Sigma$ be the projection map. It suffices to show that π is a quotient map in the d_Σ-topology, since the quotient topology is the only topology in which this holds. Let $B_{\tilde{d}_{\tilde{\Sigma}}}(R, \epsilon)$ and $B_{d_\Sigma}(M, \epsilon)$, respectively, denote $\tilde{d}_{\tilde{\Sigma}}$ and d_Σ open metric balls of radius ϵ around R and $M = \pi(R)$ in $\widetilde{\Sigma}$ and Σ. We have $\pi(B_{\tilde{d}_{\tilde{\Sigma}}}(R_1, \epsilon)) = B_{d_\Sigma}(\pi(R_1), \epsilon)$ because

$$M \in B_{d_\Sigma}(M_1, \epsilon) \Leftrightarrow \exists R \in \pi^{-1}(M) \text{ s.t. } R \in B_{\tilde{d}_{\tilde{\Sigma}}}(R_1, \epsilon), \quad (44)$$

where $M_1 = \pi(R_1)$. Thus π is an open map in the d_Σ-topology, since it maps the basis elements of the $\tilde{d}_{\tilde{\Sigma}}$-topology to open sets in the d_Σ-topology. Moreover, π is continuous in the d_Σ-topology, because $\pi^{-1}(B_{d_\Sigma}(M, \epsilon)) = \cup_{R\in\pi^{-1}(M)} B_{\tilde{d}_{\tilde{\Sigma}}}(R, \epsilon)$ is open in the $\widetilde{\Sigma}$-topology. It follows that since π is surjective, it is a quotient map in the d_Σ-topology, as claimed.

Finally, to see statement 3, first, assume that every closed and bounded set in $(\widetilde{\Sigma}, \tilde{d}_{\tilde{\Sigma}})$ is compact. Let $\{P_i\}_i$ be a sequence such that $\lim_i \tilde{d}_{\tilde{\Sigma}}(P_i \circ R_1, R_2) = \inf_P \tilde{d}_{\tilde{\Sigma}}(P \circ R_1, R_2)$. For large i_o, the points $P_i \circ R_1$ with $i \geq i_0$ are in a bounded ball around R_2 and since the orbit $G \circ R_1$ is a closed set in $\widetilde{\Sigma}$ the points lie in a closed and bounded set, hence, compact. Thus, there is a converging subsequence of $\{P_i \circ R_1\}_i$ and since the orbit $G \circ R_1$ is closed, the limit is of the form $P \circ R_1$, which means that the infimum is achieved. The same argument applies if G is compact, since (due to continuity of the action) any G-orbit is also closed in $\widetilde{\Sigma}$. ■

Acknowledgment

The authors thank Laurent Younes for helpful comments and discussions. B. Afsari thanks Prof. P. S. Krishnaprasad for some early references, and Arash Komaee for helpful comments. The authors are also grateful to the anonymous reviewers for diligent reviews and constructive comments that helped to improve the quality of this paper significantly.

References

[1] R. E. Kalman, "Mathematical description of linear dynamical systems," *Journal of the Society for Industrial and Applied Mathematics Series A Control*, vol. 1, no. 2, pp. 152–192, 1963.

[2] D. F. Delchamps, *State space and input-output linear systems*. Springer-Verlag, 1988.

[3] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ error bounds," *International Journal of Control*, vol. 39, no. 6, pp. 1115–1193, 1984.

[4] 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.

[5] 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.

[6] 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.

[7] M. Hazewinkel and R. E. Kalman, "Moduli and canonical forms for linear dynamical systems," Econometric Institute, Erasmus University, Rotterdam, Tech. Rep. 7504, 1975.

[8] ——, "On invadants, 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.

[9] M. Hazewinkel, "Moduli and canonical forms for linear dynamical systems II: The topological case," *Mathematical Systems Theory*, vol. 10, pp. 363–385, 1977.

[10] ——, “Moduli and canonical forms for linear dynamical systems III: The algebraic-geometric case,” in *The 1976 AMES Research Center (NASA) Conference on Geometric Control Theory*, C. Martin and R. Hermann, Eds. Brookline, Mass.: Math Sci Press 1977, pp. 229–276.

[11] R. W. Brockett, “Some geometric questions in the theory of linear systems,” in *IEEE Conference on Decision and Control*, vol. 14, 1975, pp. 71–76.

[12] J. M. C. Clark, “The consistent selection of parameterizations in system identification,” in *Proceedings of Joint Automatic Control Conference*, 1976, pp. 576–580.

[13] P. S. Krishnaprasad, “Geometry of minimal systems and the identification problem,” Ph.D. dissertation, Harvard University, 1977.

[14] C. I. Byrnes and N. Hurt, “On the moduli of linear dynamical systems,” *Advances in Mathematical Studies in Analysis*, vol. 4, no. 83-122, 1979.

[15] D. F. Delchamps, “The geometry of spaces of linear systems with an application to the identification problem,” Ph.D. dissertation, Harvard University, 1982.

[16] ——, “Global structure of families of multivariable linear systems with an application to identification,” *Mathematical Systems Theory*, vol. 18, pp. 329–380, 1985.

[17] 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.

[18] B. Hanzon, *Identifiability, Recursive Identification and Spaces of Linear Dynamical Systems*. Centrum voor Wiskunde en Informatica (CWI), 1989, vol. 63-64.

[19] R. Peeters, *System Identification Based on Riemannian Geometry Theory and Algorithms*, ser. Tinbergen Institute Research Series. Amsterdam: Thesis Publishers, 1994, vol. 64.

[20] J. De Caigny, J. F. Camino, and J. Swevers, “Interpolation-based modeling of MIMO LPV systems,” *IEEE Transactions on Control Systems Technology*, vol. 19, no. 1, pp. 46–63, 2011.

[21] B. Afsari, R. Chaudhry, A. Ravichandran, and R. Vidal, “Group action induced distances for averaging and clustering linear dynamical systems with applications to the analysis of dynamic visual scenes,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

[22] B. Afsari and R. Vidal, “The alignment distance on spaces of linear dynamical systems,” in *IEEE Conference on Decision and Control*, 2013.

[23] ——, “Model order reduction for discrete-time LTI systems using the alignment distance,” in *American Control Conference*, 2015, pp. 2868–2875.

[24] ——, “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.

[25] M. Petreczky and R. Vidal, “Metrics and topology for nonlinear and hybrid systems,” in *Hybrid Systems: Computation and Control*. Springer Verlag, 2007.

[26] M. Petreczky and R. Peeters, “Spaces of nonlinear and hybrid systems representable by recognizable formal power series,” in *Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems–MTNS*, vol. 5, no. 9, 2010.

[27] D. G. Kendall, D. Barden, T. K. Carne, and H. Le, *Shape and Shape Theory*, ser. Wiley Series In Probability And Statistics. John Wiley & Sons, 1999.

[28] L. Younes, *Shapes and Diffeomorphisms*, ser. Applied Mathematical Sciences. Springer, 2010, vol. 171.

[29] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry Volume I*, ser. Wiley Classics Library Edition. John Wiley & Sons, 1963.

[30] R. W. Sharpe, *Differential geometry: Cartan’s generalization of Klein’s Erlangen program*. Springer Science & Business Media, 1997, vol. 166.

[31] J. R. Munkres, *Topology*. Prentice Hall, 2000.

[32] J. M. Lee, *Introduction to Smooth Manifolds*, ser. Graduate Texts in Mathematics. Springer, 2002.

[33] R. S. Palais, “On the existence of slices for actions of non-compact lie groups,” *Annals of mathematics*, pp. 295–323, 1961.

[34] G. E. Bredon, *Introduction to compact transformation groups*. Academic press, 1972, vol. 46.

[35] M. Hazewinkel, “On the (internal) symmetry groups of linear dynamical systems,” in *Groups, Systems and Many-Body Physics*, P. Kramer and M. D. Cin, Eds. Vieweg, 1980, pp. 362–404.

[36] M. Deistler, B. O. Anderson, A. Filler, C. Zinner, and W. Chen, “Generalized linear dynamic factor models: An approach via singular autoregressions,” *European Journal of Control*, vol. 3, pp. 211–224, 2010.

[37] C. Chou and B. Hanzon, “On diffeomorphisms between classes of linear systems,” *Systems & control letters*, vol. 26, no. 4, pp. 289–300, 1995.

[38] B. Hanzon, “On the differentiable manifold of fixed order stable linear systems,” *Systems and Control Letters*, vol. 13, pp. 345 – 352, 1989.

[39] U. Helmke, “A global parametrization of asymptotically stable linear systems,” *Systems & control letters*, vol. 13, no. 5, pp. 383–389, 1989.

[40] ——, “The topology of a moduli space for linear dynamical systems,” *Commentarii Mathematici Helvetici*, vol. 60, no. 1, pp. 630–655, 1985.

[41] R. J. Ober, “Balanced realizations: canonical form, parametrization, model reduction,” *International Journal of Control*, vol. 46, no. 2, pp. 643–670, 1987.

[42] K. Glover and J. C. Willems, “Parametrizations of linear dynamical systems: Canonical forms and identifiability,” *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 640–645, November 1974.

[43] D. Bauer and M. Deistler, “Balanced canonical forms for system identification,” *IEEE Transactions on Automatic Control*, vol. 44, no. 6, pp. 1118–1131, Juner 1999.

[44] B. Hanzon and R. J. Ober, “Overlapping block-balanced canonical forms for various classes of linear systems,” *Linear Algebra and its Applications*, vol. 281, no. 1, pp. 171–225, 1998.

[45] D. F. Delchamps, “A note on the analyticity of the Riccati metric,” in *Lectures in applied mathematics*, C. I. Bymes and C. F. Martin, Eds. American Mathematical Society, 1980, vol. 18, pp. 37–41.

[46] 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.

[47] 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.

[48] 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.

[49] E. A. Jonckheere and L. M. Silverman, “A new set of invariants for linear systems-application to reduced order compensator design,” *IEEE Transactions on Automatic Control*, vol. 28, no. 10, pp. 953–964, 1983.

[50] H. Karcher, “Riemannian center of mass and mollifier smoothing,” *Communications on pure and applied mathematics*, vol. 30, no. 5, pp. 509–541, 1977.

[51] M. Moakher, “A differential geometric approach to the geometric mean of symmetric positive-definite matrices,” *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 3, pp. 735–747, 2005.

[52] B. Afsari, “Riemannian L^p center of mass: Existence, uniqueness, and convexity,” *Proceedings of the AMS*, vol. 139, no. 2, pp. 655–673, 2011.

[53] S. Lang, *Fundamentals of differential geometry*. Springer-Verlag New York, 1999, vol. 191.

[54] U. Helmke and J. B. Moore, *Optimization and dynamical systems*. Springer-Verlag, 1994.

[55] E. Verriest, “Minimum sensitivity implementations for multi-mode systems,” in *Proceedings of the 27th IEEE Conference on Decision and Control*. IEEE, 1988, pp. 2165–2170.

[56] M. Green, “Balanced stochastic realizations,” *Linear Algebra and its Applications*, vol. 98, pp. 211–247, 1988.

[57] U. Helmke and K. Hüper, “A Jacobi-type method for computing balanced realizations,” *Systems & Control letters*, vol. 39, no. 19-30, 2000.

[58] B. O’Neill, “The fundamental equations of a submersion,” *The Michigan Mathematical Journal*, vol. 13, pp. 459–469, 1966.

[59] B. Hanzon and S. I. Marcus, “Riemannian metrics on spaces of stable linear systems, with applications to identification,” in *Proceedings of IEEE Conference on Decision & Control*, 1982, pp. 1119–1124.

[60] 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.

[61] ——, “Between controllable and uncontrollable,” *Systems & control letters*, vol. 4, no. 5, pp. 263–264, 1984.

[62] M. Wicks and R. DeCarlo, “Computing the distance to an uncontrollable system,” in *the 28th IEEE Conference on Decision and Control*. IEEE, 1989, pp. 2361–2366.

[63] 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.

[64] P. Parrilo and L. Ljung, “Initialization of physical parameter estimates,” in *Proceedings of the 13th IFAC Symposium on System Identification*, 2003, pp. 1524–1529.

[65] A. Lindquist and G. Picci, *Linear Stochastic Systems*. Springer, 2015.