ALIGNING SETS OF TEMPORAL SIGNALS WITH RIEMANNIAN GEOMETRY AND KOOPMAN OPERATOR

Ohad Rahamim and Ronen Talmon

Viterbi Faculty of Electrical Engineering
Technion – Israel Institute of Technology
Haifa 3200003, Israel

ABSTRACT

In this paper, we consider the problem of aligning data sets of short temporal signals without any a-priori known correspondence. We present a method combining Koopman operator theory and the Riemannian geometry of symmetric positive-definite (SPD) matrices. First, by taking a Koopman operator theory standpoint, we build feature matrices of the signals using dynamic mode decomposition (DMD). Second, we align these features using parallel transport of SPD matrices, built from the DMD feature matrices. We showcase the performance of the proposed method on simulated observations of a mechanical system and on two real-world applications: sleep stage identification and pre-epileptic seizure prediction.

Index Terms— Koopman operator, dynamic mode decomposition, domain adaptation, Riemannian geometry, parallel transport

1. INTRODUCTION

Aligning sets of measurements is a long standing problem in applied science, which has emerged over the years in many forms in a broad range of fields. Notable examples include domain adaptation in data analysis [1, 2], transfer learning in machine learning [3], and batch effect removal in bioinformatics [4], to name but a few. Here, we focus on aligning data sets of short temporal signals.

Broadly, when analyzing signals in the time domain, it is often useful to view the signals as observations of a dynamical system, assuming that the analysis of the dynamical system provides important information on the signals. To this end, one could take advantage of a recent line of methods that attempt to learn the dynamical system from observations in a model-free manner, e.g. [5, 6, 7]. Among these methods, perhaps the most prominent approach is based on Koopman operator theory (see [8] and references therein).

Koopman operator theory has assumed an increasing role in the study of dynamical systems in recent years, and particularly, it has important utility in data-driven dynamical systems analysis from observations. The Koopman operator is an infinite-dimensional *linear* operator that represents a possibly *non-linear* dynamical system by describing the linear evolution of functions defined on the state space. Over the years, several methods to construct a finite-dimensional operator (matrix), which approximates the Koopman operator from system observations, have been proposed. Perhaps the most commonly-used is the dynamic mode decomposition (DMD) [9, 10], which has opened the door to a line of extensions, e.g. [11, 12].

In this paper, we present a two-step method for aligning data sets of short temporal signals. At the first step, we compute a DMD matrix approximation of the Koopman operator of each short temporal signal in the data sets at hand. These matrices are viewed as features that capture the dynamical information of the signals. At the second stage, we define a symmetric positive-definite (SPD) matrix based on each DMD matrix. The Riemannian geometry of SPD matrices has been studied for many years (see [13] and references therein). Presumably, the fact that many basic operations, such as the logarithmic and exponential maps and the Riemannian distance, have explicit closed-form expressions has made SPD matrices convenient and useful features. Indeed, SPD matrices have been commonly-used, and their Riemannian geometry has been incorporated in computational methods for various tasks, often achieving state of the art performance [14, 15, 16, 17]. Here, we propose to use parallel transport of SPD matrices [18], which has an explicit expression as well and can be efficiently implemented [16], in order to align sets of DMD matrices. Specifically, in order to align two sets, we propose to parallel transport the SPD features of the signals of one set from their Riemannian mean to the Riemannian mean of the other set along the unique geodesic path. Seemingly, parallel transport is nothing more than the Riemannian counterpart of mean subtraction. Yet, we show that by applying such a parallel transport to the SPD features, we attain a useful alignment of the respective DMD matrices in terms of the signal dynamics. In addition, we describe a natural extension for aligning more than two sets of signals.

We demonstrate the properties of the proposed method on observations of a simulated mechanical system. In addition, we present two applications: sleep stage identification and pre-epileptic seizure detection, where we show accurate alignment of signals recorded from different subjects in a completely unsupervised, data-driven, and model-free fashion.

2. PROBLEM FORMULATION

Consider a dynamical system controlled by parameters, which is described by the following differential equation

$$\dot{\mathbf{x}} = F\left(\mathbf{x}; \alpha\right) \tag{1}$$

where $\mathbf{x} \in \mathcal{M}^{\alpha}$ is the system state, $\mathcal{M}^{\alpha} \subseteq \mathbb{R}^d$ is the state space embedded in a d-dimensional Euclidean space, and $\alpha = \{\alpha_1, \ldots, \alpha_N\}$ is a set of parameters. Suppose we have two sets of observations of the dynamical system. Let $\{\mathbf{X}_i^{(1)}\}_{i=1}^{N_1}$ denote the set of N_1 short temporal trajectories of observations, where the system is configured with the set of parameters α_1. Each trajectory $\mathbf{X}_i^{(1)} \in \mathbb{R}^{d \times M}$ consists of M time samples of the state $\mathbf{x} \in \mathbb{R}^d$ of the respective system at a discrete uniform sampling grid, where the system is initialized at $\mathbf{x}_{i,0}^{(1)}$. Similarly, let $\{\mathbf{X}_j^{(2)} \in \mathbb{R}^{d \times M}\}_{j=1}^{N_2}$ denote the set of N_2 short temporal trajectories of observations, where the system is configured

Algorithm 1: Proposed Alignment

Input: Sets of temporal signals $\{\mathbf{X}_i^{(1)}\}_{i=1}^{N_1}$, $\{\mathbf{X}_j^{(2)}\}_{j=1}^{N_2}$

Output: Sets of aligned matrices $\{\widetilde{\mathbf{P}}_i^{(1)}\}_{i=1}^{N_1}$, $\{\mathbf{P}_j^{(2)}\}_{j=1}^{N_2}$

for $i = 1, ..., N_1$ **do**
  Build the delay coordinates of $\mathbf{X}_i^{(1)}$ according to (2)
  Set $\hat{\mathbf{X}}_i^{(1)}$ and $\hat{\mathbf{Y}}_i^{(1)}$ according to (3) and (4)
  Compute the DMD matrix $\mathbf{K}_i^{(1)}$ according to (5)
  Build the SPD matrix $\mathbf{P}_i^{(1)}$ according to (6)
end
Compute $\overline{\mathbf{P}}^{(1)} \leftarrow \texttt{RiemannianMean}(\{\mathbf{P}_i^{(1)}\}_{i=1}^{N_1})$
repeat Steps 1-7 for $\mathbf{X}_j^{(2)}$, $j = 1, \ldots, N_2$
Set $\mathbf{E} = \left(\overline{\mathbf{P}}^{(2)}(\overline{\mathbf{P}}^{(1)})^{-1}\right)^{1/2}$
for $i = 1, ..., N_1$ **do**
  Compute $\widetilde{\mathbf{P}}_i^{(1)} = \mathbf{E}\mathbf{P}_i^{(1)}\mathbf{E}^T$
end

with the set of parameters α_2. Our goal is to devise a scheme that aligns the two sets of observations $\{\mathbf{X}_i^{(1)}\}_{i=1}^{N_1}$ and $\{\mathbf{X}_j^{(2)}\}_{j=1}^{N_2}$ based on their respective dynamics in a purely data-driven fashion, without any a-priori known correspondence.

To make this task more concrete, we present an illustrative toy problem. Consider the mechanical system of coupled pendula with equal masses and equal lengths, and assume a linear regime pertaining to only small perturbations of the system around the equilibrium point (see [6] for details). Such a system is controlled by three parameters: the pendula length ℓ, the pendula mass m, and the spring constant k. The state of such a system is 4-dimensional, consisting of the positions and the velocities of the two pendula. Solving the ordinary differential equation describing the evolution of the pendula positions within the linear regime yields that the position of each pendulum in time is given by

$$c_1 \cos(\omega_1 t) + c_2 \cos(\omega_2 t)$$

where $\omega_1 = \sqrt{g/\ell}$, $\omega_2 = \sqrt{g/\ell + 2k/m}$, g represents gravity, and c_1 and c_2 are coefficients determined by the initial conditions. For the purpose of illustration, consider two dynamical regimes: `regime 1` where $c_1 > c_2$ and `regime 2` where $c_1 \leq c_2$. Now, let $\{\mathbf{X}_i^{(1)}\}_{i=1}^{N_1}$ be a set of trajectories of temporal state observations, where the system is configured with the set of parameters $\alpha_1 = \{\ell_1, m, k\}$, and let $\{\mathbf{X}_j^{(2)}\}_{j=1}^{N_2}$ be a set of trajectories of temporal state observations, where the system is configured with the set of parameters $\alpha_2 = \{\ell_2, m, k\}$, assuming $\ell_1 \neq \ell_2$. In the context of this problem, our goal is to align the two sets, so that trajectories from different sets but from the same dynamical regime are matched.

3. KOOPMAN OPERATOR THEORY

There is a vast literature on Koopman operator theory, and particularly, on estimating the Koopman operator from temporal observations in a model-free manner (e.g., see [8, 9, 10]). For brevity, we describe here a particular method, called the exact DMD approximation [10], which is used in our empirical study; other alternatives, e.g., [11, 12], can be implemented and used in the context of this work in a similar manner.

Let $\mathbf{X}_i \in \mathbb{R}^{d\times M}$ be a single trajectory of M consecutive temporal observations. To facilitate accurate DMD approximation, Takens' delay coordinates [19] are considered, i.e., we define

$$\hat{\mathbf{x}}_i(n) = [\mathbf{x}_i^T(n), \ldots, \mathbf{x}_i^T(n+l-1)]^T \in \mathbb{R}^{d\cdot l},\ n = 1, \ldots, M-l+1 \tag{2}$$

where $\mathbf{x}_i(n) \in \mathbb{R}^d$ is the n-th sample (column) in the trajectory $\mathbf{X}_i$, and $l > 1$ is a tunable integer parameter. The delay coordinates are then organized into two matrices as follows:

$$\hat{\mathbf{X}}_i = [\hat{\mathbf{x}}_i(1), \hat{\mathbf{x}}_i(2), \ldots, \hat{\mathbf{x}}_i(M-l)] \in \mathbb{R}^{(d\cdot l)\times(M-l)} \tag{3}$$

$$\hat{\mathbf{Y}}_i = [\hat{\mathbf{x}}_i(2), \hat{\mathbf{x}}_i(3), \ldots, \hat{\mathbf{x}}_i(M-l+1)] \in \mathbb{R}^{(d\cdot l)\times(M-l)} \tag{4}$$

The matrix approximation $\mathbf{K}_i \in \mathbb{R}^{(d\cdot l)\times(d\cdot l)}$ of the Koopman operator satisfies $\hat{\mathbf{Y}}_i = \mathbf{K}_i\hat{\mathbf{X}}_i$. Accordingly, the exact DMD approximation [10] is given by

$$\mathbf{K}_i \triangleq \hat{\mathbf{Y}}_i\hat{\mathbf{X}}_i^+ \tag{5}$$

where $\hat{\mathbf{X}}_i^+$ denotes the Moore-Penrose pseudo-inverse of $\hat{\mathbf{X}}_i$, given by $\hat{\mathbf{X}}_i^+ = \hat{\mathbf{X}}_i^T\left(\hat{\mathbf{X}}_i\hat{\mathbf{X}}_i^T\right)^{-1}$.

4. PROPOSED METHOD

The proposed method for aligning two sets of short temporal signals consists of two main steps. At the first step, given the two sets $\{\mathbf{X}_i^{(1)}\}_{i=1}^{N_1}$ and $\{\mathbf{X}_j^{(2)}\}_{j=1}^{N_2}$, for each signal $\mathbf{X}_i^{(k)}$, $k = \{1, 2\}$, we build the exact DMD approximation as described in Section 3. The resulting DMD matrix is denoted by $\mathbf{K}_i^{(k)}$, encapsulating the temporal dynamics of the signal. This procedure yields two sets of DMD matrices $\{\mathbf{K}_i^{(1)}\}_{i=1}^{N_1}$ and $\{\mathbf{K}_j^{(2)}\}_{j=1}^{N_2}$.

At the second step, we define an SPD matrix based on each DMD matrix

$$\mathbf{P}_i^{(k)} = \mathbf{K}_i^{(k)}(\mathbf{K}_i^{(k)})^T \tag{6}$$

yielding the sets $\{\mathbf{P}_i^{(1)}\}_{i=1}^{N_1}$ and $\{\mathbf{P}_j^{(2)}\}_{j=1}^{N_2}$.

Considering SPD matrices as features entails many advantages. Here, we make use of the Riemannian distance defined between any two SPD matrices $\mathbf{P}_i$ and $\mathbf{P}_j$ by

$$d_R(\mathbf{P}_i, \mathbf{P}_j) = \left\|\log\left(\mathbf{P}_j^{-\frac{1}{2}}\mathbf{P}_i\mathbf{P}_j^{-\frac{1}{2}}\right)\right\|_F^2 \tag{7}$$

where $\log(\cdot)$ is the matrix logarithm and $\|\cdot\|_F$ is the Frobenius norm. Based on the Riemannian distance, for each set $k = \{1, 2\}$, we compute the corresponding Riemannian mean $\overline{\mathbf{P}}^{(k)}$ using an iterative algorithm proposed in [14], which is based on [20]. Then, in order to align the two sets, we parallel transport (PT) the set $\{\mathbf{P}_i^{(1)}\}_{i=1}^{N_1}$ from $\overline{\mathbf{P}}^{(1)}$ to $\overline{\mathbf{P}}^{(2)}$ along the unique geodesic path connecting $\overline{\mathbf{P}}^{(1)}$ and $\overline{\mathbf{P}}^{(2)}$. We use the efficient implementation presented in [16], circumventing the logarithmic and exponential maps, where the transported matrices are explicitly given by

$$\widetilde{\mathbf{P}}_i^{(1)} = \mathbf{E}\mathbf{P}_i^{(1)}\mathbf{E}^T \tag{8}$$

where

$$\mathbf{E} = \left(\overline{\mathbf{P}}^{(2)}(\overline{\mathbf{P}}^{(1)})^{-1}\right)^{1/2} \tag{9}$$

Consequently, the two set $\{\widetilde{\mathbf{P}}_i^{(1)}\}_{i=1}^{N_1}$ and $\{\mathbf{P}_j^{(2)}\}_{j=1}^{N_2}$ are considered aligned. The alignment algorithm is summarized in Algorithm 1.

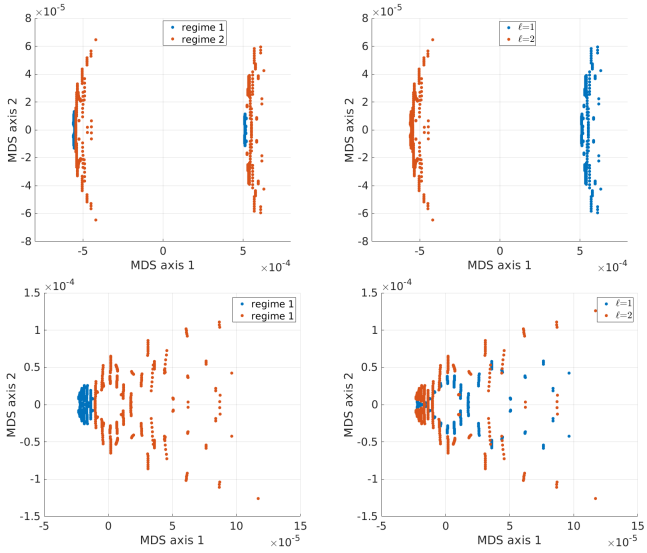

Fig. 1: MDS based on the Riemannian distances between the DMD-based features. The top row presents features before the alignment, and the bottom row presents features after the alignment. At the left column the features are colored by the dynamic regime, and at the right column by the pendula length ℓ_1 and ℓ_2.

The proposed alignment is based on the specific Riemannian geometry of SPD matrices. Yet, by combining (6) and (8), it can applied directly to the DMD matrices $\mathbf{K}_i^{(1)}$ by

$$\widetilde{\mathbf{K}}_i^{(1)} = \mathbf{E}\mathbf{K}_i^{(1)}$$

yielding the aligned sets $\{\widetilde{\mathbf{K}}_i^{(1)}\}_{i=1}^{N_1}$ and $\{\mathbf{K}_j^{(2)}\}_{j=1}^{N_2}$.

The extension of the algorithm to more than two sets is straightforward. Consider now N_s sets of trajectories of temporal observations $\{\mathbf{X}_i^{(k)}\}_{i=1}^{N_k}$, where $k = 1, ..., N_s$. Repeating the construction described earlier, let $\{\mathbf{K}_i^{(k)}\}_{i=1}^{N_k}$ denote the sets of DMD matrices and let $\{\mathbf{P}_i^{(k)}\}_{i=1}^{N_k}$ denote the corresponding sets of SPD matrices. Broadly, in order to align the N_s sets, we propose to parallel transport the matrices to the Riemannian mean of the Riemannian means.

Concretely, let $\overline{\mathbf{P}}$ denote the Riemannian mean of $\{\overline{\mathbf{P}}^{(k)}\}_{k=1}^{N_s}$, where $\overline{\mathbf{P}}^{(k)}$ is the Riemannian mean of the k-th set $\{\mathbf{P}_i^{(k)}\}_{i=1}^{N_k}$. We propose to parallel transport each set $\{\mathbf{P}_i^{(k)}\}_{i=1}^{N_k}$ along the unique geodesic path connecting $\overline{\mathbf{P}}^{(k)}$ and $\overline{\mathbf{P}}$. By (8) and (9), we now have

$$\widetilde{\mathbf{P}}_i^{(k)} = \mathbf{E}_k \mathbf{P}_i^{(k)} \mathbf{E}_k^T, \quad \mathbf{E}_k = \left(\overline{\mathbf{P}}(\overline{\mathbf{P}}^{(k)})^{-1}\right)^{1/2}$$

for $k = 1, \ldots, N_s$, yielding the following N_s aligned sets $\left\{\widetilde{\mathbf{P}}_i^{(k)}\right\}_{i=1}^{N_k}$.

5. EXPERIMENTAL RESULTS

5.1. Coupled Pendula

We revisit the coupled pendula system described in Section 2. We simulate two systems with equal mass $m = 1$ kg and spring constant $k = 750$, but with different pendula lengths: $\ell_1 = 1$ m and $\ell_2 = 2$ m. The systems are initialized with a set of $N_1 = N_2 = 441$ initial conditions, where the initial positions of the pendula are at fixed angles $(\theta_1, \theta_2) = (0.1, -0.1)$ and the initial velocities of the pendula in m/s span the uniform 2D grid $\{-1, -0.9, \ldots, 0.9, 1\}^2$. We collect the respective two sets of trajectories $\{\mathbf{X}_i^{(1)}\}_{i=1}^{N_1}$ and $\{\mathbf{X}_j^{(2)}\}_{j=1}^{N_2}$ of length $M = 1000$.

We apply Algorithm 1 to $\{\mathbf{X}_i^{(1)}\}_{i=1}^{N_1}$ and $\{\mathbf{X}_j^{(2)}\}_{j=1}^{N_2}$ using delay coordinates with $l = 2$ and present the results in Fig.1. The top row presents the multi-dimensional scaling (MDS) [21] obtained based on the Riemannian distances (7) between $\{\mathbf{P}_i^{(1)}\} \cup \{\mathbf{P}_j^{(2)}\}$, namely, before the alignment. The bottom row presents the MDS obtained based on the Riemannian distances between $\{\widetilde{\mathbf{P}}_i^{(1)}\} \cup \{\mathbf{P}_j^{(2)}\}$, namely, after the alignment. At the left column, the embedding is colored according to the regime (`regime 1` or `regime 2`), and at the right column, the embedding is colored by the system configuration (ℓ_1 or ℓ_2) .

We observe that at the top row the embedding is primarily clustered according the system configuration, where each cluster is organized according to the dynamic regime. Conversely, at the bottom row, the embedding is organized according to the regime and the system configuration is not pronounced, indicating a useful alignment between the sets with different system configurations. Importantly, we emphasize that the information about the dynamic regime was not taken into account in the alignment, and the result is obtained in a purely unsupervised fashion based on geometric considerations.

In order to further illustrate the proposed alignment, we exploit the analytic form of the Koopman operator of this system. The coupled pendula system can be recast as a linear first order differential equation in discrete time $\mathbf{x}_{n+1} = \mathbf{K}\mathbf{x}_n$, where $\mathbf{x}_n$ is a 4-dimensional state consisting of the delayed coordinates of the two pendula positions, and

$$\mathbf{K} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & \gamma & \beta \\ 0 & -1 & \beta & \gamma \end{pmatrix} \tag{10}$$

where $\beta = \frac{k}{m}(\triangle t)^2$, $\gamma = 2 - \frac{g}{\ell}(\triangle t)^2 - \beta$, and $\triangle t$ is the discrete time step. This implies that the Koopman operator of the system is finite (a matrix). In addition, (10) gives a parametric form for the estimated Koopman operators from the system observations.

Specifically, we consider

$$\mathbf{P} = \mathbf{K}\mathbf{K}^T = \begin{pmatrix} 1 & 0 & \gamma & \beta \\ 0 & 1 & \beta & \gamma \\ \gamma & \beta & 1+\gamma^2+\beta^2 & 2\gamma\beta \\ \beta & \gamma & 2\gamma\beta & 1+\gamma^2+\beta^2 \end{pmatrix} \tag{11}$$

and extract the pendulum length from the aligned matrices $\{\widetilde{\mathbf{P}}_i^{(1)}\}$ as follows

$$\hat{\ell}_i = \frac{g\,(\triangle t)^2}{2}\left(\frac{1}{2-\hat{\gamma}-\hat{\beta}} + \frac{1}{2-\hat{\gamma}'-\hat{\beta}'}\right)$$

where $\hat{\gamma} = \widetilde{\mathbf{P}}_i^{(1)}[1,3]$, $\hat{\beta} = \widetilde{\mathbf{P}}_i^{(1)}[1,4]$, $\hat{\gamma}' = \widetilde{\mathbf{P}}_i^{(1)}[2,4]$, and $\hat{\beta}' = \widetilde{\mathbf{P}}_i^{(1)}[2,3]$.

The obtained average pendulum length across all trajectories is 2.0132 with standard deviation of $6.7 \cdot 10^{-3}$. Indeed, we observe that the parametric form of the aligned matrices $\{\widetilde{\mathbf{K}}_i^{(1)}\}$ coincides with the parametric form of $\{\mathbf{K}_j^{(2)}\}$. Importantly, this alignment is obtained in a model-free manner, considering only the geometry of the DMD features.

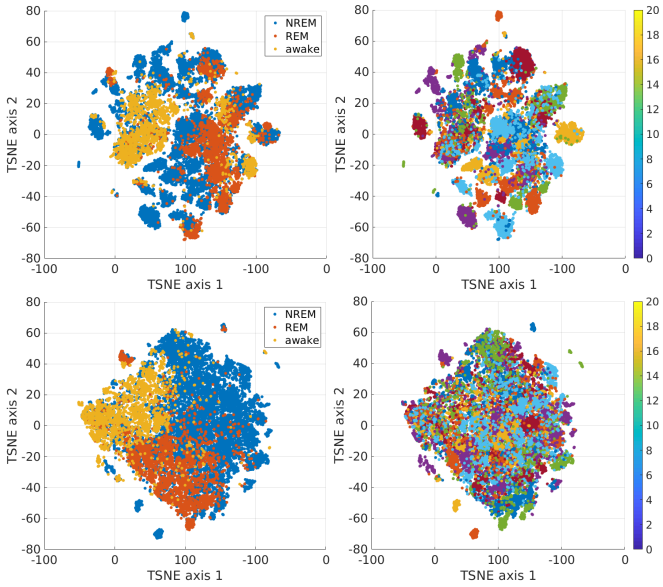

Fig. 2: tSNE plots based on the Riemannian distances between the DMD-based features. The top row presents features before the alignment, and the bottom row presents features after the alignment. At the left column, the features are colored by the sleep stage, and at the right column, the features are colored by the subject.

5.2. Sleep Stage Identification

Sleep stage identification is the task of identifying the sleep stage of a person as a function of time based on a multitude of sensor measurements acquired in special sleep clinics. The main sleep stages are awake, rapid-eye-movement (REM), and non REM (NREM) [22]. Accurate identification of the sleep stage during the sleep cycle provides important information to sleep analysis, and particularly for diagnosis of various sleep disorders, e.g. sleep apnea [23].

We use the "CAP Sleep Database" [24] from PhysioNet [25]. This dataset contains EEG recordings from 5 electrodes acquired during a night sleep from 20 subjects. The recordings are divided to segments of 30 seconds. Each segment was labeled by human experts with the corresponding sleep stage: `awake`, `REM`, and `NREM`. Our goal is to identify the sleep stage based on the EEG measurements in a *subject independent manner*. We apply the extension of Algorithm 1 described in Section 4 and present the results in Fig. 2 using tSNE [26] obtained based on the Riemannian distances between the SPD features. The top row presents features before the alignment, and the bottom row presents features after the alignment. At the left column, the features are colored by the sleep stage, which was not used in the feature extraction procedure, and at the right column, the features are colored by the subject index. We observe that the features before the alignment are clustered by the subject, whereas the features after the alignment are clustered by the sleep stage. This implies that the combination of DMD-based features and PT removes the inter-subject variability while preserving significant information on the sleep stage, in an unsupervised fashion.

To quantify this observation, we apply an SVM classifier [27] with a leave-one-subject-out cross-validation and compare the results with a similar approach using PT based on the covariance matrices of the segments [16]. The resulting classification accuracy is depicted in Table.1. Indeed, we clearly observe the contribution of the alignment. In addition, we observe that the DMD-based features facilitate better performance compared to the covariance matrices.

Table 1: Sleep stage identification accuracy.

Accuracy	Before Alignment	After Alignment
Method proposed in [16]	67.6%	72.2%
Our method	73.5%	81.2%

Table 2: Pre-epileptic seizure prediction.

Scenario	Method	Recall	Precision
Inter-subject	FT-based	64.7	97.3
	Our method	93	99.4
Intra-subject	FT-based	85	99.3
	Our method	95	99.7

5.3. Pre-epileptic Seizure Detection

Pre-epileptic seizure detection is the task of identifying the minutes before an epileptic seizure occurrence. Such a detection allows for fast treatments and life-saving actions. We use the repository from the American Epilepsy Society Seizure Prediction Challenge on Kaggle[1], which contains EEG recordings from 16 electrodes mounted on 4 dogs (subjects). We consider segments of 60 seconds, sampled at 400 Hz, and focus on two types of segments: (i) segments recorded 5 minutes prior to a seizure, labeled as `pre-seizure`, and (ii) segments recorded remotely from a seizure, labeled as as `interictal`. There are 500–1000 segments per dog, out of which approximately 10% are labeled `pre-seizure`.

Our goal is to detect the pre-epileptic segments in a *subject independent manner*. We apply the extension of Algorithm 1 described in 4 with one modification. The Riemannian means of the sets are computed based only on the `interictal` segments, because empirically we found that the `pre-seizure` segments are outliers and introduce bias. Since this procedure requires the labels, we used 25% of the `interictal` segments for this purpose and applied the alignment and subsequent detection on the remaining 75%.

To quantify the detection, we use an SVM classifier with a leave-one-subject-out cross-validation. As a baseline, we apply the SVM classifier to the recordings of each subject separately. In addition, we compare the results to one of the leading algorithms from the competition, which is based on features computed using the Fourier transform (FT). For their alignment, we use mean subtraction.

The average classification accuracy over the 4 subjects is depicted in Table. 2. We observe that in the inter-subject scenario, our method based on the DMD-based features and the geometry alignment significantly outperforms the competing method. In addition, our method is also superior in the intra-subject scenario.

6. CONCLUSIONS

In this paper, we proposed an algorithm for aligning sets of short temporal signals. The algorithm is based on the combination of Koopman operator theory and the Riemannian geometry of SPD matrices. Future work will focus on a theoretical justification of the algorithm, on handling noise, and on the incorporation of more advanced finite approximations of the Koopman operator.

[1] `http://www.kaggle.com/c/seizure-prediction`.

7. REFERENCES

[1] S. Ben-David, J. Blitzer, K. Crammer, and F. Pereira, "Analysis of representations for domain adaptation," in *Advances in neural information processing systems*, 2007, pp. 137–144.

[2] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J. Wortman Vaughan, "A theory of learning from different domains," *Machine learning*, vol. 79, no. 1-2, pp. 151–175, 2010.

[3] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[4] C. Lazar, S. Meganck, J. Taminau, D. Steenhoff, A. Coletta, C. Molter, David Y Weiss-Solís, R. Duque, H. Bersini, and Ann Nowé, "Batch effect removal methods for microarray gene expression data integration: a survey," *Briefings in bioinformatics*, vol. 14, no. 4, pp. 469–490, 2013.

[5] J. Bongard and H. Lipson, "Automated reverse engineering of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 104, no. 24, pp. 9943–9948, 2007.

[6] O. Yair, R. Talmon, R. R. Coifman, and I. G. Kevrekidis, "Reconstruction of normal forms by learning informed observation geometries from data," *Proceedings of the National Academy of Sciences*, vol. 114, no. 38, pp. E7865–E7874, 2017.

[7] T. Bertalan, F. Dietrich, I. Mezić, and I. G. Kevrekidis, "On learning hamiltonian systems from data," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 12, pp. 121107, 2019.

[8] M. Budišić, R. Mohr, and I. Mezić, "Applied koopmanism," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 22, no. 4, pp. 047510, 2012.

[9] N. J. Kutz, S. L Brunton, B. W Brunton, and J. L. Proctor, *Dynamic mode decomposition: data-driven modeling of complex systems*, SIAM, 2016.

[10] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and N. J. Kutz, "On dynamic mode decomposition: Theory and applications," *Journal of Computational Dynamics*, vol. 1, no. 2, pp. 391, 2014.

[11] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A data–driven approximation of the koopman operator: Extending dynamic mode decomposition," *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, 2015.

[12] O. Azencot, W. Yin, and A. Bertozzi, "Consistent dynamic mode decomposition," *SIAM Journal on Applied Dynamical Systems*, vol. 18, no. 3, pp. 1565–1585, 2019.

[13] X. Pennec, P. Fillard, and N. Ayache, "A riemannian framework for tensor computing," *International Journal of computer vision*, vol. 66, no. 1, pp. 41–66, 2006.

[14] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Classification of covariance matrices using a riemannian-based kernel for bci applications," *Neurocomputing*, vol. 112, pp. 172–178, 2013.

[15] P. Zanini, M. Congedo, C. Jutten, S. Said, and Y. Berthoumieu, "Transfer learning: A riemannian geometry framework with applications to brain–computer interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 5, pp. 1107–1116, 2017.

[16] O. Yair, M. Ben-Chen, and R. Talmon, "Parallel transport on the cone manifold of spd matrices for domain adaptation," *IEEE Transactions on Signal Processing*, vol. 67, no. 7, pp. 1797–1811, 2019.

[17] Or Yair, F. Dietrich, R. Talmon, and I. G. Kevrekidis, "Optimal transport on the manifold of spd matrices for domain adaptation," *arXiv preprint arXiv:1906.00616*, 2019.

[18] S. Sra and R. Hosseini, "Conic geometric optimization on the manifold of positive definite matrices," *SIAM Journal on Optimization*, vol. 25, no. 1, pp. 713–739, 2015.

[19] N. H. Packard, J. P. Crutchfield, D. J. Farmer, and R. S. Shaw, "Geometry from a time series," *Physical review letters*, vol. 45, no. 9, pp. 712, 1980.

[20] 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.

[21] M. AA. Cox and T. F. Cox, "Multidimensional scaling," in *Handbook of data visualization*, pp. 315–347. Springer, 2008.

[22] T. Lee-Chiong, *Sleep medicine: Essentials and review*, Oxford University Press, 2008.

[23] H. K. Yaggi, J. Concato, W. N. Kernan, J. H. Lichtman, L. M. Brass, and V. Mohsenin, "Obstructive sleep apnea as a risk factor for stroke and death," *New England Journal of Medicine*, vol. 353, no. 19, pp. 2034–2041, 2005.

[24] M. G. Terzano, L. Parrino, A. Sherieri, R. Chervin, S. Chokroverty, C. Guilleminault, M. Hirshkowitz, M. Mahowald, H. Moldofsky, A. Rosa, et al., "Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (cap) in human sleep.," *Sleep medicine*, vol. 2, no. 6, pp. 537–553, 2001.

[25] A. L. Goldberger, L. AN. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[26] L. van der Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[27] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.