The Alignment Distance on Spaces of Linear Dynamical Systems

Bijan Afsari and René Vidal

Abstract— We introduce a family of group action induced distances on spaces of Linear Dynamical Systems (LDSs) of fixed size and order. The distance between two LDSs is computed by finding the change of basis that best aligns the state-space realizations of the two LDSs, hence the name *alignment distance*. This distance can be computed efficiently, hence it is particularly suitable for applications in modern dynamic data analysis (e.g., video sequence classification and clustering), where a large number of high-dimensional LDSs may need to be compared. Based on the alignment distance, we also define a notion of average between LDSs of the same size and order with the property that the order and in some cases stability are naturally preserved. Various extensions to the basic notion of alignment distance are also proposed.

I. INTRODUCTION

In control and systems theory, the notion of a distance between Linear Dynamical Systems (LDSs) has a long history, and several such distances have been introduced for different applications, e.g., system identification [15], [21], robust control [11], and model reduction [12]. An important differentiating factor between such distances is the LDS *spaces* on which they are defined. For instance, while [21] and [15] are concerned with the spaces of LDSs of fixed output-input dimension (i.e., *size*) and *fixed* minimal *order*, the more commonly used gap metric is defined on the spaces of LDSs of fixed size but *arbitrary* (possibly infinite) order [11]. The distances may also differ in their computational frameworks, e.g., [21] and [15] define Riemannian distances in terms of state-space realization parameters[1], while the gap metric is defined in terms of transfer functions viewed as operators in a Hilbert space [11].

More recently, the need for comparing LDSs has arisen in fields such as computer vision. For example, it has been shown that one can use LDSs to effectively describe or model videos of human motion or activity (see, e.g., [1], [4], [5], [7], [10], and references therein). This is usually done via fitting an LDS to a video sequence for which efficient state-space system identification algorithms are available (see, e.g., [10] for details). While different people move differently, the dynamical models of two people performing the same action (e.g., walking) should be "closer" to each other than

the models of two people performing a different action (e.g., walking vs running) [5]. Obviously, here, a notion of distance between LDSs is useful for recognizing an action and other purposes.

In practice, the LDSs appearing in such applications are of a special form which we call *tall and full rank* (see §II and §II-D.1 for more details). Moreover, often all the LDSs are of the same size and order. Hence, in such an application a space of LDSs of fixed size and order is a natural choice. Typically, the LDSs are of very large output dimension ($10^3 - 10^4$) and low input dimension and order ($3 - 10$), and one usually deals with tens of thousands of them. This warrants for efficiently computable distances on the spaces of LDSs of fixed size and order, for which existing frameworks (e.g., as in [21] or [14]) are not suited, due to their inherent heavy computational cost.

A valid question is "Why not use a more readily available distance?" In fact, a target space, namely, the space of LDSs of size (p,m) and order n can be viewed as an *embedded* finite dimensional subspace in the *infinite* dimensional *ambient* space of LDSs of size (p,m) and of arbitrary order. The ambient space is a linear space, and any distance there induces a distance on the target space. An example is the commonly used distance based on the L^2 norm of the difference between two transfer functions. The main problem with this distance, the gap metric, or other similar distances, is that they are too *coarse* and blind to the order of LDSs (unless one specializes them to the target subspace, which is challenging). This problem becomes pronounced, e.g., if one wants to use such a distance to define a notion of *average* on the target space. Averaging LDSs is useful because it allows us to replace a large set of LDSs with a single representative LDS (see [1] for an application in video sequence classification and clustering). Finding an average LDS using the mentioned L^2 based distance (as the minimizer of the sum of square of distances) amounts to finding the Euclidean average of the corresponding transfer functions. Thus the average of N LDSs of order n is (generically) of order nN, which can be extremely large. It is also not clear if model reduction methods are practically applicable and justified here, where such a drastic order reduction is desired. Instead, a notion of average, which automatically preserves system order is a more natural solution, and for that a distance on the target space cognizant of system order is needed. We add that to our knowledge defining such a notion of average has not been studied systematically. An exception is the similar notion of interpolation in gain scheduling control, where one interpolates between two controllers of the same order to get a new one [25].

[1]To be clear, we should mention that in these and similar works (for system identification purposes) the main object of interest has not been a Riemannian distance itself rather its infinitesimal form (a.k.a. the Riemannian metric), which in comparison is a computationally friendly object. On the other hand, finding the Riemannian distance (essentially integrating the Riemannian metric) involves solving systems of nonlinear ODEs, hence is computationally expensive.

Contributions, scope, and outline. In this paper, we introduce a large family of easy-to-compute group action induced distances (called the *alignment* distances) on spaces of LDSs of fixed size and order. The computational advantage stems from the fact that such a distance is found by a simple minimization problem over a (matrix) rotation group. The basic idea behind the alignment distance is to compare *state-space realizations* of two LDSs considering the fact that each LDS has an equivalent class of realizations (related via a state-space change of basis). Hence, one needs to *align* the realizations (in a specific sense) and then compare them. In our earlier work [1], we introduced the alignment distance for the class of tall and full rank LDSs. In this paper (see §II), we use methods from differential geometry and the theory of balanced realizations [17] to extend this definition to other classes of LDS spaces of fixed size and order. In §III, we show that the alignment distance naturally leads to a notion of average which preserves the system order. In §IV, we show how the basic definition can be extended to other useful and application-oriented distances. The paper conclusions are presented in §V. This is a theoretical paper entirely devoted to discrete-time *deterministic* LDSs, and due to space limitation the proofs of most stated results will appear elsewhere. Some *computational* aspects and the case of *stochastic* LDSs have been addressed in [19] and [2], respectively.

II. THE ALIGNMENT DISTANCE

Consider a discrete-time LDS M of *order* n and *size* (p, m) described by:

$$
\begin{aligned}
x_t &= Ax_{t-1} + Bv_t \\
y_t &= Cx_t,
\end{aligned}
\tag{1}
$$

where v_t, x_t, and y_t are the input, state, and output signals, respectively. We explicitly distinguish between a realization $R = (A, B, C)$ and M the LDS it realizes.

Define the realization space $\widetilde{\mathcal{L}}_{m,n,p}$ as

$$
\widetilde{\mathcal{L}}_{m,n,p} = \{(A, B, C) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{p \times n}\}. \tag{2}
$$

We let $GL(n)$ denote the Lie group of $n \times n$ non-singular matrices. An important fact is that, as far as input-output relation is concerned, for any $P \in GL(n)$, R and $P \circ R$, where

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

represent the same system. Formally, one says that $GL(n)$ *acts* on the space of realizations $\widetilde{\mathcal{L}}_{m,n,p}$ via the above *action* (strictly speaking this is a *right* action because $(P_1 P_2) \circ R = P_2 \circ (P_1 \circ R)$, but for notational convenience we write it as a *left* action). Now the space of LDSs of order n and size (p, m), denoted by $\mathcal{L}_{m,n,p}$, is the quotient space of this action, i.e., $\mathcal{L}_{m,n,p} = \widetilde{\mathcal{L}}_{m,n,p}/GL(n)$. On $\mathcal{L}_{m,n,p}$ we define the notion of equality:

Definition 1: Two systems in $\mathcal{L}_{m,n,p}$ are equal if their impulse responses are equal.

If a Lie group acts on a smooth manifold *smoothly*, *freely* and *properly*, then the quotient space will be a *smooth manifold* whose dimension is equal to the difference between the

dimension of the original manifold and that of the group [22, p. 218]. The action in (3) is clearly smooth (where smoothness is induced from the standard topology of $\widetilde{\mathcal{L}}_{m,n,p}$). Recall that an action of a Lie group on a manifold $\mathcal{M}$ is free if the identity element of the group is the only element that fixes an element of $\mathcal{M}$. Properness is a more technical notion about the properties of the orbits of the action and we leave it for Appendix I. The action of a compact group is always proper. In Proposition 7[2] in Appendix I we show that the action (3) is free and proper on the subspaces of realizations which are observable, controllable, minimal, with $\text{rank}(C) = n$, or with $\text{rank}(B) = n$ (the latter two exist only when $p \geq n$ and $m \geq n$, respectively). Let us denote the realization spaces in these cases by $\widetilde{\mathcal{L}}_{m,n,p}^{\text{ob}}$, $\widetilde{\mathcal{L}}_{m,n,p}^{\text{co}}$, $\widetilde{\mathcal{L}}_{m,n,p}^{\text{min}}$, $\widetilde{\mathcal{L}}_{m,n,p}^{\text{tC}}(\subset \widetilde{\mathcal{L}}_{m,n,p}^{\text{ob}})$, and $\widetilde{\mathcal{L}}_{m,n,p}^{\text{fB}}(\subset \widetilde{\mathcal{L}}_{m,n,p}^{\text{co}})$, respectively. All these spaces are, in fact, smooth (open and dense) submanifolds of $\widetilde{\mathcal{L}}_{m,n,p}$. The corresponding (smooth quotient) manifolds of systems will be denoted with the same letters but without the $\sim$, e.g., $(\widetilde{\mathcal{L}}_{m,n,p}^{\text{min}}, \mathcal{L}_{m,n,p}^{\text{min}})$ where $\mathcal{L}_{m,n,p}^{\text{min}} = \widetilde{\mathcal{L}}_{m,n,p}^{\text{min}}/GL(n)$. We call every such *realization-system space pairs* a proper realization-system space pair. Each pair is an example of a *principal fiber bundle* with the structure $GL(n)$ [20]. The differential geometry of $(\widetilde{\mathcal{L}}_{m,n,p}^{\text{min}}, \mathcal{L}_{m,n,p}^{\text{min}})$, in particular, had been the subject of extensive studies in the 70's and 80's (see e.g., [14], [9], [8], [21] and references therein). In occasions, it is more convenient to use a generic proper pair $(\widetilde{\Sigma}_{m,n,p}, \Sigma_{m,n,p})$ rather than a specific one. The next proposition is essential, since it shows when equality in $\mathcal{L}_{m,n,p}$ translates into equal realization equivalent classes.

Proposition 2: Let $\widetilde{\Sigma}_{m,n,p}$ be any of the above proper realization spaces. Let M be a system in $\mathcal{L}_{m,n,p}$ that has a realization R_1 in $\widetilde{\Sigma}_{m,n,p}$. Then for every $P \in GL(n)$, $P \circ R_1$ is also a realization of M in $\widetilde{\Sigma}_{m,n,p}$. Conversely, for any other realization of R_2 of M in $\widetilde{\Sigma}_{m,n,p}$, there exists a unique $P \in GL(n)$ such that $R_2 = P \circ R_1$.

A. Difficulties due to the non-compactness of $GL(n)$

Let $(\widetilde{\Sigma}_{m,n,p}, \Sigma_{m,n,p})$ be a proper realization-system space pair. Defining a distance on $\Sigma_{m,n,p}$ is conceptually straightforward (see [26, Ch. 12] for a general theory): One starts with a $GL(n)$-invariant distance $\tilde{d}(\cdot, \cdot)$ on $\widetilde{\Sigma}_{m,n,p}$ and defines the distance between two systems $M_1, M_2 \in \Sigma_{m,n,p}$ as $d(M_1, M_2) = \inf_{P \in GL(n)} \tilde{d}(P \circ R_1, R_2)$, where R_1 and R_2 are any realizations of M_1 and M_2, respectively. Recall that $\tilde{d}$ is $GL(n)$-invariant if $\tilde{d}(P \circ R_1, P \circ R_2) = \tilde{d}(R_1, R_2)$ for $\forall P \in GL(n)$. Unfortunately, constructing such a $\tilde{d}$ due to the *non-compactness* of $GL(n)$ is computationally difficult. Most available distances (e.g., matrix norm-based ones) are only unitary-invariant. If the goal is to construct a Riemannian distance on $\Sigma_{m,n,p}$, one can define a $GL(n)$-invariant Riemannian structure on $\widetilde{\Sigma}_{m,n,p}$ (hence a $GL(n)$-invariant distance) and then solve the corresponding geodesic

---

[2]This, presumably, is a known result. However, we found no published proof for it. An alternative proof that LDSs of size (p, m) and minimal order n form a smooth manifold of dimension $n(m + p)$ is given in [8].

equation and compute distances between LDSs (this was proposed by Krishnaprasad in [21]). However, the computational challenge is that the geodesic equation is a system of high-dimensional *nonlinear* ODEs.

B. Reduction of the structure to maximal compact subgroup

A fundamental observation we make is that the action of $GL(n)$ on $\widetilde{\Sigma}_{m,n,p}$ can be reduced to the action of its subgroup $O(n)$, the group of $n \times n$ orthogonal matrices. This is an example of *bundle or structure group reduction* [20], and in the theory of principal fiber bundles it has a precise meaning: One can find a submanifold of $\widetilde{\Sigma}_{m,n,p}$ (not necessarily unique) denoted by $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ on which $P \circ R \in \widetilde{\mathcal{O}\Sigma}_{m,n,p}$ implies $P \in O(n)$. Moreover, $\Sigma_{m,n,p} \stackrel{\text{diff}}{=} \widetilde{\mathcal{O}\Sigma}_{m,n,p}/O(n)$, meaning that the two sides are not only equal as sets but also diffeomorphic to each other. Intuitively, the meaning of this reduction is that one can get rid of the non-compact part of $GL(n)$ (i.e., keep its *maximal* compact subgroup) while the differential-topological structure of the quotient is unchanged. We call this reduction *standardization* or *orthonormalization* of the realization space $\widetilde{\Sigma}_{m,n,p}$. This process is analogous to Gramm-Schmidt orthonormalization.

The main benefit of standardization for us is computational: constructing $O(n)$-invariant distances on $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ is straightforward (see §II-C). Assuming an $O(n)$-invariant distance on $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$, $\tilde{d}_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}$, one can define a distance on $\Sigma_{m,n,p}$ simply by

$$d_{\Sigma_{m,n,p}}(M_1, M_2) = \min_{Q \in O(n)} \tilde{d}_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}(Q \circ R_1, R_2), \quad (4)$$

where R_1 and R_2 are any realizations of M_1 and M_2 in $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$, respectively. We call this problem the *realization alignment problem*, or simply the alignment problem, and denote a solution to it by Q_{12}. We call $d_{\Sigma_{m,n,p}}$ an *alignment distance* (see Figure 1). In general, the alignment problem is not a convex minimization problem and its solution may not be unique. Recall that a similar situation may happen in a Riemannian manifold where, e.g., there may be more than one shortest geodesic between two points.

Fig. 1. Over each system in $\Sigma_{m,n,p}$ sits a realization fiber. The fibers together form the realization space (bundle) $\widetilde{\Sigma}_{m,n,p}$. $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ is a standardized subspace of $\widetilde{\Sigma}_{m,n,p}$. Constructing an $O(n)$-invariant distance on the standardized subspace and hence the corresponding alignment distance on $\Sigma_{m,n,p}$ is computationally easy (see (4)).

C. A simple $O(n)$-invariant distance on realization spaces

The most immediate example of an $O(n)$-invariant distance on $\widetilde{\Sigma}_{m,n,p}$ is the Frobenius norm based distance:

$$\tilde{d}_F^2(R_1, R_2) = \lambda_A \|A_1 - A_2\|_F^2 + \lambda_B \|B_1 - B_2\|_F^2 + \lambda_C \|C_1 - C_2\|_F^2, \quad (5)$$

where $\lambda_A, \lambda_B, \lambda_C > 0$ are weights (which could be tuned in some applications). The use of the Frobenius norm makes this distance as computationally friendly as one could hope for. In [19], a fast Jacobi-type algorithm for finding the alignment distance (4) based on $\tilde{d}_F$ is proposed. Such an algorithm is a local minimization algorithm over $O(n)$, however, an interesting experimental observation made in [19] is that this algorithm is more likely to find the global minimizer than the gradient algorithm. Another feature of $\tilde{d}_F$ is that it is a product distance, i.e., it is the sum of three separate terms coming from the product structure of the ambient space $\widetilde{\mathcal{L}}_{m,n,p} \supset \widetilde{\Sigma}_{m,n,p}$. This feature is advantageous in some related optimization problems (see §III). Finally, note that $\tilde{d}_F$ is an *extrinsic* distance on $\widetilde{\Sigma}_{m,n,p}$ (and its submanifolds) meaning that it need not respect the nonlinearities of $\widetilde{\Sigma}_{m,n,p}$, e.g., in the case of $\widetilde{\mathcal{L}}_{m,n,p}^{\min}$ the distance is blind to the minimality of the realizations. Nevertheless, see $\tilde{d}_{\text{ST}}$ in §II-D.1 which does better and yields an *intrinsic* alignment distance on $\mathcal{L}_{m,n,p}^{\text{tC}}$. In general, constructing a distance which takes into account all nonlinearities of a space is challenging. Even in the Riemannian case defining a Riemannian metric does not automatically guarantee this, and one has to judiciously design a *complete* Riemannian metric, i.e., one in which geodesics never leave the space in finite time [20].

D. Examples of standardized realization spaces

Next, we introduce some realization standardization methods that yield $O(n)$-subbundles. First, we establish some notations. Denote by $\mathcal{O}_k = [C^\top, (CA)^\top, \ldots, (CA^{k-1})^\top]^\top$ and $\mathcal{C}_k = [B, AB, \ldots, A^{k-1}B]$ the observability and controllability matrices of order k ($n \leq k \leq \infty$). The observability and controllability Gramians of order k are defined as $W_{\text{o},k} = \mathcal{O}_k^\top \mathcal{O}_k$ and $W_{\text{c},k} = \mathcal{C}_k \mathcal{C}_k^\top$, respectively. For $k = \infty$ asymptotic stability of A is needed. These matrices are realization dependent and under the action (3) they transform in obvious ways: $\mathcal{O}_k \to \mathcal{O}_k P$, $\mathcal{C}_k \to P^{-1} \mathcal{C}_k$, $W_{\text{o},k} \to P^\top W_{\text{o},k} P$, and $W_{\text{c},k} \to P^{-1} W_{\text{c},k} P^{-\top}$ (the last two coincide if $P \in O(n)$).

1) Orthonormalized observation matrix (OOM): This standardization is specific to $(\widetilde{\mathcal{L}}_{m,n,p}^{\text{tC}}, \mathcal{L}_{m,n,p}^{\text{tC}})$ which we call the tall and full rank realization-system space pair, but similarly can be applied to $(\widetilde{\mathcal{L}}_{m,n,p}^{\text{fB}}, \mathcal{L}_{m,n,p}^{\text{fB}})$. Let $R = (A, B, C) \in \widetilde{\mathcal{L}}_{m,n,p}^{\text{tC}}$ and let $C = USV^\top$ be a thin SVD of C. With the change of coordinates $R' = (A', B', C') = P' \circ R$, where $P' = (SV^\top)^{-1}$ we have $C'^\top C' = I_n$. Let us define

$$\widetilde{\mathcal{OL}}_{m,n,p}^{\text{tC}} = \{(A, B, C) \in \widetilde{\mathcal{L}}_{m,n,p}^{\text{tC}} | C^\top C = I_n\}. \quad (6)$$

We then have:

Proposition 3: $\widetilde{\mathcal{OL}}_{m,n,p}^{\text{tC}}$ is an $O(n)$-subbundle of $\widetilde{\mathcal{L}}_{m,n,p}^{\text{tC}}$ with codimension $\frac{n(n+1)}{2}$.

Interestingly, $\widetilde{\mathcal{OL}}_{m,n,p}^{\text{tC}}$ naturally appears in a PCA-based system identification method employed often in computer vision applications [10]. The realization space $\widetilde{\mathcal{OL}}_{m,n,p}^{\text{tC}}$ equipped with the distance $\tilde{d}_F$ (5) was used in [1] to define a distance on $\mathcal{L}_{m,n,p}^{\text{tC}}$. Also note that in $\widetilde{\mathcal{OL}}_{m,n,p}^{\text{tC}}$, C belongs to the Stiefel manifold $\text{ST}(p,n)$. Therefore, instead of $\tilde{d}_F$ an alternative intrinsic distance on $\widetilde{\mathcal{L}}_{m,n,p}^{\text{tC}}$ could be defined as

$$\tilde{d}_{\text{ST}}^2(R_1, R_2) = \|A_1 - A_2\|_F^2 + \|B_1 - B_2\|_F^2 + d_{\text{ST}(p,n)}^2(C_1, C_2), \tag{7}$$

where $d_{\text{ST}(p,n)}$ is the standard Riemannian distance on $\text{ST}(p,n)$. This distance induces an intrinsic (Riemannian) distance on $\mathcal{L}_{m,n,p}^{\text{tC}}$. However, for very large p (e.g., in video sequence analysis) the computations can become prohibitive.

2) Output-normalized realizations: The output-normalized realization is a generalization of the OOM realization in which instead of orthonormalizing the observation matrix C, we orthonormalize the observability matrix $\mathcal{O}_k$. Since $\text{rank}(\mathcal{O}_k) = n$, we can orthonormalize $\mathcal{O}_k$ by the change of coordinates $R' = (A', B', C') = P' \circ R$, with $P' = (SV^\top)^{-1}$, where $\mathcal{O}_k = USV^\top$. We define the space of observable realizations with orthonormalized observability matrix (the order k is implicitly assumed) as:

$$\widetilde{\mathcal{OL}}_{m,n,p}^{\text{ob,on}} = \{(A, B, C) \in \widetilde{\mathcal{L}}_{m,n,p}^{\text{ob}} | W_{\text{o},k} = I_n\}, \tag{8}$$

for which we have:

Proposition 4: $\widetilde{\mathcal{OL}}_{m,n,p}^{\text{ob,on}}$ is an $O(n)$-subbundle of $\widetilde{\mathcal{L}}_{m,n,p}^{\text{ob}}$ with codimension $\frac{n(n+1)}{2}$.

3) Balanced realizations (in the sense of Helmke [17]): We follow [17] in defining a balanced realization, which is (slightly) different from the (diagonal) balanced realization originally introduced in [23]. For $k = \infty$ the extra assumption of asymptotic stability of A is needed, in which case we consider the space of minimal asymptotically stable realizations $\widetilde{\mathcal{L}}_{m,n,p}^{\text{min,a}}$ and the corresponding system space $\mathcal{L}_{m,n,p}^{\text{min,a}}$. The space of balanced realizations is defined by

$$\widetilde{\mathcal{OL}}_{m,n,p}^{\text{min,bl}} = \{(A, B, C) \in \widetilde{\mathcal{L}}_{m,n,p}^{\text{min}} | W_{\text{o},k} = W_{\text{c},k}\}. \tag{9}$$

As shown in [17] this space can also be characterized as follows: Consider the function $h : GL(n) \to \mathbb{R}$ where $h(P; R) = \text{trace}(P^\top W_{\text{o},k} P) + \text{trace}(P^{-1} W_{\text{c},k} P^{-\top})$ and $R \in \widetilde{\mathcal{L}}_{m,n,p}^{\text{min}}$. The realization R can be transformed to a balanced one by minimizing $h(P; R)$ over $GL(n)$. Clearly $h(P; R)$ is invariant on $O(n)$. It can be shown that any critical point of h is a global minimizer. Moreover, for any two critical points $P_1, P_2 \in GL(n)$ there is a unique $\Theta \in O(n)$ such that $P_1 = P_2\Theta$, and the realization $P_1 \circ R$ belongs to $\widetilde{\mathcal{OL}}_{m,n,p}^{\text{min,bl}}$ [17]. Based on this we can show:

Proposition 5: $\widetilde{\mathcal{OL}}_{m,n,p}^{\text{min,bl}}$ is an $O(n)$-subbundle of $\widetilde{\mathcal{L}}_{m,n,p}^{\text{min}}$ with codimension $\frac{n(n+1)}{2}$.

The original realization balancing (more precisely *diagonally* balancing) as defined in [23] makes the controllability and observability Gramians not only equal but also diagonal. Therefore, such a realization also belongs to $\widetilde{\mathcal{OL}}_{m,n,p}^{\text{min,bl}}$, and any algorithms for diagonal balancing can be used to balance

realizations in the sense of Helmke. Also we add that the above variational approach has been used in [17], [18] to define a large class of balanced realizations which could be shown to correspond to standardization or bundle reduction as defined here. Accordingly, they all can be used to define a large family of distances on the LDS spaces.

E. Digression on continuous canonical forms

The reader might ask why we are not using canonical forms as vehicles to define distances. Recall that a (global) canonical form $s : \Sigma_{m,n,p} \to \widetilde{\Sigma}_{m,n,p}$ is a function that assigns to each system $M \in \Sigma_{m,n,p}$ a unique realization $s(M)$ in the equivalent class of realizations of M in $\widetilde{\Sigma}_{m,n,p}$. For our applications we need the canonical form to be additionally *continuous* (otherwise the realizations of two nearby systems might be far). Having such a continuous canonical form and a distance $\tilde{d}$ on $\widetilde{\Sigma}_{m,n,p}$, we can define the distance $d(M_1, M_2)$ as $\tilde{d}(s(M_1), s(M_2))$ for every $M_1, M_2 \in \Sigma_{m,n,p}$. However, a fundamental problem is that for most of our realization-system space pairs no continuous canonical form exists. As mentioned before the realization space $\widetilde{\Sigma}_{m,n,p}$ is a principal fiber bundle over $\Sigma_{m,n,p}$ with the structural group $GL(n)$. The non-existence of (*global*) continuous canonical forms is the consequence of a basic fact in the theory of principal fiber bundles that only a trivial principal fiber admits a continuous cross section. Triviality means that $\widetilde{\Sigma}_{m,n,p}$ must be (globally) diffeomorphic to the product $\widetilde{\Sigma}_{m,n,p} \times GL(n)$. It has been established that the realization bundles $\widetilde{\mathcal{L}}_{m,n,p}^{\text{ob}}, \widetilde{\mathcal{L}}_{m,n,p}^{\text{co}}, \widetilde{\mathcal{L}}_{m,n,p}^{\text{min}}$ are trivial only when $m = 1$ or $p = 1$ hence they do not admit continuous canonical forms when $m > 1$ or $p > 1$ (see [16]). This fact is, actually, a major culprit in making the identification of MIMO systems a challenging task (see e.g., [13] and [3]). On the other hand, it is easy to see that $\widetilde{\mathcal{L}}_{m,n,p}^{\text{tC}}$ and $\widetilde{\mathcal{L}}_{m,n,p}^{\text{fB}}$ are trivial; and hence we can have continuous canonical forms over $\mathcal{L}_{m,n,p}^{\text{tC}}$ and $\mathcal{L}_{m,n,p}^{\text{fB}}$. However, even in this case the alignment distance is preferable. To see this consider Figure 2, where a global continuous canonical form s exists on the realization bundle and a group-invariant distance $\tilde{d}_{\widetilde{\Sigma}_{m,n,p}}(\cdot, \cdot)$ is available. Consider two systems M_1 and M_2 with canonical forms $s(M_1) = R_1$ and $s(M_2) = R_2$. Also assume that Q_{12} aligns R_1 to R_2. In general, the canonical form is not adapted to the distance and we have $\tilde{d}_{\widetilde{\Sigma}_{m,n,p}}(R_1, R_2) > \tilde{d}_{\widetilde{\Sigma}_{m,n,p}}(Q_{12} \circ R_1, R_2)$, which indicates that by aligning the realizations we get a more realistic distance.

III. How to average LDSs?

Now, we briefly address the LDS averaging problem introduced in §I. We are given an LDS set $\{M_i\}_{i=1}^N \subset \Sigma_{m,n,p}$. Next, we choose a suitable standardized realization space $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ and find corresponding realizations $\{R_i\}_{i=1}^N \subset \widetilde{\mathcal{O}\Sigma}_{m,n,p}$. For simplicity assume that we use the Frobenius based alignment distance pair $(\tilde{d}_F, d_{\Sigma_{m,n,p}})$. We define the average (Fréchet mean) of $\{M_i\}_{i=1}^N$ as the (not necessarily
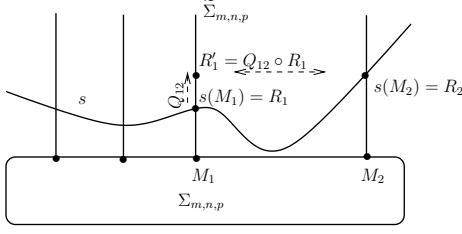
Fig. 2. If a canonical form exists on a realization-system space pair and a group-invariant distance is available, then still the group action induced distance is more suitable than the distance induced by the canonical form.

unique) system

$$\bar{M} = \operatorname{argmin}_{M \in \Sigma_{m,n,p}} \sum_{i=1}^{N} d_{\Sigma_{m,n,p}}^2 (M_i, M). \qquad (10)$$

We give a simple approximate iterative algorithm for solving this problem. Let $\bar{M}^0$ be an initial estimate of the average with a realization $\bar{R}^0 = (\bar{A}^0, \bar{B}^0, \bar{C}^0) \in \widetilde{\mathcal{O}\Sigma}_{m,n,p}$. At step $k+1$, where $\bar{R}^k$ is available, we align each R_i to $\bar{R}^k$ and find the aligning matrix $Q_i^k \in O(n)$ (see (4)). Next we need to solve $\min_{R \in \widetilde{\mathcal{O}\Sigma}_{m,n,p}} \sum_{i=1}^{N} \tilde{d}_F^2(Q_i^k \circ R_i, R)$. Despite the apparent decoupling of the three terms for A, B, and C in $\tilde{d}_F^2$ since $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ is, in general, non-trivial this (constrained) minimization would be difficult to solve. Instead, we solve the unconstrained version of the problem, i.e., assuming $R = (A, B, C) \in \widetilde{\mathcal{L}}_{m,n,p}$ and then project the result onto $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$. Since the unconstrained problem decouples, one needs to find simple Euclidean averages, namely, $\hat{A}^k = \frac{1}{N} \sum_{i=1} Q_i^{k\top} A_i Q_i^k$, $\hat{B}^k = \frac{1}{N} \sum_{i=1} Q_i^{k\top} B_i$ and $\hat{C}^k = \frac{1}{N} \sum_{i=1} C_i Q_i^k$. Next, one finds $\bar{R}^{k+1}$ by projecting $\hat{R}^k$ to $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ and then repeats the above steps.

This algorithm for $\mathcal{L}_{m,n,p}^{\text{tc}}$ was used in [1] for classification/clustering of video sequences. The projection step may be costly, but if the realization space $\widetilde{\Sigma}_{m,n,p}$ is dense in $\widetilde{\mathcal{L}}_{m,n,p}$, instead of the projection we may simply perform standardization of $\bar{R}^k$ (the two are, in general, different). If the elements of $\Sigma_{m,n,p}$ are asymptotically stable, then a standardization of $\hat{R}^k$ might not be enough. However, if $\|A_i\|_2$, the 2-norm of the A_i, is smaller than 1 for every i, then $\|\hat{A}^k\|_2 < 1$ (hence $\hat{A}^k$ is asymptotically stable). Interestingly, the use of $\widetilde{\mathcal{O}\mathcal{L}}_{m,n,p}^{\text{min,bl}}$ automatically takes this into account, since in this subbundle any asymptotically stable LDS has realizations with $\|A\|_2 \leq 1$ and generically the strict inequality holds (see II-D.3 and [24]); hence, generically order and stability will be preserved. The convergence properties of the algorithm have yet to be studied.

IV. GENERALIZATIONS AND EXTENSIONS

The introduced methodology of group action induced distances is quite flexible and general in the sense that it allows us to define a large family of distances over a given space of systems. Here, we briefly discuss some extensions and generalizations. The most immediate extension is the use of distances other than $\tilde{d}_F$. We already alluded to $\tilde{d}_{\text{ST}}$ in

(7). Another interesting example could be the matrix nuclear norm. Next, we discuss some less trivial extensions.

A. General input, internal, and output symmetries

The action $\circ$ is an example of what we call *internal symmetry*. More generally, one can think of other internal, input, or output symmetries. For example, in the case of stochastic systems the group $O(m)$ induces a symmetry at the input. This is because when v_t is the standard Gaussian input noise so is Θv_t for any $\Theta \in O(m)$. Therefore, $R = (A, B, C)$ and $(Q, \Theta) \bullet R = (Q^\top A Q, Q^\top B \Theta, C Q)$ $((Q, \Theta) \in O(n) \times O(m))$ when stimulated with v_t yield the same power spectral density and hence the same output processes (see [2] for more details).

As another example, we could consider a situation where a group $G = G_{\text{out}} \times G_{\text{int}} \times G_{\text{in}}$ acts on a realization space $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$, where the groups G_{int}, G_{in}, and G_{out} act at the input, internally, and at the output, respectively. Denote the action by $\star$. If a G-invariant distance $d_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}$ is given, we can define a distance as:

$$d_{\Sigma_{m,n,p}}(M_1, M_2) = \min_{\Omega \in G} \tilde{d}_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}(\Omega \star R_1, R_2). \qquad (11)$$

We mention that in certain cases the group G may be a discrete group in which case solving this minimization problem would more challenging than our previous examples.

There are interesting cases where G is non-compact and finding a G-invariant distance $\tilde{d}_{\widetilde{\mathcal{O}\Sigma}_{m,n,p}}$ is difficult, yet (11) could give a distance-like measure. One such example is scaling at the input and output, in which case (A, B, C) and $(Q, \alpha) \star (A, B, C) = (Q^\top A Q, \alpha Q^\top B, \alpha C Q)$ are considered equivalent for every $Q \in O(n)$ and $\alpha \in \mathbb{R}^+$. Another interesting case is when two LDSs are considered equivalent if one is a slow (or fast) copy of another one (e.g., in the video sequence application the LDSs modeling slow walking and fast walking may be considered equal). Assuming the discrete-time LDSs are coming from a continuous-time stable systems, one might consider $(Q, l) \star (Q^\top A^l Q, Q^\top B, C Q)$ and (A, B, C) equivalent for every l in $\mathbb{R}^+$ (or nonnegative integers) and solve a minimization problem similar to (11) to define a "speed invariant" distance measure. This is another example of what we called internal symmetry. There are other types of symmetries associated with LDSs studied in [6], which could potentially be used to define interesting distances in like manner as above.

B. Distances via diffeomorphic and related spaces

Here, we consider spaces diffeomorphic to $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ or some subspaces in this space to define new distances. For example, for any positive integer l, $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ is diffeomorphic to $\{(A, A^2, \ldots, A^l, B, C) | (A, B, C) \in \widetilde{\mathcal{O}\Sigma}_{m,n,p}\}$ on which $GL(n)$ acts in an obvious way. One can define a new distance on $\widetilde{\mathcal{O}\Sigma}_{m,n,p}$ (and hence on $\Sigma_{m,n,p}$) by considering the powers of A:

$$\tilde{d}^2(R_1, R_2) = \sum_{i=1}^{l} \|A_1^i - A_2^i\|_F^2 + \|B_1 - B_2\|_F^2 + \|C_1 - C_2\|_F^2. \qquad (12)$$

Such a distance might be useful in certain applications where the role of A is important. A similar idea in the case of asymptotically stable LDSs yields a new distance where in (5) instead of the first term we have $\|(I_n - A_1)^{-1} - (I_n - A_2)^{-1}\|_F^2$. As another example, we might consider the subspace $\{(A,B,C) \in \widetilde{\mathcal{OL}}_{m,n,p}^{\mathrm{min,bl}} | \|B\|_F = \|C\|_F = 1\}$ and use (5) or any of the other distances on this subspace to get an alignment distance on the LDS space. This is an alternative way to define a scale-invariant distance.

V. Conclusions

We introduced a class of group action induced distances on spaces of LDSs of fixed size and order, which we collectively called alignment distances. Our emphasis was on extrinsic distances since they are computationally favorable. Our framework is extendible and flexible, and can be used to construct new distances and perform statistical analysis (e.g., find averages) on spaces of LDSs. We did not address any specific control application here, but we expect that the alignment distance and the notion of average LDS can be useful in areas such as system identification, robust control, model order reduction, and linear parameter varying systems.

Appendix I
Proper Actions

The following is a working definition of a proper action [22, Ch. 9]:

Definition 6: Let $\circ$ be a smooth action of a Lie group G on a manifold $\mathcal{M}$. The action $\circ$ is called proper if the following holds: If $\{m_i\}_i$ is a convergent sequence in $\mathcal{M}$ and $\{g_i\}_i$ is a sequence in G such that $\{g_i \circ m_i\}_i$ converges in $\mathcal{M}$, then a subsequence of $\{g_i\}_i$ converges in G.

Proposition 7: Let $\widetilde{\Sigma}_{m,n,p}$ be either $\widetilde{\mathcal{L}}_{m,n,p}^{\mathrm{ob}}$, $\widetilde{\mathcal{L}}_{m,n,p}^{\mathrm{co}}$, $\widetilde{\mathcal{L}}_{m,n,p}^{\mathrm{min}}$, $\widetilde{\mathcal{L}}_{m,n,p}^{\mathrm{tC}}$, and $\widetilde{\mathcal{L}}_{m,n,p}^{\mathrm{fB}}$ or their stable or minimum phase submanifolds. Then $GL(n)$ acts properly and freely on $\widetilde{\Sigma}_{m,n,p}$ via (3).

Proof: We only prove the result for $\widetilde{\mathcal{L}}_{m,n,p}^{\mathrm{ob}}$, the rest are similar. First, let $R \in \widetilde{\mathcal{L}}_{m,n,p}^{\mathrm{ob}}$, and $P \circ R = R$ for some $P \in GL(n)$. It follows that $\mathcal{O}P = \mathcal{O}$, where $\mathcal{O}$ is the (full rank) observation matrix of order n. This implies freeness of the action, because we must have $P = I_n$. Now, to see properness, assume that $\{R_i\}_i$ is a sequence in $\widetilde{\mathcal{L}}_{m,n,p}^{\mathrm{ob}}$ converging to $R \in \widetilde{\mathcal{L}}_{m,n,p}^{\mathrm{ob}}$ and $\{P_i \circ R_i\}_i$ is converging to $\bar{R} \in \widetilde{\mathcal{L}}_{m,n,p}^{\mathrm{ob}}$. 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 $\bar{R}$. 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 to $\bar{\mathcal{O}}$. Let $\mathcal{O}_i^\dagger$ denote the left pseudo inverse of $\mathcal{O}_i$ (this is unique since $\mathcal{O}_i$ is full rank). Note that since $\{\mathcal{O}_i\}_i$ converges to $\mathcal{O}$ and $\mathcal{O}$ is full rank, for large enough i there is K such that $\|\mathcal{O}_i^\dagger\|_2 < K$, where $\|.\|_2$ is the matrix 2-norm. Therefore, we have $\|P_i - \mathcal{O}_i^\dagger \bar{\mathcal{O}}\|_2 \leq K\|\mathcal{O}_i P_i - \bar{\mathcal{O}}\|_2$ for large enough i. This implies that $\{P_i\}_i$ converges to $\mathcal{O}^\dagger \bar{\mathcal{O}} \in GL(n)$, which suffices for properness. ∎

References

[1] 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.

[2] B. Afsari and R. Vidal. Group action induced distances on spaces of high-dimensional linear stochastic processes. In *Geometric Science of Information*, volume 8085 of *LNCS*, pages 425–432. Springer, 2013.

[3] D. Bauer and M. Deistler. Balanced canonical forms for system identification. *IEEE Transactions on Automatic Control*, 44(6):1118–1131, Juner 1999.

[4] B. Béjar, L. Zappella, and R. Vidal. Surgical gesture classification from video data. In *Medical Image Computing and Computer Assisted Intervention*, pages 34–41, 2012.

[5] A. Bissacco, A. Chiuso, Y. Ma, and S. Soatto. Recognition of human gaits. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 52–58, 2001.

[6] R. Brockett and P. Krishnaprasad. A scaling theory for linear systems. *IEEE Transactions on Automatic Control*, 25(2):197–207, 1980.

[7] R. Chaudhry, A. Ravichandran, G. Hager, and R. Vidal. Histograms of oriented optical flow and binet-cauchy kernels on nonlinear dynamical systems for the recognition of human actions. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

[8] J. M. C. Clark. The consistent selection of parameterizations in system identification. In *Proceedings of Joint Automatic Control Conference*, pages 576–580, 1976.

[9] D. F. Delchamps. *The geometry of spaces of linear systems with an application to the identification problem*. PhD thesis, Harvard University, 1982.

[10] G. Doretto, A. Chiuso, Y. Wu, and S. Soatto. Dynamic textures. *International Journal of Computer Vision*, 51(2):91–109, 2003.

[11] A. El-Sakkary. The gap metric: Robustness of stabilization of feedback systems. *IEEE Transactions on Automatic Control*, 30(3):240–247, 1985.

[12] K. Glover. All optimal hankel-norm approximations of linear multivariable systems and their L^∞ error bounds. *International Journal of Control*, 39(6):1115–1193, 1984.

[13] K. Glover and J. C. Willems. Parametrizations of linear dynamical systems: Canonical forms and identifiability. *IEEE Transactions on Automatic Control*, 19(6):640–645, November 1974.

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

[15] B. Hanzon and S. I. Marcus. Riemannian metrics on spaces of stable linear systems, with applications to identification. In *IEEE Conference on Decision & Control*, pages 1119–1124, 1982.

[16] M. Hazewinkel. Moduli and canonical forms for linear dynamical systems II: the topological case. *Mathematical Systems Theory*, 10:363–385, 1977.

[17] U. Helmke. Balanced realizations for linear systems: a variational approach. *SIAM Journal on Control and Optimization*, 31(1):1–15, January 1993.

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

[19] N. D. Jimenez, B. Afsari, and R. Vidal. Fast Jacobi-type algorithm for computing distances between linear dynamical systems. In *European Control Conference*, pages 3682 – 3687, 2013.

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

[21] P. S. Krishnaprasad. *Geometry of Minimal Systems and the Identification Problem*. PhD thesis, Harvard University, 1977.

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

[23] B. C. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions On Automatic Control*, 26:17–32, 1981.

[24] L. Pernebo and L. Silverman. Model reduction via balanced state space representations. *IEEE Transactions on Automatic Control*, 27(2):382–387, 1982.

[25] D. J. Stilwell and W. J. Rugh. Stability preserving interpolation methods for the synthesis of gain scheduled controllers. *Automatica*, 36(5):665–671, 2000.

[26] L. Younes. *Shapes and Diffeomorphisms*, volume 171 of *Applied Mathematical Sciences*. Springer, 2010.