The SB-ST decomposition in the study of Developmental Coordination Disorder

Leonardo Claudino  
Department of Computer Science  
U. of MD, College Park, MD, USA  
Email: claudino@cs.umd.edu

Jane E. Clark  
Department of Kinesiology  
U. of MD, College Park, MD, USA  
Email: jeclark@umd.edu

Yiannis Aloimonos  
Department of Computer Science  
U. of MD, College Park, MD, USA  
Email: yiannis@cs.umd.edu

Abstract—To deal with redundancy and high-dimensionality that are typical of movement data, we propose to decompose action matrices in two decoupled steps: first, we discover a set of key postures, that is, vectors corresponding to key relationships among degrees of freedom (like angles between body parts) which we call spatial basis (SB) and second, we impose a parametric model to the spatio-temporal (ST) profiles of each SB vector. These two steps constitute the SB-ST decomposition of an action: SB vectors represent the key postures, their ST profiles represent trajectories of these postures and ST parameters express how these postures are being controlled and coordinated. SB-ST shares elements in common with computational models of motor synergies and biological motion perception, and it relates to human manifold models that are popular in machine learning. We showcase the method by applying SB vectors and ST parameters to study vertical jumps of adults, typically developing children and children with Developmental Coordination Disorder obtained with motion capture. Using that data, we also evaluate SB-ST alone and against other techniques in terms of reconstruction ability and number of dimensions used.

I. INTRODUCTION

The means to obtain movement data are getting cheaper, more diverse and achieving higher throughput. These data are high-dimensional and highly redundant, both at the level of degrees of freedom (or dofs, for example, angles between body parts) and in terms of how often spatial arrangements of these dofs (postures) are recruited in the timeline of the action. It is thus very hard to analyze raw movement data, and in practice movement analysts will discard dofs, look at a single dof at a time, or assume the existence of a single external variable (an unknown direct or indirect function of dofs) being controlled during the action. A typical approach is uncontrolled manifold analysis (UCM) [1], [2], a framework designed to investigate whether a certain performance variable is being controlled during movement by factoring the variance (or covariance) of one (or more) elemental variables (or dofs) at different instants of a task performance into two manifolds: one that is tangent to the trajectory ($V_{UCM}$) and another that is orthogonal to it ($V_{ORT}$). When most of the variance projects onto $V_{UCM}$, the performance variable is expected to change little in face of flexible configurations of the considered dofs. For example, when studying vertical jumps, one could use the center of mass trajectory as the performance variable and, within the UCM framework, verify how stable or relatively invariant that variable is when typically (TD) or atypically (ATD) developing children perform, given a number of jumps obtained from both populations.

Despite being a great tool to study the stability of performance variables, UCM was not designed to identify ensembles of dofs and/or parametrize its relative timings, plus it will often rely on multiple trials to calculate manifolds. With that in mind we propose an alternative representation obtained by decomposing a single trial action matrix $Y_{T \times J}$ ($T$ = time instants, $J$ = dofs) in two decoupled steps: first, we discover a set of vectors spanning the $J$ space of $Y$, which we call spatial basis (SB) because they are supposed to represent key relationships between dofs, or key postures. Second, we impose a parametric model to the spatio-temporal (ST) profiles of each SB vector. Spatio-temporal profiles of SB vectors are 1-D signals expressing their temporal correlation with $Y$; a high correlation of a vector at time $t$ indicates strong recruitment or activation of the vector at that time. These two steps constitute the SB-ST decomposition of an action: SB vectors represent the key postures, their ST profiles represent trajectories of these postures, and ST parameters express how much (control) and when (coordination) these postures are being recruited in each case. Going back to the jump example, we can now use SB-ST to compare jumps of TD and ATD children simultaneously in terms of dof recruitment, trajectories, control and coordination.

Dimensionality reduction of movement data has been studied in the context of different disciplines: for example, in motor neuroscience, the time-varying muscle synergy (TVMS) model was originally designed to study laboratory data from frog jumps [3] and walking data from humans [4]. In psychology of vision, the locomotory model of Troje [5] was used to characterize point-light displays\(^1\) of walkers, and for the synthesis of new walking displays. In machine learning, human manifold models like the GPLVM family [6], [7] were shown to perform very well in tasks such as tracking and pose estimation. Like SB-ST, the first two approaches produce representations that decouple space and time. The latter reduces dimensionality in both spaces simultaneously.

Our contributions are threefold: (1) we present a very unique application of dimensionality reduction: the analysis

\(^1\)A typical point-light display is a video with an actor dressed up with dark clothes and white spherical markers in a way that only the markers are visible. The result is a moving point cloud.
II. THE SB-ST ACTION DECOMPOSITION

SB-ST is computed in 2 major steps: (1) given an input action matrix, we first we extract spatial basis (SB) vectors and compute their spatio-temporal (ST) profiles, and (2) we fit a parametric model to the ST profiles of each vector, as follows. See Fig. 1 for an overview of the method.

A. Spatial basis SB and spatio-temporal profiles

Let $Y_{T \times J}$ be a multi-dimensional action signal, for example, a $T$-length sequence of $J$ degrees of freedom (dofs). The $k$-th order approximation of that signal by SVD, in matrix notation is:

$$
\hat{Y}_{T \times J} = z_1 v_1^T + z_2 v_2^T + \ldots + z_N v_N^T,
$$

where $v_i$ is one of the top $k$ right singular vectors of $Y$, therefore spanning the column space of that matrix, and projection $z_i = Y v_i$ corresponds to the spatio-temporal profile of $v_i$, that is a one-dimensional time series that expresses the correlations of the particular spatial configuration represented by $v_i$ along the timeline of the action\(^2\). For each $i$, let $\{\Phi(\tau_{i,j}, t) : j = 1 \ldots N_i\}$ be a family of $N_i$ Gaussians with fixed standard deviations and $\Phi_{\tau_i}$ to be the corresponding $T \times N_i$ matrix such that each function is evaluated at $T$ instants and it becomes a column of that matrix. We will parametrize $z_i$ by fitting a linear combination of the columns of $\Phi_{\tau_i}$ with linear parameters $c_i = \{c_{i,1}, c_{i,2} \ldots c_{i,N_i}\}$:

$$
\hat{Y}_{T \times J} = (\Phi_{\tau_1} c_1) v_1^T + (\Phi_{\tau_2} c_2) v_2^T + \ldots + (\Phi_{\tau_N} c_N) v_N^T,
$$

and we have $z_i = \Phi_{\tau_i} c_i$. Equivalently, the posture produced by the model at time $t$ is:

$$
\hat{Y}(t) = \tilde{z}_1(t) v_1^T + \tilde{z}_2(t) v_2^T + \ldots + \tilde{z}_k(t) v_k^T,
$$

where:

$$
\tilde{z}_i(t) = c_{i,1} \Phi(\tau_{i,1}, t) + c_{i,2} \Phi(\tau_{i,2}, t) + \ldots + c_{i,N_i} \Phi(\tau_{i,N_i}, t).
$$

See Fig. 2 for a schematic view. Vector $v_i$ corresponds to the $i$-th spatial basis (SB) vector of $Y$ or SB-$i$. Each $v_i$ expresses relationships between dofs (principal postures). Basis functions $\Phi(\tau_{i,j}, t)$ (and, equivalently, its matrix version $\Phi_{\tau_i}$) together with the mean vector $\tau_i$ and the linear parameter vector $c_i$ constitute what we call the $i$-th spatio-temporal representation (ST) of $Y$ or ST-$i$. These parameters map local temporal patterns and describe how a spatial vector $v_i$ is controlled and coordinated.

\(^2\)Note that, for right singular vector $v_i$, $z_i = Y v_i = \sigma_i u_i$, with $\sigma_i$ being the $i$-th singular value and $u_i$ the $i$-th left singular vector. We chose to use $Y v_i$ rather than $u_i$ just to emphasize that vector $z_i$ expresses a time series of correlations between the data matrix $Y$ and the particular $v_i$. 

Fig. 1: SB-ST action decomposition. In this example, $J$-dimensional spatial basis vector $v_i$ encodes a linear combination of joint angles $\theta_1$, $\theta_2$ and $\theta_3$ computed with SVD, as shown by the leftmost figure. The projection $z_i = Y v_i$ of action matrix $Y_{T \times J}$ onto $v_i$ results into an often smooth temporal series of correlations that represents the activity of that particular spatial arrangement (posture) along the timeline of the action (center figure). We use VARPRO to produce a compact parametric representation for this temporal behavior by fitting a mixture of $z_i = \Phi_{\tau_i} c_i$ to $z_i$ (right figure) which results in parameter vectors $\tau_i = \{\tau_{i,1}, \tau_{i,2}, \tau_{i,3}\}$ and $c_i = \{c_{i,1}, c_{i,2}, c_{i,3}\}$. An action matrix is therefore fully characterized by each spatial basis vector $v_i$ and corresponding set of spatio-temporal parameter vectors $\tau_i$ and $c_i$.

Fig. 2: Generating postures with SB-ST. A posture $Y(t)$ results of a linear combination of spatial basis vectors $v_1, v_2 \ldots v_k$ (dashed lines), as in Equation 3. Coefficients $\tilde{z}_i(t)$ of each vector $v_i$ are the product of the $t$-th time row of its spatio-temporal matrix $\Phi_{\tau_i}$ and respective linear parameter vector $c_i$ (solid lines).

of motion capture data of vertical jumps performed by adults, TD children and children with Developmental Coordination Disorder (DCD); there is an increasing demand for this kind of study, in response to recent scientific findings correlating movement abnormalities in childhood and the later development of neuro-developmental disorders [8]. Using the jump data we: (2) introduce a framework to study actions and actors based on SB vectors and ST parameters and present evidence that the major differences between TD and DCD jumps are more likely to reside in the spatio-temporal facet of the behavior, plus (3) evaluate and compare SB-ST with alternative techniques. For example, as opposed to SB-ST, TVMS does not work well on individual trials, and both Troje’s method and GPLVM miss local temporal features that are crucial to the study of behavior [4].

This paper is organized as follows: we begin by introducing SB-ST (Sec. II) followed by previous work (Sec. III), to facilitate comparing the proposed method with alternative techniques by having presented its structure first. Next, we discuss our experiments and conclusions.
Algorithm 1: SB_ST(Y, k, \([N_1...k]\))

Compute \([U, \Sigma, V]\) = SVD of \(Y\)
for \(i = 1\) to the first \(k\) columns \(v_i\) of \(V\) do
Form \(z_i = Y v_i\)
Form approximation \(\tilde{z}_i\) by:
1. running NLLS solver that calls \([\tau_i, \tilde{c}_i, r, J]\) = VARPRO_loop(\(\tau_i, z_i, [N_1...k]\)) w/random initial \(\tau_i\) (solver minim. \(r^2\) using Jacob. \(J\)),
2. recalculating \(\Phi_{\tau_i}\) from optimal \(\tau_i\) and fixed stds,
3. making \(\tilde{z}_i = \Phi_{\tau_i} \tilde{c}_i\) using optimal \(\tilde{c}_i\),
Update approximation \(Y_{\tilde{z}_i} = Y_{\tilde{z}_i} + \tilde{z}_i v_i^T\)
end for
Return \(v_i, \tilde{c}_i, \tau_i, (i = 1 \ldots k)\) and \(Y\)

Algorithm 2: VARPRO_loop(\(\tau_i, z_i, [N_1...k]\))

Compute matrix \(\Phi_{\tau_j}\) from \(\tau_j\) and fixed stds
Compare \([U, \Sigma, V]\) = SVD of \(\Phi_{\tau_i}\)
Make \(\tilde{\epsilon}_i = V^T \Sigma^{-T} U^T z_i\)
Compute current \(\tilde{z}_i = \Phi_{\tau_i} \tilde{\epsilon}_i\) and residual \(r = z_i - \tilde{z}_i\)
for \(j = 1\) to \(N_i\) Gaussians of \(\Phi_{\tau_j}\) do
Form matrix with partial derivatives \(D_j = \frac{\partial(\Phi_{\tau_j})}{\partial \tau_j}\)
Make \(a_j = D_j \tilde{c}_i - U(U^T (D_j \tilde{c}_i))\) and \(b_j = U(\Sigma^{-1} (V^T (D_j^T r)))\)
Add \(a_j\) and \(b_j\) and form the \(j\)-th column of \(J\) as in Eqs. 9 to 11
end for
Return \(\tau_i, \tilde{c}_i, r, J\)

B. Spatio-temporal representations ST

Because we made \(\Phi(\tau_{j,i}, t)\) a family of single-parameter Gaussians, this problem turns out to be a separable least-squares regression problem, which allows us to solve for \(\tau_i\) and \(c_i\) using variable projection (VARPRO) [9]. The method exploits the linear substructure of this particular case of nonlinear least squares (NLLS) regression: if you fix the set of non-linear parameters \(\tau_i\), the problem turns out to be linear in \(c_i\) and can be solved for the latter using linear least squares (LLS). In other words, parameter \(c_i\) becomes a function of parameters \(\tau_i\) and so, instead of solving:

\[
\min_{\tau_i, c_i} \|z_i - \tilde{z}_i(\tau_i, c_i)\|_2, \tag{5}
\]

we solve a less parametrized problem:

\[
\min_{\tau_i} \|z_i - \tilde{z}_i(c_i(\tau_i))\|_2. \tag{6}
\]

In the LLS stage, the pseudo-inverse solution for \(c_i\) is:

\[
\tilde{c}_i = [\Phi_{\tau_i}]^T z_i, \tag{7}
\]

where \(\tilde{z}_i\) is VARPRO’s approximation to \(z_i = Y v_i\). The solution can be expressed in terms of the SVD of \(\Phi_{\tau_i}\):

\[
\tilde{c}_i = V \Sigma^{-1} U^T z_i. \tag{8}
\]

The LLS solution is then directly embedded in the calculation of the Jacobian of \(\tilde{z}_i(c_i(\tau_i))\) for the NLLS part of the optimization. The Jacobian can be expressed as a sum of two matrices [10]:

\[
J = -(A + B), \tag{9}
\]

where each of their \(N_i\) columns are:

\[
a_j = D_j \tilde{c}_i - U(U^T (D_j \tilde{c}_i)), \tag{10}
\]

\[
b_j = U(\Sigma^{-1} (V^T (D_j^T r))), \tag{11}
\]

where \(D_j\) has zeros at all columns but \(j\), which will have the partial derivatives of the \(j\)-th Gaussian \(\Phi(\tau_{j,i}, t)\) (or the \(j\)-th column of matrix \(\Phi_{\tau_j}\) with respect to \(\tau_{j,i}\), evaluated at all \(\tau_i, \Sigma^{-1} 1\) and \(V\) are the SVD factors of \(\Phi_{\tau_j}\) (Eq. 8), and \(r\) is the residual \(z_i - \tilde{z}_i\). Operations were grouped so that only matrix-vector products are required. The full SB-ST decomposition is summarized in Algorithms 1 and 2.

III. RELATED WORK

A. Motor synergies

In the field of motor neuroscience, many agree that the central nervous system (CNS) organizes behavior by solving a dimensionality reduction problem known as Bernstein’s degrees of freedom (dofs) problem [11] or how to manage multiple dofs in space and time. One hypothesis is that the CNS controls dofs synergistically as opposed to individually, and that a small number of such motor synergies is sufficient [12], [13], [14], [3], [15]. There are various theories around the nature of motor synergies; SB-ST has more aspects in common with computational models involving matrix factorizations, in particular the time-varying muscle synergies model (TVMS). Like SB-ST, TVMS also approximates the temporal evolution of a multi-dimensional action vector with \(k\) components, which according to our notation would be:

\[
\bar{Y}(t) = z_1 v_1 (t - \tau_1)^T + z_2 v_2 (t - \tau_2)^T + \ldots + z_k v_k (t - \tau_k)^T, \tag{12}
\]

where the synergy vectors \(v_i(t - \tau_i)\) are columns of synergy matrices like the \(V_i\)’s of Fig.3a. These matrices correspond to short-length sequences of postures that are time-shifted by \(\tau_i\) and scaled by a fixed value \(z_i\) (Fig.3b). In contrast, SB vectors \(v_i\) correspond to individual postures with time-varying scaling magnitudes \(z_i(t)\) (compare with Eq. 3).

Both methods have in common the use of explicit local parametrization for spatio-temporal profiles; the importance of this choice can be illustrated by the studies of Ivanenko et al. who use EMG data of human locomotion to look for compositional differences between walking alone and walking combined with voluntary behaviors, such as kicking a ball or overcoming an obstacle [16], [4]. Their results showed that all behaviors agreed upon the same five first profiles – which happened to be very similar to walking – but not upon the sixth, whose synergy activation times varied across behaviors. They proposed an additive model tailored to their observations, where each profile was parametrized by a single Gaussian with standard deviation fixed at 6% of the walking cycle duration. SB-ST, on the other hand, represents these profiles with mixtures and thus allows for more than one activation in the timeline of the action.

B. Biological motion perception

The perception of movement is also believed to be founded on compact representations. In the pioneer experiment of Gunnar Johansson [17] point-light displays (pld) of moving actors were presented to completely naïve observers who all reported seeing a walking human, despite the lack of form
information in the visual stimuli. He then proposed vector analysis (VA) as a model to explain the phenomenon, in which a body part is modeled as a pendulum fixed at the body part it attaches to, and the whole stimuli results in a hierarchy of moving pendulums perceived as a single gestalt unit. This study is considered to have started the biological motion perception research framework, and the same pld setting has been used to study more complex classes of activities [18]. Of particular interest, Troje [5] proposed a computational method to create and manipulate synthetic plds of walking data. His eigenpostures, or the 4 principal components of a single-walker data matrix, are equivalent to the SB described in the previous section. He modeled the temporal occurrences of the eigenpostures with a family of sine functions. His sine functions are thus a special case of our spatio-temporal representation, because it will only pick up patterns that are global to the whole timeline of the action, and will miss local events that can reveal control and coordination differences across populations.

C. Human manifold models

SB-ST parametrizes trajectories projected on a low-dimensional space, so it also relates to human manifold models. Especially, Gaussian process latent variable models (GPLVM) are a family of models that map low-dimensional latent points $X_{T \times J}$ to observed data $Y_{T \times J}$ by maximizing the likelihood of $Y$ given $X$ [19], where $J$ is the number of latent dimensions. GPLVM extends principal component analysis (PCA) and probabilistic PCA in it allows for non-linear mappings by kernelizing the process covariance function. Faster extensions to GPVLM were shown to improve sparsification in the latent space [6] and to model time-dependency in $X$, like Gaussian Process Dynamic Models (GPDM) [7]. Conceptually, the columns of $X$ produced by GPLVM and the like are analogous to SB-ST’s spatio-temporal profiles $z_i = Y v_i$ obtained with SVD (note the similarity of $x_1$ and $z_1$ in Fig. 5(top-I)). Regarding GPDM, although it models dynamics, it still produces the same one-to-one $X \rightarrow Y$ kind of mapping as GPLVM, because the model marginalizes out the basis functions $f(\cdot)$ that relate one latent posture to its preceding ones and $g(\cdot)$ that models how latent variables relate to observed postures (Eqs. 1, 2 in [19]). In the end, the method creates a representation that merges space and time within the same manifold, and although this unifying approach has proven adequate for various human movement tasks [20], [21], [22], explicit local parametrization of dynamics – in contrast, present in SB-ST – is key to uncover aspects of control and coordination that are not otherwise accessible (see Motor synergies, Sec. III).

IV. EXPERIMENTS AND RESULTS

The first goal of our experiments was to examine data reconstruction performance of SB-ST alone and in comparison with (1) methods that, as SB-ST, decouple space and time (Troje and TVMS) and (2) a method that does not (GPLVM). Our second goal was to illustrate how SB-ST can used to provide insights to both actions and actors involved.

Although any kind of action could have been chosen, we looked at vertical jumps, a non-trivial behavior that requires strength, coordination and balance. Our 39 participants were first setup with 34 infrared markers and next told to jump as high as possible and try to reach for a visual target, while being recorded by an Optitrack (NaturalPoint Inc.) motion capture system with ten V100 and V100:R2 Flex cameras. The Arena software (included) was used to export its proprietary data format to BVH (Biovision Hierarchy). BVH data were later processed by code\(^3\) written in MATLAB\(^5\) (versions R2010b and R2011a).

We were able to collect a total of 358 jumps: 9 typically developing female children (TD-F, 98) 6 adult females (AD-F, 61) 10 TD male children (TD-M, 88) 5 adult males (AD-M, 52) and 7 children diagnosed with Developmental Coordination Disorder (DCD, 59) [23]. DCD data were collapsed across gender to make the sample bigger. Children were in the broad age range of 5.1 to 14 years old. Adults (AD) were in their early 20’s.TD and DCD groups were both assessed with the MABC (Movement Assessment Battery for Children) test [24], with scores $<5^{th}$ percentile and $>29^{th}$ percentile, respectively.

All jump trials were decomposed into a spatial basis of 3 vectors SB-1, SB-2 and SB-3. Regarding ST basis functions, standard deviations were fixed to $\sigma_i = \{1/(2 \cdot 1), 1/(2 \cdot 2) \ldots 1/(2 \cdot N)\} \times T$, with $T \approx 80$ rows (about .8 seconds) and $J = 6$ columns: left and right hips, knees and ankles. We only used the flexion/extension intersegmental joint angles. Each individual trial was manually segmented by an expert in the vertical jump movement, so that they span the same postural range: all poses captured within the initial and final peak knee flexions. Prior to parameter estimation, each $z_i$ was normalized into a unit vector. When using VARPLO, $\tau_i$ was constrained to $[0, 1]$, and no constraints were applied to $c_i$.

Overall, SB-ST achieved an average reconstruction accuracy of $R^2 \geq 0.95$ for all $N$ tried, where $R^2$ is the coefficient of determination (Fig. 4 (top)). We will next discuss how SB-ST performed against competing models.

A. Reconstruction: comparing with TVMS

From TVMS results, it appears that a single synergy matrix $V_2$, with hips-ankles as agonists and knees as antagonists would explain most of the jump trial ($V_1$ and $V_3$ are mostly all zeros in Fig.3). However, we were often unable to get satisfactory reconstruction of our data using TVMS, as shown in Fig. 3c: a significant part of the signal is not covered by the

---

\(^3\)Although we have used SVD to create our SB vectors, other factorizations that are not PCA-like could have been used, i. e. eigenpostures are just one possible set of SB vectors.

\(^4\)Written informed consent was obtained from all subjects/parents/legal representatives.

\(^5\)Most of the code we used to parse the BVH files is part of Prof. Neil Lawrence’s motion capture toolbox, which can be downloaded for free by registering at the author’s website. The toolbox is currently hosted at: http://staffwww.dcs.shef.ac.uk/people/N.Lawrence/mocap/.
resulting synergies, resulting in a very low $R^2$. To rule out the cause of the problem to be the poor selections of $N$, $Q$, we ran TVMS on the whole data using different combinations of these quantities, but the low $R^2$ still persisted, as illustrated by the statistics of $R^2$ shown in Fig. 3d. As a result, we discontinued the analysis based on that method. We then conjecture that the reconstruction problems of TVMS on our data should result from not using more than a single trial to compute synergies and other parameters. TVMS was designed under the assumption that there exists latent repertoires of synergies/coordinate parameters that span both multiple behaviors [25] and others that are behavior-dependent [3] and thus constrained their optimization to obtain factors that are faithful to these assumptions; synergies and parameters are supposed to be obtained from minimizing reconstruction errors across several trials. Because SB-ST operates on a per-trial basis, to be able to properly compare the two methods, we had to run TVMS on a single-trial basis.

B. Reconstruction: comparing with Troje-inspired

Fig. 4 (top) shows the performances of SB-ST against a Troje-inspired decomposition. To clarify: Troje [5] fits the time series of his eigenpostures with a single fundamental harmonic, which he finds sufficient to model locomotion. A natural extension to non-periodic actions like jumps is to select as many harmonics as needed to obtain good approximations. This is what we call a Troje-inspired decomposition. For a certain $N$, the decomposition consisted in selecting the top-$N$ responding Fourier harmonics via FFT of $z_i$ and using only these harmonics to reconstruct the original $z_i$ via IFFT. Note that, for a certain $N$, the number of parameters needed to reconstruct $z_i$ is the same in both cases making these methods comparable: SB-ST fits a mixture of $N$ Gaussians of fixed scales, therefore resulting in $N$ pairs $\tau_i, c_i$ (ST-i parameters) while Troje uses $N$ pairs of Fourier harmonics along with respective responses. Our results show that SB-ST outperformed Troje-inspired approximation of $z_i$ when $4 \leq N \leq 7$, which could be considered the range with the best trade-off between number of parameters and reconstruction error (note the change of slope in both methods when $N$ moves from 3 to 4, as well as the dramatic decrease in $R^2$ variances). Fig. 5(top-I) also shows superior qualitative performance: for the same $N = 8$, SB-ST fits the local details of $z_i$ better than its competitor.

C. Reconstruction: comparing with GPLVM

To evaluate a GPLVM computed for an action matrix $Y$, we used two steps. (1) With the resulting set of latent vectors $\hat{X}$ (see Sec. IV-C) we pseudo-inverted the $5^{th}$ Equation of [19]
to get the approximation $\hat{Y}$:

$$\hat{Y} = \hat{J} \cdot X (Y^\top K_x^{-1} X)^\top,$$

(13)

where, $K_x$ is the kernelized covariance matrix, and $\hat{J}$ is the number of columns of $X$ used in the approximation. (2) We computed $R^2$ from $\hat{Y}$ and $Y$.

Fig. 4 (bottom) shows statistics of $R^2$ on the full jump data: from left to right, more parameters are being used to compute $\hat{Y}$, that is, the larger $J$ the more columns of $X$ are being used to compute $\hat{Y}$. Note that the best result $R^2 = 0.92$ is still lower than any of the SB-ST scores in Fig. 4 (top). Moreover, we note that a GPLVM setup with $J = 1$ will result in $\hat{J} \cdot \hat{T} + 3 = 83$ parameters ($\hat{T} = 80$ is approximately the average length of $X$ obtained from our jumps) while an SB-ST configuration with $N = 4$ scoring $R^2 > 0.95$ has exactly $k(J + 2N) = 42$ parameters ($k$ is the number of SB vectors and $2N$ is the number of pairs of ST parameters): with half as many parameters, SB-ST performs better than GPLVM, which is also visible from the qualitative example of Fig. 4 (bottom) where a SB-ST configuration with $k = 3$ and $N = 10$ (78 parameters) fits the local details of the joint signals better than its competing one-latent vector GPLVM (154 parameters).

But more interesting than the $R^2$ differences between the two methods is that the best GPLVM configuration (other than the full-dimensional, $\hat{J} = J = 6$) is the one with a single latent vector ($\hat{J} = 1$) and that increasing $\hat{J}$ from 2 to 5 (except for $\hat{J} = 3$) makes $R^2$ decrease, which we found somewhat counterintuitive. A possible explanation would be that one major GPLVM latent variable is enough to represent the major features of the jumps as seen with synergy matrix $V_2$ of TVMS and as will be seen next with SB-1.

D. Data exploration: looking at jumps and jumpers based on the SB-ST parameters

In our second experiment, we used SB-ST to explore our jump data. As in Fig. 6, SB-1 coefficient statistics demonstrate that over 50% of the explained variances in the vertical jump come from 2 main groups of rotations: hips-ankles and knees. SB-1 thus works by clustering leg joints into groups of agonist and antagonist motions, and these distributions seem to generalize across all populations, given the tight clusters. Fig. 6 also reveals that both SB-2 and SB-3 coefficients are almost zero-centered and have high variances, meaning they provide no clear interpretation of the action, so the remaining of the analysis focus on spatio-temporal aspects of SB-1 alone, that is, the statistics of ST-1’s $\tau_1 = \{\tau_{1,1}, \ldots, \tau_{1,N}\}$ and $c_1 = \{c_{1,1}, \ldots, c_{1,N}\}$. We may also call $\tau_1$ and $c_1$ coordination and control parameters respectively, because the former “places” each of the Gaussians along the timeline, so they match the local features of the spatio-temporal profile of SB-1, while the latter scales these Gaussians in accordance to the intensities of SB-1 activation. We ran VARPRO with two settings ($N = 4, 5$) just to illustrate how the choice of $N$ can affect parameter distributions. After smoothing all distributions, we looked at how jumpers at different developmental stages agreed on ST-1. Data were collapsed across gender to increase the number of subjects per population of interest.

To be considered to agree, two distributions must have similar shape and/or about the same peak abscissa. We judged that, for the present purposes, visual inspection was enough to assess agreement. As seen from Fig. 7, frequent agreements between adults (AD) and typically developing children (TD) plus partial and full disagreements with DCD children suggest that the 3 populations may be controlling and coordinating SB-1 distinctly: for example, when $N = 5$, all populations seem to be recruiting the fourth Gaussian early in the timeline, since all $\tau_{1,4}$ peak at about the same time but $c_{1,4}$ do not (orange selection, Fig. 7) so we can hypothesize that (1) there may be inter-population discrepancies related to spatial configuration SB-1 taking place somewhat early in the jump and that (2) the more you move to the right on $c_{1,4}$, the less mature the jump is, since the sequence of peaks is AD→TD→DCD. The movement analyst could then manipulate $c_{1,4}$, reconstruct the jumps and inspect the effects near $\tau_{1,4}$.
V. CONCLUSIONS AND FUTURE WORK

This paper describes the SB-ST decomposition and how it factors action matrices. Conceptually speaking, SB-ST can be seen as a synergy model of single postures with time-varying scaling magnitudes, and it generalizes spatio-temporal profiles proposed to explain locomotory data in motor neuroscience [4] and psychology of vision [5]. Local parametrization of spatio-temporal profiles, although proven critical in the study of actions [16], [4], is not present in human manifold models like GPLVM and GPDM, but it is a feature of SB-ST.

Comparative reconstruction of vertical jumps suggested that: (1) SB-ST can be more adequate than TVMS to factor single-trials, (2) SB-ST can outperform Troje-inspired at the best the trade-off between number of basis functions and $R^2$, (3) it do as well or better than GPLVM with half the representation size. In a second experiment, we showed that SB-ST can be a good tool to study actions and actors, and results revealed that (1) despite conceptual differences, TVMS, GPLVM and SB-ST all agreed that jumps are mostly loading on a single factor. (2) SB-1 coefficients were consistent among all populations, suggesting jumpers are recruiting the same major synergy regardless of jump maturity (age, presence of disorder) or gender. By inspecting ST-1 statistics, we saw that (3) one of the Gaussians is consistently coordinated by all populations to be at the beginning of the trial, but it is controlled differently. We note that to discern what exactly these differences mean as well as their significance would require a more thorough analysis and rigorous statistical testing, which surpasses the scope of this work.

In ST parameter estimation, we use a family of $N$ Gaussians with fixed standard deviations (stds) to facilitate the comparison among populations, because we could establish correspondences between Gaussians based on corresponding stds (as we did in Sec. IV-D when we fixed the fourth Gaussian and looked at differences in $c_{1,4}$ and $\tau_{1,4}$). Therefore, reconstruction results could improve further if we also optimized for stds; a future development would be to add std optimization, discretize these stds into bins and correspond Gaussians based on the bins. Another interesting future experiment would be to compare the performance of our VARPRO-based ST representation with an ST learned with the dynamic primitives proposed by Ijspeert et al. [26].

ACKNOWLEDGMENTS

Work described in this paper was supported by NIH grant R21DA024323. We thank Dr. Melissa Pangelinan for helping with scheduling and testing the subjects, and for customizing the IR marker setup to suit all participants. We thank Dr. Brad King for his contributions during the data discussion meetings.

REFERENCES