SPDALIGN: INTERPRETABLE RIEMANNIAN ALIGNMENT FOR EEG FORWARD MODELING SHIFTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Electroencephalography (EEG) based brain-computer interfaces enable direct brain-to-device communication for applications such as rehabilitation and communication. However, their practical utility is often limited as the non-stationary nature of the EEG data introduces distribution shifts across domains (e.g., sessions and subjects). Adapting machine learning models to be invariant to these shifts in an unsupervised way, without using costly labeled calibration data, would drastically improve the utility of EEG data. In this work, we use a classic generative model of EEG to study distribution shifts introduced by the domain-specific forward process, which is associated with factors such as head geometry. We theoretically show that such distribution shifts can be recovered solely through linear transformations on the Symmetric Positive Definite manifold. Building on this insight, we propose SPDAlign, an interpretable framework for promoting domain-invariant EEG learning. SPDAlign first aligns the domain-specific means and corrects global rotations across domains using a recent optimal transport technique called Wasserstein Procrustes. We systematically study the proposed approach through simulations and demonstrate its competitive performance on extensive public EEG datasets. Additionally, SPDAlign is a globally linear framework and is intrinsically interpretable, so that the framework can identify frequency ranges of interest, determine the spatial patterns reflecting source–sensor relationships, and address cross-subject variability.

1 INTRODUCTION

Electroencephalography (EEG) records multi-channel brain activity from the scalp, providing non-invasive access to neuronal dynamics with millisecond resolution (Niedermeyer & da Silva, 2005). EEG-based brain–computer interfaces (BCIs) enable direct brain-to-device communication for applications such as rehabilitation (Ang & Guan, 2016), emotion recognition (Lazarou et al., 2018), and assistive technology (Wolpaw et al., 2002). Despite these promising applications, their practical utility is limited by poor generalization due to distribution shifts across subjects (Fairclough & Lotte, 2020).

Conventional strategies for mitigating distribution shifts in EEG applications involve collecting labeled calibration data and training domain-specific models (Lotte et al., 2018). While effective in controlled settings, this approach poses significant challenges to the scalability of the technology (Wei et al., 2022). As an alternative, domain adaptation aims to learn a model from source domains that generalizes effectively to unseen but related target domains, offering a principled alternative to training domain-specific models (Ben-David et al., 2010). Within the BCI field, domain adaptation mainly addresses the cross-subject transfer learning problem (Wu et al., 2020), which is one of the grand challenges in EEG research (Fairclough & Lotte, 2020).

Unsupervised domain adaptation (UDA) (Wilson & Cook, 2020) without requiring target domain labels is considered the key to tackling this challenge (Wei et al., 2022). Current literature casts it as a multi-source and multi-target UDA problem, attempting to address it with Riemannian alignment frameworks (Ju et al., 2025). Operating on symmetric positive definite (SPD) manifolds, these alignment frameworks mitigate domain discrepancies by aligning first-order (Zanini et al., 2017) and second-order (Kobler et al., 2022b) statistics, as well as rotational components (Rodrigues et al., 2019) across domains. Once aligned, a model trained on source domains typically generalizes to

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

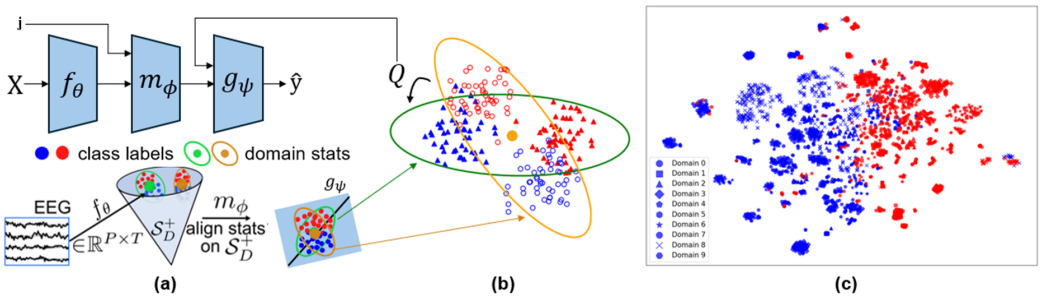

Figure 1: **SPDAlign Overview. (a) Architecture:** an end-to-end Riemannian learning framework on the SPD manifold that aligns domain-specific means and applies orthogonal transformations across domains for multi-source, multi-target unsupervised domain adaptation. **(b) Rotational alignment:** Global rotation shifts between domains are resolved by estimating an orthogonal transformation Q via Wasserstein Procrustes. **(c) t-SNE visualization:** Classification space for the driving fatigue dataset across all subjects. Color represents labels, and shape denotes domains.

related target domains (Rodrigues et al., 2019). However, despite achieving state-of-the-art (SotA) performance (Li et al., 2025; Mellot et al., 2024), these heuristic alignment methods lack theoretical foundations and thus offer limited explanatory power.

In this work, we theoretically demonstrate that Riemannian alignment frameworks can compensate for distribution shifts introduced by the EEG forward modeling process in Eq. (1), which characterizes the observed EEG data as the linear mixing of latent source signals (Sabbagh et al., 2019; Haufe et al., 2014). This scenario represents a form of latent model shift, where discrepancies are driven by the domain-specific forward matrices associated with subject-specific variations such as head geometry (Mellot et al., 2023; Sabbagh et al., 2020). By accounting for domain-specific forward matrices, we show that robust latent representations can be recovered solely through linear transformations on the SPD manifold. Specifically, this involves mean matching followed by a global rotation transformation. The resulting representations encode the latent source covariance, aligning with neuroscience studies that utilize source power as a robust biomarker (Blankertz et al., 2007; Grosse-Wentrup & Buss, 2008).

Building on these theoretical insights and recent advances in end-to-end SPD learning (Kobler et al., 2022a), we propose SPDAlign (Fig. 1 (a)), an end-to-end Riemannian learning framework designed to facilitate EEG domain-invariant representation learning. SPDAlign aligns the domain-specific means (Fig. 1 (a)) and subsequently applies a global rotation transformation across domains (Fig. 1 (b)). These rotations are estimated using orthogonal Procrustes analysis for labeled source domains and Wasserstein Procrustes for unlabeled target domains (Grave et al., 2019). Unlike prior semi-supervised frameworks limited to single-source, single-target settings (Rodrigues et al., 2019; Bleuzé et al., 2022), SPDAlign generalizes to unsupervised multi-source, multi-target adaptation without requiring target domain labels. By addressing forward modeling shifts (Fig. 1 (c)), it aligns representations across subjects and steps toward domain-invariant representations.

SPDAlign is a globally linear framework, comprising a linear feature extractor, linear transformations, and a linear classifier, and therefore is intrinsically interpretable. Classifier weights can be mapped back to spatial patterns and source contributions, supporting transparent and clinically meaningful decision-making (Haufe et al., 2014). We systematically evaluate SPDAlign through simulations under the assumed generative model and show its competitive performance in EEG cross-subject decoding via extensive experiments on public datasets spanning motor imagery, workload, mental stress, and driving fatigue. Moreover, using the mental workload task as a case study, we show that SPDAlign successfully identifies task-discriminative source activities peaking in the theta band within the frontal and parietal regions in Fig. 4, which aligns with neurophysiological markers of mental workload (Borghini et al., 2014; Dehais et al., 2020). These results further demonstrate that SPDAlign can identify frequency ranges of interest, determine the spatial patterns reflecting source–sensor relationships, and address cross-subject variability.

2 PRELIMINARIES

2.1 MULTI-SOURCE MULTI-TARGET UDA

Let x, y, and j denote random variables representing the features, associated true labels, and domains, respectively. Given M labeled source domain datasets $\mathcal{S} = \{(x_i, y_i, j_i) \mid x_i \in \mathcal{X},\, y_i \in \mathcal{Y},\, j_i \in \mathcal{J}_s\}_{i=1}^{L_s}$ and N unlabeled target domain datasets $\mathcal{T} = \{(x_i, j_i) \mid x_i \in \mathcal{X},\, j_i \in \mathcal{J}_t\}_{i=1}^{L_t}$, where $\mathcal{X}$ denotes the input space, $\mathcal{Y}$ the label space, and $\mathcal{J}$ the set of domain identifiers. Here, L_s and L_t denote the number of samples, and y_i and j_i represent the associated label and domain of sample i. In the multi-source multi-target UDA setting, we assume that all source and target domains share the same input and label spaces, and the objective is to learn a model h from $\mathcal{S}$ that generalizes to unseen, unlabeled $\mathcal{T}$ via unsupervised domain adaptation.

2.2 EEG FORWARD MODELING

The EEG signals $x_i \in \mathcal{X} \subseteq \mathbb{R}^{P\times T}$, consisting of epochs of multivariate time-series with T temporal samples and P spatial channels, are modeled as a linear mixing of informative source plus additive noise (Nunez & Srinivasan, 2006):

$$x_i = \mathbf{A}_s \mathbf{z}_i + \mathbf{n}_i, \tag{1}$$

where $z_i \in \mathbb{R}^{Q\times T}$ denotes the activity of informative sources and $\mathbf{n}_i \in \mathbb{R}^P$ the additive noise. The matrix $\mathbf{A}_s \in \mathbb{R}^{P\times Q}$ contains the Q source patterns. As in Sabbagh et al. (2020); Haufe et al. (2014), the brain signals are assumed to arise from activity of uncorrelated sources, and the noise $\mathbf{n}_i = \mathbf{A}_n \boldsymbol{\nu}_i$ is stationary, uncorrelated with the sources, and spans a subspace $\left(\mathbf{A}_n \in \mathbb{R}^{P\times(P-Q)}\right)$ that is shared across observations. The generative model can then be written as:

$$x_i = A_s z_i + A_n \boldsymbol{\nu}_i = A_j s_i. \tag{2}$$

where $s_i \in \mathbb{R}^{P\times T}$ denotes the activity of uncorrelated latent sources; $A_j \in \{A \in \mathbb{R}^{P\times P}\}$ is a domain-specific forward model reflecting latent source activation patterns and is assumed to be invertible (Sabbagh et al., 2020; Kobler et al., 2021; Mellot et al., 2023).

The task label y_i is modeled as a function of the spatial covariance of the latent sources $E_i = \mathrm{Cov}(s_i)$:

$$y_i = b^T f(E_i) + b_0 + \varepsilon_i, \tag{3}$$

where $f(\cdot)$ is a fixed function, b is a weight vector, b_0 is a bias term, and $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ is the noise. Following the log-linear relationships typically observed in oscillatory brain activity (Buzsáki & Mizuseki, 2014), we set $f(\cdot) = \log(\cdot)$. Assuming zero-mean signals (i.e., $\sum_{t=1}^{T} x_t = 0$), the empirical covariance matrix for each multivariate epoch $x_i \in \mathcal{X}$ is defined as:

$$C_{ij} = \mathrm{Cov}(x_i) = \frac{1}{T} x_i x_i^\top \in \mathcal{S}_P^+. \tag{4}$$

Given the linear mixing between observed features and latent sources (Eq. (1)), the sensor-space covariance matrices satisfy:

$$C_{ij} = A_j E_i A_j^\top. \tag{5}$$

Since E_i encodes the target y_i and A_j is invertible, the empirical covariances C_{ij} serve as sufficient statistics for decoding y_i.

2.3 SPD LEARNING IN EEG

SPD-valued tangent space mapping (TSM) learning frameworks have been widely adopted in EEG analysis (Lotte et al., 2018). A nutshell introduction to SPD manifold is presented in Sec. A. Their popularity stems from several favorable properties, including invariance to linear mixing (Congedo et al., 2017), statistical consistency (Sabbagh et al., 2019), and interpretability (Kobler et al., 2021). In this context, the empirical covariance matrix C_i (Eq. (4)) serves as a natural SPD representation (Barachant et al., 2011). Typical TSM learning frameworks first estimate the Fréchet mean $\bar{C}$ of the set $\mathcal{C} = \{\mathrm{Cov}(x_i) \mid x_i \in \mathcal{D}\}$ using Eq. (14), projecting each C_i into the tangent space at $\bar{C}$ via Eq. (16), and finally aligning the data around the identity matrix I via parallel transport Eq. (18) for classification. Related work is discussed in Sec. A.1.

2.4 WASSERSTEIN PROCRUSTES

Orthogonal Procrustes analysis. The orthogonal Procrustes problem seeks an orthogonal matrix Q that best aligns a given set of points $X \in \mathbb{R}^{n\times d}$ with another set of points $Y \in \mathbb{R}^{n\times d}$. When the correspondences between the two point sets are *known* (i.e., which point of X corresponds to which point of Y), the problem:

$$Q^\star =: \arg\min_{Q\in\mathcal{O}_d} \|XQ - Y\|_F^2, \tag{6}$$

admits a closed-form solution (Schönemann, 1966) via the singular value decomposition (SVD) of $X^\top Y$. Let $X^\top Y = USV^\top$; then the optimal transformation is:

$$Q^\star = UV^\top. \tag{7}$$

In the above, $\mathcal{O}_d$ denotes the set of $d \times d$ orthogonal matrices and $\|\cdot\|_F$ is the Frobenius norm. The orthogonality constraint ensures that pairwise distances are preserved, restricting the alignment to a global rotation.

Wasserstein distance. When the correspondences between the two point sets X and Y are *unknown*, the Wasserstein distance provides a principled way to infer correspondences by solving an optimal transport problem (Kuhn, 1955). It finds a transport plan that distributes mass from points in X to points in Y in a way that minimizes the total transportation cost. The squared 2-Wasserstein distance is defined as the distance between two discrete probability distributions $\mathbf{X} = \{x_i\}_{i=1}^n$ with weights $a = (a_1, \ldots, a_n)$ and $\mathbf{Y} = \{y_j\}_{j=1}^m$ with weights $b = (b_1, \ldots, b_m)$ as:

$$W_2^2(X, Y) \;=\; \min_{\pi\in\Pi(a,b)} \sum_{i=1}^{n}\sum_{j=1}^{m} \pi_{ij}\, \|x_i - y_j\|_2^2, \tag{8}$$

where $\Pi(a, b)$ denotes the set of all transport plans $\pi \in \mathbb{R}_+^{n\times m}$ whose row sums equal a and column sums equal b.

Wasserstein Procrustes. In our setting, neither the correspondence between the two point sets nor the underlying orthogonal transformation is known. Our objective is to learn an orthogonal matrix $Q \in \mathcal{O}_d$ that aligns the set X with the set Y, enabling the recovery of correspondences. To measure the similarity between the two sets, we employ the Wasserstein distance and integrate it with the orthogonal Procrustes formulation following (Grave et al., 2019). This yields the *Wasserstein Procrustes* problem, where the goal is to estimate the orthogonal transformation Q and the correspondences π jointly:

$$\min_{Q\in O_d} W_2^2(XQ,\, Y)\,. \tag{9}$$

Exact solutions exist for each sub-optimization problem when the other is fixed (Grave et al., 2019).

3 METHODS

3.1 EEG FORWARD MODELING SHIFTS

Following Li et al. (2025); Sabbagh et al. (2020), we focus on EEG forward modeling shifts where the marginal distribution of the observed data varies across domains while assuming the conditional source-to-label relationship remains invariant. Under this assumption, a shared regression model in Eq. (3) can generalize effectively across domains once these distributional discrepancies are appropriately mitigated. Specifically, we examine shifts introduced by the domain-specific forward matrix A_j (Eq. (1)). This is a case of latent model shifting, where variations in the forward modeling process (head geometry, sensor placement, and cognitive states) induce distribution shifts in the observed sensor data (Mellot et al., 2023). Changes in neural dynamics, such as impedance drift, may also occur but are typically modeled as noise (Akalin Acar & Makeig, 2013). By leveraging the uniqueness of the polar decomposition for square invertible matrices, we decompose A_j as:

$$A_j = Q_j \exp(P_j) \tag{10}$$

where the transformation A_j is expressed as a domain-specific transition $\exp(P_j)$ and orthogonal matrix $Q_j \in \{Q \in \mathbb{R}^{P\times P} : Q^TQ = I_P\}$. We select the polar decomposition to characterize

these shifts because it allows decomposing the possible shifts to translations and rotations on the SPD manifold, therefore naturally linking EEG forward modeling to common geometric operations on SPD matrices (Bhatia & Congedo, 2019). Specifically, mean shifts correspond to global power changes that occur similarly across all channels (Congedo et al., 2017), and rotation shifts correspond to scale shifts that differ across classes.

Let $\mathcal{C} = \{\mathrm{Cov}(x_i) \mid (x_i) \in \mathcal{D}\}$ denote the set of empirical covariance matrices. We define the first stage of our alignment framework, m_{ϕ_1}, as the combination of TSM and recentering (RCT, Eq. (18)) to the identity matrix on the manifold. This stage aligns the domain-specific first-order moments (Fréchet means) $\bar{C}_j$ to the identity matrix for each domain j:

$$m_{\phi_1}(C_{ij}) = \Gamma_{\bar{C}_j \to I} \circ \mathrm{Log}_{\bar{C}_j}(C_{ij}) = \log\left(\bar{C}_j^{-\frac{1}{2}}\, C_{ij}\, \bar{C}_j^{-\frac{1}{2}}\right) \tag{11}$$

For the EEG forward modeling described in Sec. 2.2, an additional rotation transformation in Eq. (11) is beneficial to compensate for the distribution shifts induced by the domain-specific forward matrix A. This transformation recovers robust latent representations $\log(E_i)$ that maintain a linear relationship to the labels y_i via Eq. (3):

$$m_{\phi}(C_{ij}) = \hat{Q}_j^T \log\left(\bar{C}_j^{-\frac{1}{2}} C_{ij} \bar{C}_j^{-\frac{1}{2}}\right) \hat{Q}_j = \log(E_i) \tag{12}$$

In our framework, $\hat{Q}_j$ is estimated via orthogonal Procrustes analysis (Eq. (7)) using labels during training, and via Wasserstein Procrustes (Eq. (9)) without labels during target inference. The outputs $m_\phi(C_i)$ are the robust latent representations and are forwarded to the regression model Eq. (3).

3.2 SPDAlign Framework

As demonstrated by Eq. (12), robust latent representations $\log(E_i)$ can be recovered by compensating for distribution shifts introduced by the domain-specific forward matrix A_j (Eq. (1)) via alignment of domain-specific first-order moments followed by a global rotation. Building on these theoretical insights, we propose SPDAlign: an end-to-end Riemannian deep learning framework on the SPD manifold designed for multi-source/multi-target unsupervised domain adaptation. SPDAlign addresses forward modeling shifts by leveraging source domain labels when available, while operating entirely without labels during target domain inference.

Source training. Following Kobler et al. (2022a), we constrain the hypothesis class $\mathcal{H}$ to functions $h : \mathcal{X} \times \mathcal{D} \to \mathcal{Y}$ that can be decomposed into a composition of a shared feature extractor $f_\theta : \mathcal{X} \to \mathcal{S}_D^+$, an alignment module $m_\phi : \mathcal{S}_d^+ \times \mathcal{D} \to \mathbb{R}^{D(D+1)/2}$, and a shared linear classifier $g_\psi : \mathbb{R}^{D(D+1)/2} \to \mathcal{Y}$ with parameters $\Theta = \{\theta, \phi, \psi\}$. We parametrize $h = g_\psi \circ m_\phi \circ f_\theta$ as a neural network (details in Sec. E) and learn the entire model in an end-to-end fashion, see Fig. 2. To tackle forward modeling shifts, m_ϕ tracks the running domain-specific Fréchet mean during training and recenters these first-order moments to the identity matrix; and employs orthogonal Procrustes analysis in Eq. (6) to estimate the rotation between each domain data point and a reference domain, with labels providing known correspondences. Finally, g_ψ serves as the shared linear classifier.

Model training proceeds in two stages (*algorithm 1*). In the first stage (70% of the training iterations), we train the model without applying the rotation alignment. This stage learns a good feature extractor that can extract common representations across domains. In the second stage, we freeze the parameters θ of the feature extractor, estimate and apply the domain-specific rotation alignment, and subsequently fine-tune the classifier parameters ψ on the aligned representations. To estimate the rotation transformation, the Orthogonal Procrustes (Eq. (6)) is applied during source training using labels to construct class-constrained sample correspondences. We select the reference domain randomly. For each remaining domain, we compare its class-wise sample counts with those of the reference domain and, for any class with unequal sample counts, downsample the larger subset to match the smaller one.

SPDAlign builds on standard simplifying assumptions regarding the linear generative process of EEG (Sabbagh et al., 2019; Kobler et al., 2021). In practice, the method is applied to covariance representations obtained after feature extraction. Because convolutions act as linear operators, this extraction step inherently preserves the structural form of the spatial generative model, so our theoretical analysis remains applicable (Haufe et al., 2014).

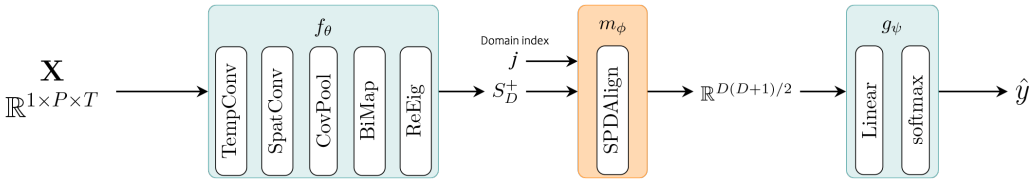

Figure 2: **SPDAlign architecture.** SPDAlign optimizes parameters $\Theta = \{\theta, \phi, \psi\}$ in an end-to-end fashion: (i) a feature extractor f_θ maps raw EEG signals onto the SPD manifold S_D^+; (ii) the SPDAlign module m_ϕ addresses shifts by aligning means and correcting rotations (Eqs. (11) and (12)); and (iii) a shared linear classifier g_ψ outputs predictions $\hat{y}$. By maintaining global linearity, SPDAlign is equivalent to the Tangent Space Mapping framework, providing interpretability.

Target inference. During inference, SPDAlign freezes the learned feature extractor f_θ and the shared linear classifier g_ψ. When samples from previously unseen target domains arrive, it leverages the entire target domain data to estimate the domain-specific mean (Eq. (14)) and the rotation transformation between the reference source domain (set in the training stage) and the target domain data via Wasserstein Procrustes (Eq. (9)). In cases of label shifts, the Wasserstein component can be replaced with Unbalanced Optimal Transport (UOT) (Peyré & Cuturi, 2019; Koç et al., 2026), which relaxes marginal constraints. The target rotations are estimated totally without target labels, and the resulting optimization procedure is detailed in *algorithm 2*. At each iteration t, we compute the optimal transport coupling between the entire reference domain and the target domain data. The resulting transport plan is then used to compute the gradient with respect to the orthogonal matrix $\hat{Q}$, which is updated via projected gradient descent on the orthogonal group. We initialize $\hat{Q}$ to the identity matrix for simplicity and iterate for 50 epochs with a learning rate of 0.01.

Interpretability. Kobler et al. (2021) demonstrated that the coupling of latent sources can be recovered by solving a generalized eigenvalue problem between the Fréchet mean and classifier patterns (Haufe et al., 2014). With an additional linear orthogonal transformation, SPDAlign maintains global linearity, equivalent to the typical TSM learning framework. This global linearity ensures the model preserves a direct correspondence between classifier weights and latent source contributions: the learned representations can be mapped back to the sensor space, and the resulting spatial patterns reflect source–sensor relationships and source contributions. Such interpretability is crucial for determining the specific brain regions (measurement channels) and frequencies that reflect certain cognitive processes and for supporting trustworthy clinical BCI systems (Haufe et al., 2014).

4 Experiments

To evaluate the effectiveness of SPDAlign, we simulated binary classification tasks using balanced accuracy as the evaluation metric, based on the forward model (implementation details in Sec. C). We compared the performance of SPDAlign against RCT (Zanini et al., 2017) across a range of rotational shifts. We varied the parameters related to the forward model: class separability Fig. 5, sample size Fig. 6, dimensionality Fig. 10, number of source domains Fig. 11, informative source Fig. 7, data availability Fig. 9, and label shifts Fig. 8. We discuss the main findings here, and detailed results are provided in the Appendix Sec. C. As expected, classification performance declines as the magnitude of rotational shifts increases, and SPDAlign is superior to RCT across all scenarios. The simulations varying parameters suggest that SPDAlign is robust within typical EEG regimes, where moderate sample sizes are available, low SNR, and sufficiently informative sources. These conditions are consistent with standard BCI datasets, where SPDAlign achieves strong performance.

4.1 Main experiments

We consider four public benchmark EEG datasets: motor imagery (Tangermann et al., 2012), workload estimation (Roy et al., 2022), mental stress (Zyma et al., 2019), and driving fatigue (Cao et al., 2019) (see Sec. D.1 for detailed descriptions and preprocessing). We compare SPDAlign against a comprehensive suite of benchmarks: (i) six deep learning models: EEGNet (Lawhern et al., 2018),

Table 1: **Cross-subject decoding results on four EEG tasks.** Results are reported as the grand average of all test-set balanced accuracy (%) ± standard deviation. The best performing model for each dataset is highlighted. Note that SPDAlign (Oracle) serves as an empirical upper bound, using target domain labels only for optimal rotation estimation through empirical risk minimization.

Model	UDA	Motor Imagery	Workload	Stress	Fatigue
FBCNet	w/o	69.5 ± 10.0	49.0 ± 11.2	67.5 ± 19.7	75.2 ± 8.1
EEGConformer	w/o	67.7 ± 10.4	46.5 ± 11.1	72.0 ± 15.5	74.8 ± 8.5
ATCNet	w/o	67.2 ± 14.9	46.6 ± 9.2	66.8 ± 15.4	75.8 ± 6.3
EEGDeformer	w/o	64.5 ± 8.8	48.1 ± 9.0	68.6 ± 17.4	74.7 ± 10.6
LaBraM	w/o	70.8 ± 9.7	46.7 ± 6.8	75.3 ± 11.7	76.6 ± 9.9
CBraMod	w/o	**73.9 ± 7.3**	49.7 ± 8.0	79.1 ± 14.4	77.4 ± 11.2
LUNA	w/o	68.9 ± 8.6	48.1 ± 10.3	74.7 ± 16.2	75.2 ± 8.9
REVE	w/o	72.5 ± 8.7	48.9 ± 9.6	78.1 ± 10.2	75.7 ± 11.5
EEGNet	w/o	66.9 ± 13.0	48.8 ± 10.7	74.9 ± 19.5	75.9 ± 7.2
	EA	73.2 ± 12.0	55.2 ± 8.6	76.9 ± 18.9	75.6 ± 8.2
	STMA	70.6 ± 9.8	56.1 ± 10.3	80.1 ± 19.2	78.3 ± 8.3
TSLANet	w/o	67.1 ± 13.7	50.5 ± 7.7	68.5 ± 16.7	72.2 ± 7.0
	EA	67.2 ± 12.0	54.3 ± 8.0	71.2 ± 22.6	72.9 ± 10.6
	STMA	68.0 ± 10.7	55.7 ± 8.0	74.3 ± 18.5	73.2 ± 7.9
SPDMBN	w/o	65.5 ± 11.6	47.2 ± 9.2	69.9 ± 19.2	73.4 ± 10.7
	RC	69.0 ± 14.0	56.5 ± 9.9	82.4 ± 22.4	76.6 ± 12.8
TSMNet	SPDDSBN	70.6 ± 11.6	52.5 ± 13.0	76.4 ± 20.8	78.4 ± 7.3
HEEGNet	DSMDBN	73.0 ± 15.5	53.0 ± 10.9	81.5 ± 15.1	76.4 ± 8.1
SPDAlign	-	72.7 ± 13.4	**57.6 ± 10.6**	**84.5 ± 18.7**	**82.7 ± 11.9**
SPDAlign (Oracle)	-	92.6 ± 5.6	88.3 ± 5.5	95.8 ± 4.3	94.8 ± 5.5

EEGConformer (Song et al., 2022), ATCNet (Altaheri et al., 2022), TSLANet (Eldele et al., 2024), EEGDeformer (Ding et al., 2024), and FBCNet (Mane et al., 2021); (ii) four EEG foundation models: LaBraM (Jiang et al., 2024), CBraMod (Wang et al., 2025), Luna (Döner et al., 2025) and Reve (El Ouahidi et al., 2025); (iii) two geometric frameworks: SPDMBN (Kobler et al., 2022b) and its rotation-corrected variant (SPDMBN+RC); and (iv) UDA frameworks: TSMNet (Kobler et al., 2022a), HEEGNet (Li et al., 2026), Euclidean alignment (EA) (He & Wu, 2019) and spatio-temporal Monge alignment (STMA) (Gnassounou et al., 2024) combined with EEGNet and TSLANet (see Sec. D.2 for detailed implementation). We evaluate all models under the same evaluation scheme for cross-subject decoding within each dataset. To establish an empirical upper bound for the Procrustes alignment in the target domain (Koc et al., 2025), we additionally evaluate a variant of SPDAlign in which target labels are used only during the estimation of the orthogonal transformation through empirical risk minimization (ERM); the rest follows the standard source-ERM protocol.

Results. Tab. 1 summarizes the results across all EEG tasks, presenting the grand average of test-set scores. SPDAlign achieves competitive performance, achieving SotA performance in three of four downstream tasks. The alignment trajectory for EEG data is visualized using t-SNE in Fig. 3. Fig. 4 summarizes SPDAlign interpretation results, where patterns are extracted from the workload dataset and class *low* (cross-subject setting). Results for the class *high* can be found in Fig. 12. For the mental workload tasks, the top contributing source activity peaks in the theta band, and source activities are discriminative primarily in frontal and parietal areas. These findings are consistent with neuroscience studies identifying that both of these areas and the theta band are relevant markers of workload (Borghini et al., 2014; Dehais et al., 2020). Beyond competitive predictive performance, SPDAlign can identify frequency ranges of interest, determine the specific brain regions (measured by EEG channels), and address cross-subject variability.

Model Validation. The high performance of the *Oracle* in Tab. 1 supports that forward modeling shifts are the primary cause of distribution shifts in EEG BCI datasets. In this setting, the classifier

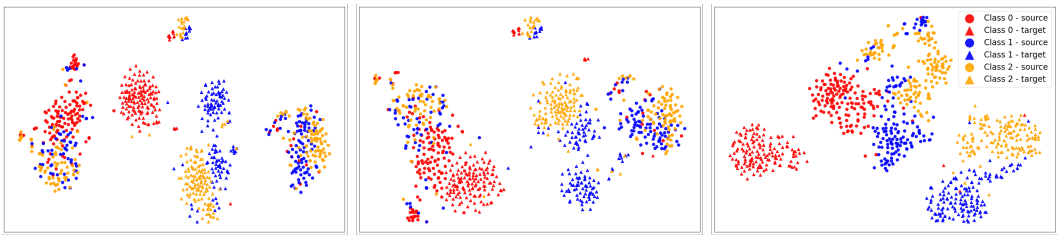

Figure 3: **SPDAlign steps toward domain-invariant representations.** The t-SNE visualization illustrates the alignment process for the workload dataset in the classification space. **(a) Moment alignment:** aligning the domain-specific Fréchet means (Eq. (11)) to the identity matrix (previous work). **(b) SPDAlign** and **(c) SPDAlign (Oracle):** applying the orthogonal transformation Q estimated via Wasserstein Procrustes (Eq. (9)) and empirical risk minimization. SPDAlign enhances class separability, addressing forward modeling shifts without requiring labels. The remaining gap in the unsupervised setting is primarily due to suboptimal rotation estimation.

is fixed and shared across domains, while target labels are used only to estimate the optimal rotation for alignment. The resulting improvements across datasets suggest that a substantial portion of the cross-subject distribution shift can be addressed through conditional rotation alignment. These results suggest that the main driver of the domain shift can be explained by transformations in our model, and the remaining gap in the unsupervised setting is primarily due to suboptimal rotation estimation (visualized in Fig. 3 b and c), not model mismatch. Correcting for these shifts nearly closes the cross-subject decoding gap, which provides strong empirical validation for the generative model in Eq. (1). Furthermore, the high performance implies that the underlying latent source covariance is stable across individuals. These results emphasize the necessity of tackling forward modeling shifts, provide empirical support for our theoretical result, and confirm that the forward model Eq. (1) effectively characterizes the physical processes of EEG.

Room for improvement. While Wasserstein Procrustes estimation allows the unsupervised SPDAlign to achieve state-of-the-art performance, a gap remains relative to the ERM version. This discrepancy arises because the Wasserstein Procrustes in algorithm 2 relies on distribution-level optimal transport to infer point correspondences, which can be ambiguous in the absence of label supervision. When class-conditional structures overlap or are weakly separated, the estimated transport plan may associate semantically mismatched points, leading to a biased estimation of the global rotation (Courty et al., 2016). Such erroneous correspondences can accumulate during the iterative optimization and prevent the recovered rotation from reaching the empirical optimum. Our simulation results (Fig. 5) further show that unsupervised rotation estimation degrades under extreme rotation shifts or limited class separability. These results highlight the potential to develop more robust rotation estimation techniques to bridge the remaining gap toward the oracle upper bound.

Ablation. Tab. 2 summarizes the combination of moment and rotation alignment within SPDAlign. We highlight the three important observations: (i) the full configuration (both components enabled) achieves the best performance, confirming that both alignment steps are necessary; (ii) rotation alignment appears to be particularly critical, as its removal leads to a statistically significant performance drop ($p = 0.02$); and (iii) the omission of both components results in the largest decrease in accuracy, underscoring their effect in mitigating distribution shifts. A global reference mean is used across domains in our implementation; otherwise, the performance will degrade to chance level, which in turn makes it difficult to analyze other parts. This reduces the apparent impact of explicit moment alignment, as partial moment alignment is already implicitly enforced.

5 Discussions

In this work, we presented SPDAlign, a global linear Riemannian learning framework for multi-source/multi-target UDA. By bridging the physical generative model of EEG with Riemannian geometry, we demonstrated that distribution shifts induced by domain-specific forward modeling can

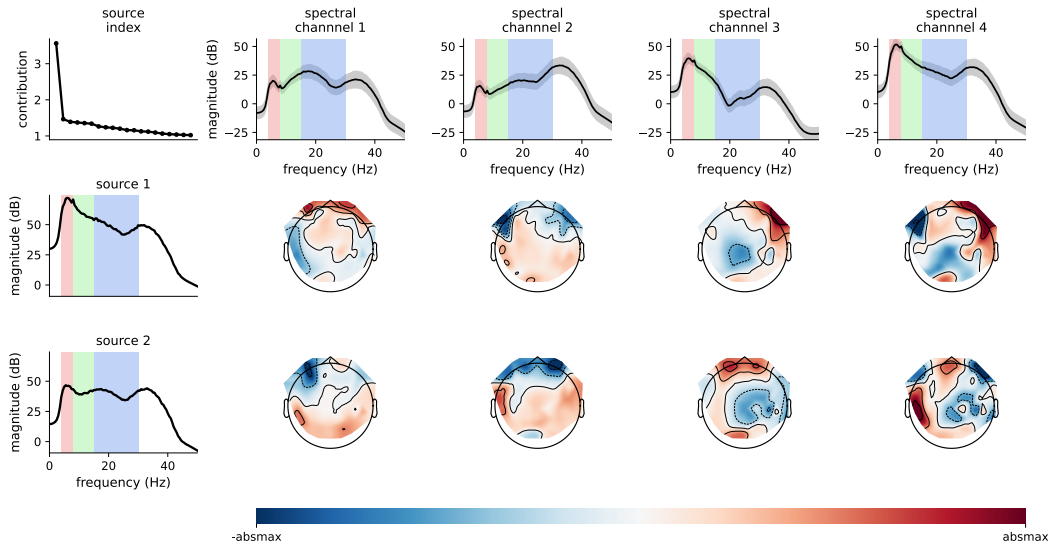

Figure 4: **SPDAlign interpretability.** Patterns extracted from the *workload* dataset and class *low* (cross-subject) are shown above. The top-left panel details the relative contribution of each extracted latent source. For clarity, only the two most discriminative sources are displayed in the left column panels. The top row panels summarize the frequency profile of each spectral channel (derived from the output of the four temporal convolution layers in f_θ). Topographic plots show the projection of source activity to regions covered by the EEG channels; areas shaded in darker blue or red indicate greater discriminative source activity. These results align with prior research, as the frontal and parietal areas and the theta band serve as biomarkers of mental workload (Borghini et al., 2014).

Table 2: **Ablation study on the components of SPDAlign.** Grand average change in balanced accuracy (%) for all test sets relative to the full configuration. Moment alignment refers to Fréchet mean matching, while rotation refers to the (Wasserstein) Procrustes analysis. Permutation-paired t-tests, with t-values and p-values indicating the performance impact of omitting specific components.

moment	rotation	mean $\pm$ std (%)	t-val (p-val)
✓	✓	–	–
✗	✓	-2.5 $\pm$ 20.9	0.82 (0.4)
✓	✗	-5.6 $\pm$ 16.3	2.37 (0.02)
✗	✗	-11.3 $\pm$ 18.4	3.58 (0.0006)

be effectively mitigated through unsupervised alignment on the SPD manifold. Our results highlight the critical role of rotational alignment; as shown in our ablation study (Tab. 2), omitting this step significantly degrades performance. By achieving SotA results on three of four benchmark datasets (Tab. 1), we show that our approach effectively recovers robust latent representations without requiring target labels. Furthermore, we demonstrated the interpretability of SPDAlign; the framework can identify frequency ranges of interest and determine the specific brain regions. Altogether, these results demonstrate the effectiveness of our framework, SPDAlign, which achieves highly competitive performance while retaining inherent interpretability. The limitation of our framework, which applies to any method involving eigendecompositions, is the computational cost, which limits its application to high-dimensional SPD features. While we limited this study to offline evaluations, our framework could be readily extended to online UDA by tracking the running statistics (Kobler et al., 2022a) and replacing the mini-batch Wasserstein distance estimation with online variants (Mensch & Peyré, 2020). Future work could explore more robust rotation estimation methods and extend the framework to online BCI applications.

AI USE STATEMENT

In this work, we used generative AI tools for polishing the grammar and writing. We have not used generative AI tools for any other purpose, such as drafting the paper. We have reviewed all AI-assisted work. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

REFERENCES

P-A Absil, Robert Mahony, and Rodolphe Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2008.

Dinesh Acharya, Zhiwu Huang, Danda Pani Paudel, and Luc Van Gool. Covariance pooling for facial expression recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 367–374, 2018.

Zeynep Akalin Acar and Scott Makeig. Effects of forward model errors on eeg source localization. *Brain topography*, 26(3):378–396, 2013.

Hamdi Altaheri, Ghulam Muhammad, and Mansour Alsulaiman. Physics-informed attention temporal convolutional network for eeg-based motor imagery classification. *IEEE transactions on industrial informatics*, 19(2):2249–2258, 2022. doi: 10.1109/TII.2022.3197419.

Kai Keng Ang and Cuntai Guan. Eeg-based strategies to detect motor imagery for control and rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(4): 392–401, 2016.

Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Multiclass brain–computer interface classification by riemannian geometry. *IEEE Transactions on Biomedical Engineering*, 59(4):920–928, 2011. doi: 10.1109/TBME.2011.2172210.

Gary Bécigneul and Octavian-Eugen Ganea. Riemannian adaptive optimization methods. *arXiv preprint arXiv:1810.00760*, 2018.

Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. A theory of learning from different domains. *Machine learning*, 79:151–175, 2010. doi: 10.1007/s10994-009-5152-4.

Rajendra Bhatia. *Positive definite matrices*. Princeton university press, 2009. ISBN 978-0-691-12918-1.

Rajendra Bhatia. The Riemannian Mean of Positive Matrices. In Frank Nielsen and Rajendra Bhatia (eds.), *Matrix Information Geometry*, pp. 35–51. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. ISBN 978-3-642-30232-9. doi: 10.1007/978-3-642-30232-9_2.

Rajendra Bhatia and Marco Congedo. Procrustes problems in riemannian manifolds of positive definite matrices. *Linear Algebra and its Applications*, 563:440–445, 2019.

Benjamin Blankertz, Ryota Tomioka, Steven Lemm, Motoaki Kawanabe, and Klaus-Robert Muller. Optimizing spatial filters for robust eeg single-trial analysis. *IEEE Signal processing magazine*, 25(1):41–56, 2007.

Alexandre Bleuzé, Jérémie Mattout, and Marco Congedo. Tangent space alignment: Transfer learning for brain-computer interface. *Frontiers in human neuroscience*, 16:1049985, 2022.

Gianluca Borghini, Laura Astolfi, Giovanni Vecchiato, Donatella Mattia, and Fabio Babiloni. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*, 44:58–75, 2014.

György Buzsáki and Kenji Mizuseki. The log-dynamic brain: how skewed distributions affect network operations. *Nature reviews neuroscience*, 15(4):264–278, 2014.

Zehong Cao, Chun-Hsiang Chuang, Jung-Kai King, and Chin-Teng Lin. Multi-channel eeg recordings during a sustained-attention driving task. *Scientific data*, 6(1):19, 2019.

Marco Congedo, Alexandre Barachant, and Rajendra Bhatia. Riemannian geometry for eeg-based brain-computer interfaces; a primer and a review. *Brain-Computer Interfaces*, 4(3):155–174, 2017. doi: 10.1080/2326263X.2017.1297192.

Nicolas Courty, Rémi Flamary, Devis Tuia, and Alain Rakotomamonjy. Optimal transport for domain adaptation. *IEEE transactions on pattern analysis and machine intelligence*, 39(9):1853–1865, 2016.

Frédéric Dehais, Alex Lafont, Raphaëlle Roy, and Stephen Fairclough. A neuroergonomics approach to mental workload, engagement and human performance. *Frontiers in neuroscience*, 14: 268, 2020.

Yi Ding, Yong Li, Hao Sun, Rui Liu, Chengxuan Tong, Chenyu Liu, Xinliang Zhou, and Cuntai Guan. Eeg-deformer: A dense convolutional transformer for brain-computer interfaces. *IEEE Journal of Biomedical and Health Informatics*, 2024.

Berkay Döner, Thorir Mar Ingolfsson, Luca Benini, and Yawei Li. Luna: Efficient and topology-agnostic foundation model for eeg signal analysis. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Yassine El Ouahidi, Jonathan Lys, Philipp Thölke, Nicolas Farrugia, Bastien Pasdeloup, Vincent Gripon, Karim Jerbi, and Giulia Lioi. Reve: A foundation model for eeg-adapting to any setup with large-scale pretraining on 25,000 subjects. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Emadeldeen Eldele, Mohamed Ragab, Zhenghua Chen, Min Wu, and Xiaoli Li. Tslanet: rethinking transformers for time series representation learning. In *Proceedings of the 41st International Conference on Machine Learning*, pp. 12409–12428, 2024.

Stephen H Fairclough and Fabien Lotte. Grand challenges in neurotechnology and system neuroergonomics. *Frontiers in Neuroergonomics*, 1:602504, 2020. doi: 10.3389/fnrgo.2020.602504.

Théo Gnassounou, Antoine Collas, Rémi Flamary, Karim Lounici, and Alexandre Gramfort. Multi-source and test-time domain adaptation on multivariate signals using spatio-temporal monge alignment. *arXiv preprint arXiv:2407.14303*, 2024.

Edouard Grave, Armand Joulin, and Quentin Berthet. Unsupervised alignment of embeddings with wasserstein procrustes. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 1880–1890. PMLR, 2019.

Moritz Grosse-Wentrup and Martin Buss. Multiclass common spatial patterns and information theoretic feature extraction. *IEEE transactions on Biomedical Engineering*, 55(8):1991–2000, 2008.

Stefan Haufe, Frank Meinecke, Kai Görgen, Sven Dähne, John-Dylan Haynes, Benjamin Blankertz, and Felix Bießmann. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage*, 87:96–110, 2014.

He He and Dongrui Wu. Transfer learning for brain–computer interfaces: A euclidean space data alignment approach. *IEEE Transactions on Biomedical Engineering*, 67(2):399–410, 2019. doi: 10.1109/TBME.2019.2913914.

He He and Dongrui Wu. Different set domain adaptation for brain-computer interfaces: A label alignment approach. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28 (5):1091–1108, 2020.

Zhiwu Huang and Luc Van Gool. A Riemannian Network for SPD Matrix Learning. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, AAAI'17, pp. 2036–2042. AAAI Press, 2017.

Wei-Bang Jiang, Li-Ming Zhao, and Bao-Liang Lu. Large brain model for learning generic representations with tremendous eeg data in bci. *arXiv preprint arXiv:2405.18765*, 2024.

Ce Ju, Reinmar J Kobler, Antoine Collas, Motoaki Kawanabe, Cuntai Guan, and Bertrand Thirion. Spd learning for covariance-based neuroimaging analysis: Perspectives, methods, and challenges. *arXiv preprint arXiv:2504.18882*, 2025.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Reinmar Kobler, Jun-ichiro Hirayama, Qibin Zhao, and Motoaki Kawanabe. Spd domain-specific batch normalization to crack interpretable unsupervised domain adaptation in eeg. *Advances in Neural Information Processing Systems*, 35:6219–6235, 2022a.

Reinmar J Kobler, Jun-Ichiro Hirayama, Lea Hehenberger, Catarina Lopes-Dias, Gernot R Müller-Putz, and Motoaki Kawanabe. On the interpretation of linear riemannian tangent space model parameters in m/eeg. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 5909–5913. IEEE, 2021.

Reinmar J. Kobler, Jun-ichiro Hirayama, and Motoaki Kawanabe. Controlling The Fréchet Variance Improves Batch Normalization on the Symmetric Positive Definite Manifold. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3863–3867, Singapore, Singapore, 2022b. IEEE. doi: 10.1109/ICASSP43922.2022.9746629.

Okan Koc, Alexander Soen, Chao-Kai Chiang, and Masashi Sugiyama. Domain adaptation and entanglement: an optimal transport perspective. In *International Conference on Artificial Intelligence and Statistics*, pp. 3034–3042. PMLR, 2025.

Okan Koç, Alexander Soen, Shanglin Li, and Masashi Sugiyama. Incorporating importance weighting in optimal transport based domain alignment. In *Forty-third International Conference on Machine Learning*, 2026.

Max Kochurov, Rasul Karimov, and Serge Kozlukov. Geoopt: Riemannian optimization in pytorch. *arXiv preprint arXiv:2005.02819*, 2020.

Harold W Kuhn. The hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2):83–97, 1955.

Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance. Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces. *Journal of neural engineering*, 15(5):056013, 2018. doi: 10.1088/1741-2552/aace8c.

Ioulietta Lazarou, Spiros Nikolopoulos, Panagiotis C Petrantonakis, Ioannis Kompatsiaris, and Magda Tsolaki. Eeg-based brain–computer interfaces for communication and rehabilitation of people with motor impairment: a novel approach of the 21 st century. *Frontiers in human neuroscience*, 12:14, 2018.

Shanglin Li, Motoaki Kawanabe, and Reinmar J Kobler. Spdim: Source-free unsupervised conditional and label shift adaptation in eeg. In *The Thirteenth International Conference on Learning Representations*, 2025.

Shanglin Li, Shiwen Chu, Okan Koç, Yi Ding, Qibin Zhao, Motoaki Kawanabe, and Ziheng Chen. Heegnet: Hyperbolic embeddings for eeg. In *The Fourteenth International Conference on Learning Representations*, 2026.

Fabien Lotte, Laurent Bougrain, Andrzej Cichocki, Maureen Clerc, Marco Congedo, Alain Rakotomamonjy, and Florian Yger. A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year update. *Journal of neural engineering*, 15(3):031005, 2018. doi: 10.1088/1741-2552/aab2f2.

Ravikiran Mane, Effie Chew, Karen Chua, Kai Keng Ang, Neethu Robinson, A Prasad Vinod, Seong-Whan Lee, and Cuntai Guan. Fbcnet: A multi-view convolutional neural network for brain-computer interface. *arXiv preprint arXiv:2104.01233*, 2021.

Apolline Mellot, Antoine Collas, Pedro LC Rodrigues, Denis Engemann, and Alexandre Gramfort. Harmonizing and aligning m/eeg datasets with covariance-based techniques to enhance predictive regression modeling. *Imaging Neuroscience*, 1:1–23, 2023. doi: 10.1162/imag_a_00040.

Apolline Mellot, Antoine Collas, Sylvain Chevallier, Alexandre Gramfort, and Denis A Engemann. Geodesic optimization for predictive shift adaptation on eeg data. *arXiv preprint arXiv:2407.03878*, 2024.

Arthur Mensch and Gabriel Peyré. Online sinkhorn: Optimal transport distances from sample streams. *Advances in Neural Information Processing Systems*, 33:1657–1667, 2020.

Ernst Niedermeyer and FH Lopes da Silva. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005. ISBN 978-0190228484.

Paul L Nunez and Ramesh Srinivasan. *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA, 2006. ISBN 9780195050387.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011. doi: 10.5555/1953048.2078195.

Gabriel Peyré and Marco Cuturi. *Computational optimal transport: With applications to data science*. Now Foundations and Trends, 2019.

Pedro Luiz Coelho Rodrigues, Christian Jutten, and Marco Congedo. Riemannian Procrustes Analysis: Transfer Learning for Brain–Computer Interfaces. *IEEE Transactions on Biomedical Engineering*, 66(8):2390–2401, 2019. doi: 10.1109/TBME.2018.2889705.

Raphaëlle N. Roy, Marcel F. Hinss, Ludovic Darmet, Simon Ladouce, Emilie S. Jahanpour, Bertille Somon, Xiaoqi Xu, Nicolas Drougard, Frédéric Dehais, and Fabien Lotte. Retrospective on the First Passive Brain-Computer Interface Competition on Cross-Session Workload Estimation. *Frontiers in Neuroergonomics*, 3:838342, 2022. doi: 10.3389/fnrgo.2022.838342.

David Sabbagh, Pierre Ablin, Gaël Varoquaux, Alexandre Gramfort, and Denis A Engemann. Manifold-regression to predict from meg/eeg brain signals without source modeling. *Advances in Neural Information Processing Systems*, 32, 2019.

David Sabbagh, Pierre Ablin, Gaël Varoquaux, Alexandre Gramfort, and Denis A Engemann. Predictive regression modeling with meg/eeg: from source power to signals and cognitive states. *NeuroImage*, 222:116893, 2020. doi: j.neuroimage.2020.116893.

Peter H Schönemann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31(1):1–10, 1966.

Yonghao Song, Qingqing Zheng, Bingchuan Liu, and Xiaorong Gao. Eeg conformer: Convolutional transformer for eeg decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:710–719, 2022. doi: 10.1109/TNSRE.2022.3230250.

Michael Tangermann, Klaus-Robert Müller, Ad Aertsen, Niels Birbaumer, Christoph Braun, Clemens Brunner, Robert Leeb, Carsten Mehring, Kai J Miller, Gernot R Müller-Putz, et al. Review of the bci competition iv. *Frontiers in neuroscience*, 6:55, 2012.

Jiquan Wang, Sha Zhao, Zhiling Luo, Yangxuan Zhou, Haiteng Jiang, Shijian Li, Tao Li, and Gang Pan. Cbramod: A criss-cross brain foundation model for eeg decoding. In *The Thirteenth International Conference on Learning Representations*, 2025.

Xiaoxi Wei, A. Aldo Faisal, Moritz Grosse-Wentrup, Alexandre Gramfort, Sylvain Chevallier, Vinay Jayaram, Camille Jeunet, Stylianos Bakas, Siegfried Ludwig, Konstantinos Barmpas, Mehdi Bahri, Yannis Panagakis, Nikolaos Laskaris, Dimitrios A. Adamos, Stefanos Zafeiriou, William C. Duong, Stephen M. Gordon, Vernon J. Lawhern, Maciej Śliwowski, Vincent Rouanne, and Piotr Tempczyk. 2021 BEETL Competition: Advancing Transfer Learning for Subject Independence and Heterogenous EEG Data Sets. In Douwe Kiela, Marco Ciccone, and Barbara Caputo (eds.), *Proceedings of the NeurIPS 2021 Competitions and Demonstrations Track*, volume 176 of *Proceedings of Machine Learning Research*, pp. 205–219. PMLR, 2022.

Garrett Wilson and Diane J Cook. A survey of unsupervised deep domain adaptation. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 11(5):1–46, 2020.

Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791, 2002. doi: 10.1016/S1388-2457(02)00057-3.

Dongrui Wu, Yifan Xu, and Bao-Liang Lu. Transfer learning for eeg-based brain–computer interfaces: A review of progress made since 2016. *IEEE Transactions on Cognitive and Developmental Systems*, 14(1):4–19, 2020. doi: 10.1109/TCDS.2020.3007453.

Paolo Zanini, Marco Congedo, Christian Jutten, Salem Said, and Yannick Berthoumieu. Transfer learning: A riemannian geometry framework with applications to brain–computer interfaces. *IEEE Transactions on Biomedical Engineering*, 65(5):1107–1116, 2017. doi: 10.1109/TBME.2017.2742541.

Igor Zyma, Sergii Tukaev, Ivan Seleznov, Ken Kiyono, Anton Popov, Mariia Chernykh, and Oleksii Shpenkov. Electroencephalograms during mental arithmetic task performance. *Data*, 4(1):14, 2019.

A SPD MANIFOLD IN A NUTSHELL

The SPD manifold is defined as $\mathcal{S}_D^+ = \{C \in \mathbb{R}^{D\times D} : C^T = C, C \succ 0\}$, equipped with an inner product on the tangent space $\mathcal{T}_C\mathcal{S}_D^+$ at each point $C \in \mathcal{S}_D^+$, it forms a Riemannian manifold M. Tangent spaces possess a Euclidean structure, providing a local linear approximation to Riemannian distances on $\mathcal{S}_D^+$ (Absil et al., 2008). In this study, we adopt the affine-invariant Riemannian metric $g_C^{\text{AIRM}}(S_1, S_2)$=$\text{Tr}(C^{-1}S_1C^{-1}S_2)$ as the chosen inner product, leading to the distance (Bhatia, 2009):

$$\delta(C_1, C_2) = \| \log(C_1^{-\frac{1}{2}} C_2 C_1^{-\frac{1}{2}}) \|_F, \tag{13}$$

where C_1 and C_2 denote two SPD matrices, $\log(\cdot)$ is the matrix logarithm, $\text{Tr}(\cdot)$ is the trace, and $\|\cdot\|_F$ is the Frobenius norm. Given a set $\mathcal{C} = \{C_i \in \mathcal{S}_D^+\}_{i\leq N}$, the Fréchet mean is the minimizer of the average squared distances:

$$\bar{C} = \arg\min_{G\in\mathcal{S}_D^+} \frac{1}{N}\sum_{i=1}^{N} \delta^2(G, C_i). \tag{14}$$

For $N = 2$, there is a closed-form solution expressed as

$$\bar{C}(\gamma) = C_1 \#_\gamma C_2 = C_1^{\frac{1}{2}} \left(C_1^{-\frac{1}{2}} C_2 C_1^{-\frac{1}{2}} \right)^\gamma C_1^{\frac{1}{2}}. \tag{15}$$

with weight $\gamma = 0.5$. Choosing $\gamma \in [0, 1]$ computes weighted means along the geodesic (i.e., the shortest curve) that connects both points.

The logarithmic map $\text{Log}_{\bar{C}} : \mathcal{S}_D^+ \to \mathcal{T}_{\bar{C}}\mathcal{S}_D^+$ projects points from the manifold onto the tangent space at $\bar{C}$ and exponential map $\text{Exp}_{\bar{C}} : \mathcal{T}_{\bar{C}}\mathcal{S}_D^+ \to \mathcal{S}_D^+$ performs the inverse projection.

$$\text{Log}_{\bar{C}}(C_i) = \bar{C}^{\frac{1}{2}} \log\left(\bar{C}^{-\frac{1}{2}} C_i \bar{C}^{-\frac{1}{2}} \right) \bar{C}^{\frac{1}{2}}, \tag{16}$$

$$\text{Exp}_{\bar{C}}(S_i) = \bar{C}^{\frac{1}{2}} \exp(\bar{C}^{-\frac{1}{2}} S_i \bar{C}^{-\frac{1}{2}}) \bar{C}^{\frac{1}{2}}. \tag{17}$$

Parallel transport $\Gamma_{\bar{C}\to\bar{C}_\phi} \mathcal{T}_{\bar{C}}S_D^+ \to \mathcal{T}_{\bar{C}_\phi}S_D^+$ transports tangent space vectors $S_i \in \mathcal{T}_{\bar{C}}\mathcal{S}_D^+$ at $\bar{C}$ to the tangent space of an arbitrary point $\bar{C}_\phi$ on the manifold:

$$\Gamma_{\bar{C}\to\bar{C}_\phi}(S_i) = \left(\bar{C}_\phi \bar{C}^{-1}\right)^{\frac{1}{2}} S_i \left(\bar{C}_\phi \bar{C}^{-1}\right)^{\frac{1}{2}}. \tag{18}$$

A.1 RELATED WORK

Riemannian geometry-aware alignment frameworks have dominated in EEG cross-domain generalization problems (Roy et al., 2022; Ju et al., 2025; Mellot et al., 2023). Within the typical sequential TSM learning framework, (Rodrigues et al., 2019; Bleuzé et al., 2022) match a source and a target distribution using simple geometric transformations: first- and second-order moment alignment and a global rotation transformation over the data points in a semi-supervised fashion. More recently, (Kobler et al., 2022a) proposed tracking the domain-specific running Fréchet mean during training, thereby achieving end-to-end learning of feature extractors, TSM, and a tangent space classifier.

While our framework adopts similar geometric transformations, it introduces several key distinctions. Unlike the previous end-to-end learning framework of (Kobler et al., 2022a), which aligns only domain-specific moments, SPDAlign additionally accounts for rotation transformations. We further show the importance of rotation alignment in multi-source, multi-target settings that was overlooked in the framework in (Kobler et al., 2022a) through simulations (Fig. 11) and ablation (Tab. 2). In contrast to existing rotation alignment methods (Rodrigues et al., 2019; Bleuzé et al., 2022) that rely on target domain labels for estimation, SPDAlign is the first unsupervised framework capable of estimating these rotations in EEG without label access. Additionally, SPDAlign addresses the multi-source, multi-target adaptation setting, reflecting real-world BCI scenarios where data from multiple recorded subjects are leveraged to enable continuous adaptation to new users (He & Wu, 2020).

B Derivation of domain-invariant latent representations

Leveraging the linear relationship between the latent covariance and the empirical covariance (Eq. (5)), together with the congruence invariance property of the Fréchet mean (Bhatia, 2013), we obtain:

$$\bar{C}_j = A_j \bar{E}_j A_j^T \tag{19}$$

where $\bar{E}_j$ is the Fréchet mean of the latent covariance.

In previous work, Li et al. (2025) prove a proposition that for the set of examples $\mathcal{E}_i = \{E_i | E_i \in \mathcal{S}_P^+\}_{i=1}^L$, the Fréchet mean $\mathcal{E}_j$, defined in (14), converges as $L \to \infty$ to the identity matrix I_P (i.e., $\bar{E}_j = I_P$ for all j). Using the proposition $\bar{E}_j = I_P$ for all j and substituting (Eq. (10)) for A_j, the expression in (Eq. (19)) simplifies to:

$$\begin{aligned}
\bar{C}_j &= A_j I_P A_j^T && (20)\\
&= Q_j \exp(P_j) \exp(P_j) Q_j^T && (21)\\
&= Q_j \exp(P_j)^2 Q_j^T && (22)\\
\Rightarrow \bar{C}_j^{-\frac{1}{2}} &= Q_j \exp(P_j)^{-1} Q_j^T && (23)
\end{aligned}$$

The two-stage transformations, as defined in Eq. (11) and Eq. (12), consisting of recentering the data $C_i \in \mathcal{C}_j$ and a global rotation transformation for each domain j, together with Eq. (23), Eq. (10), and Eq. (18) gives:

$$\begin{aligned}
m_\phi(C_{ij}) &= Q_j^T (\Gamma_{\bar{C}_j \to I} \circ \mathrm{Log}_{\bar{C}_j}(C_{ij})) Q_j && (24)\\
&= Q_j^T (\Gamma_{\bar{C}_j \to I} \circ (\bar{C}_j^{\frac{1}{2}} \log\Big(\bar{C}_j^{-\frac{1}{2}} C_{ij} \bar{C}_j^{-\frac{1}{2}}\Big) \bar{C}_j^{\frac{1}{2}})) Q_j && (25)\\
&= Q_j^T (\log\Big(\bar{C}_j^{-\frac{1}{2}} C_{ij} \bar{C}_j^{-\frac{1}{2}}\Big)) Q_j && (26)\\
&= Q_j^T \log\left(Q_j \exp(P_j)^{-1} Q_j^T C_{ij} Q_j \exp(P_j)^{-1} Q_j^T\right) Q_j && (27)\\
&= Q_j^T \log\left(Q_j \exp(P_j)^{-1} A_j E_i A_j^T \exp(P_j)^{-1} Q_j^T\right) Q_j && (28)\\
&= Q_j^T \log\left(Q_j \exp(P_j)^{-1} Q_j^T Q_j \exp(P_j) E_i \exp(P_j) Q_j^T Q_j \exp(P_j)^{-1} Q_j^T\right) Q_j && (29)\\
&= Q_j^T \log\left(Q_j E_i Q_j^T\right) Q_j && (30)\\
&= Q_j^T Q_j \log(E_i) Q_j^T Q_j && (31)\\
&= \log(E_i) && (32)
\end{aligned}$$

where Eq. (5) is used for C_{ij}, for any invertible matrix G and any $X \in \mathcal{S}_P^+$, the matrix logarithm satisfies $\log(GXG^{-1}) = G \log(X) G^{-1}$.

C SIMULATIONS

To construct the covariance matrices $C_i \in \mathcal{S}_P^+$ with $P = 2$, we initiated the generation process in the log-space. Specifically, log-space features, denoted as $\mathrm{upper}\log(E_i)$, were synthesized using the `scikit-learn` function `make_classification`, with one dimension configured to encode label-relevant information. These features were subsequently standardized to achieve zero mean and unit variance. To map these points onto the SPD manifold $\mathcal{S}_P^+$, we applied the inverse vectorization operator upper^{-1} and the exponential map Exp_{I_P} as defined in (17). The resulting data were partitioned across source and target domains, with each domain receiving 500 observations. Finally, the data were projected into the channel space via domain-specific forward matrices A_j, following (10). To introduce rotation shifts, we either artificially varied the rotation degree using angle-parameterized rotation matrices or generated random orthogonal matrices.

Class separability Fig. 5: in the absence of rotation, moment alignment alone effectively compensates for the domain-specific transition effects introduced by EEG forward modeling shifts (Eq. (10)). As expected, classification performance declines as the magnitude of rotational shifts increases across all scenarios, though higher class separability confers greater robustness. Notably, while the RCT method (green) is highly sensitive to even mild rotations, SPDAlign (red) demonstrates superior resilience, maintaining stable accuracy across a wider range of shifts. The source-domain performance (blue line) represents the empirical upper bound for these unsupervised adaptation tasks.

Sample size Fig. 6: Performance degrades gradually as the number of samples decreases, primarily affecting the estimation of covariance matrices and Wasserstein Procrustes. However, SPDAlign maintains a clear advantage once moderate sample sizes are available (e.g., typical EEG epoch counts), even when the rotation is severe.

Informative sources (Fig. 7): we vary the number of label-informative latent sources to evaluate the effect of distributed task-relevant information. Across different settings, classification performance decreases as rotational shifts become more severe, while SPDAlign generally remains more robust than RCT. The advantage is particularly evident when sufficient informative source structure is available, supporting the robustness of SPDAlign under typical EEG regimes.

Label shifts (Fig. 8): the ratio represents the degree of class imbalance in the simulation. We compared the optimal transport (OT) method used in our paper with its unbalanced variant, UOT. Performance was tested under four different ratios, ranging from mild to severe label shifts. As expected, label shifts degraded performance, but UOT, which relaxes the marginal constraints, still provided performance gains, indicating that UOT is a suitable choice when class imbalance is present.

Data availability (Fig. 9): to examine the effect of data availability in an online learning setting, we progressively vary the proportion of target-domain samples available for alignment. With limited incoming data, the estimation is less reliable, resulting in a smaller advantage over RCT. As more target samples become available, SPDAlign consistently shows advantages over RCT. This result suggests that SPDAlign can be easily extended to online settings, where its alignment quality progressively improves as data accumulate.

SNR (Fig. 10): As SNR decreases (i.e., increasing proportion of noise sources in the simulation), all methods' performance declines. SPDAlign shows a clear advantage under mild noise conditions (-3 dB), remains beneficial under moderate noise (-10.5 dB), and maintains similar performance without degrading results under more extreme noise levels (-14.6 dB, -17.4 dB).

Domain numbers (Fig. 11): we evaluate performance on consistent target domains across a varying number of source domains, comparing the impact of enabling or disabling source domain rotation correction. In the absence of such correction, training performance degrades as the number of source domains increases, due to the accumulation of conflicting domain-specific rotation shifts, underscoring the importance of source-domain alignment in multi-source settings that was overlooked in the framework proposed by Kobler et al. (2022a).

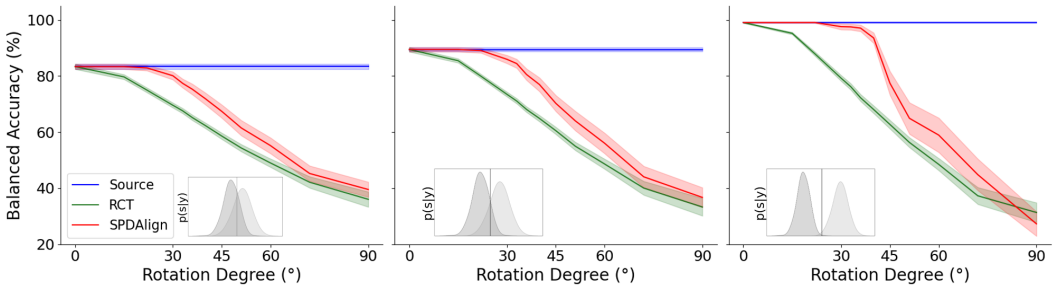

Figure 5: **Simulation performance evaluation across varying class separability.** The panels illustrate the balanced accuracy (y-axis) of different alignment methods as a function of the rotational shift in degrees (x-axis) introduced by domain-specific forward modeling shifts. Each panel represents a distinct level of class separability, with the degree of class overlap visualized in the gray insets. We compare the performance of SPDAlign against recentering to the identity (RCT) (Zanini et al., 2017). The blue line indicates source domain performance, providing an empirical upper bound for the UDA task.

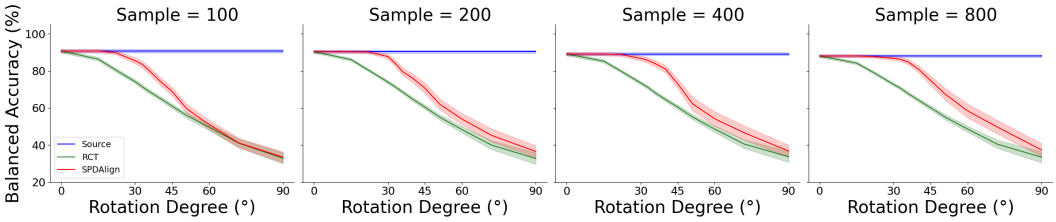

Figure 6: **Simulation performance over the number of data samples.** Balanced accuracy scores (higher is better) across rotation degree on the x-axis and number of data samples across panels. Larger sample sizes lead to a clearer advantage of SPDAlign over RCT under moderate rotational shifts.

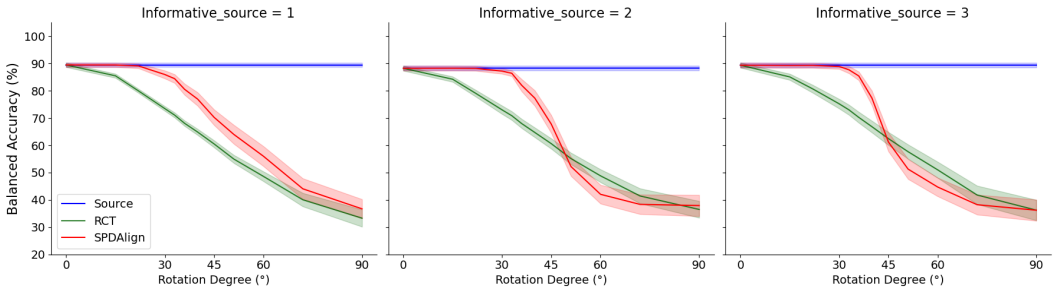

Figure 7: **Simulation performance over the number of informative sources.** Balanced accuracy scores (higher is better) across rotation degree on the x-axis and number of informative sources across panels. SPDAlign generally remains more robust than RCT; the advantage is particularly evident when sufficient informative source structure is available, supporting the robustness of SPDAlign under typical EEG regimes.

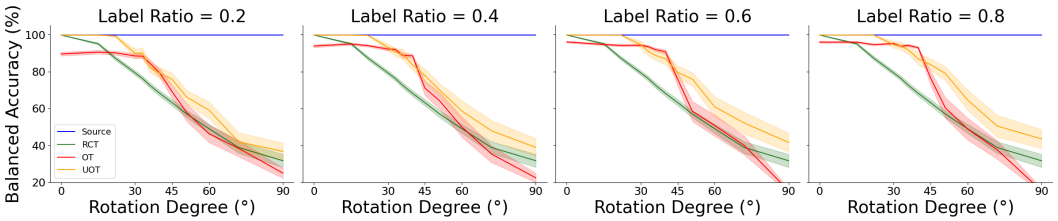

Figure 8: **Simulation performance over label shift ratios.** Balanced accuracy scores (higher is better) across rotation degree on the x-axis. As expected, label shifts degraded performance, but UOT, which relaxes the marginal constraints, still provided performance gains, indicating that UOT is a suitable choice when class imbalance is present.

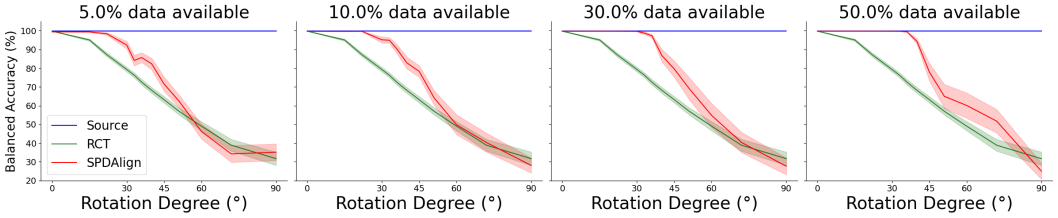

Figure 9: **Simulation performance over available data percentage.** Balanced accuracy scores (higher is better) across rotation degree on the x-axis. With limited incoming data, the estimation is less reliable, resulting in a smaller advantage over RCT. As more target samples become available, SPDAlign consistently shows advantages over RCT.

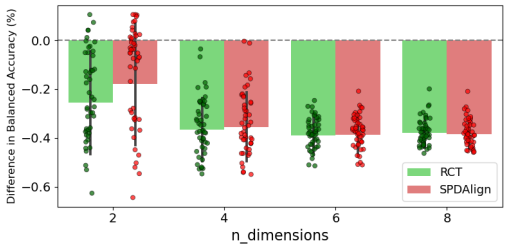

Figure 10: **Simulation performance over the number of data dimensions.** Balanced accuracy scores (higher is better) across the number of data dimensions on the x-axis. As the proportion of noise sources increases (i.e., as SNR decreases), all methods' performance declines. SPDAlign shows a clear advantage under mild noise conditions, remains beneficial under moderate noise, and maintains similar performance without degrading results under more extreme noise levels.

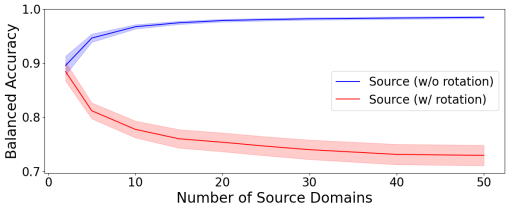

Figure 11: **Simulation performance over the number of source domains.** Balanced accuracy scores (higher is better) across the number of source domains on the x-axis. In the absence of source domain rotation correction, training performance degrades as the number of source domains increases, due to the accumulation of conflicting domain-specific rotation shifts.

D EXPERIMENT DETAILS

D.1 DATASET DETAILS

Motor imagery: (Tangermann et al., 2012) is a motor imagery dataset containing EEG from 9 subjects recorded with 22 channels at 250 Hz. The paradigm involves two motor imagery tasks: left hand and right hand, with 144 trials per class per subject. Data were collected across two sessions on separate days, each comprising six runs of 48 trials. Each trial lasted 4 seconds.

Workload: (Roy et al., 2022) containing EEG from 15 subjects recorded with 64 channels at 500 Hz. The paradigm involves three distinct mental workload levels: low, medium, and high, induced by the Multi-Attribute Task Battery II (MATB-II) task. Each participant completed three sessions on different days, with each workload condition lasting 5 minutes per session.

Mental Stress: (Zyma et al., 2019) containing EEG from 36 subjects recorded with 19 channels at 500 Hz using the Neurocom system. The paradigm involves two cognitive states: a resting state and a mental stress state induced by an intensive serial subtraction arithmetic task. Each participant completed a single session consisting of a 3-minute resting period followed by 4 minutes of mental arithmetic.

Driving Fatigue: (Cao et al., 2019) containing EEG from 11 subjects recorded with 32 channels at 500 Hz. The paradigm involves a prolonged 90-minute dynamic driving simulation designed to induce varying levels of fatigue. Continuous EEG recordings were collected, with lane departure events and reaction times monitored to assess alertness and fatigue throughout the session.

Preprocessing. For motor imagery, signals were resampled to 256 Hz, band-pass filtered between 4–36 Hz, and segmented into 3-second epochs corresponding to the two most discriminative labels. For workload, signals were band-pass filtered at 4–36 Hz, resampled to 200 Hz, and segmented into 2-second epochs. For driving fatigue, signals were band-pass filtered at 1–50 Hz, resampled to 128 Hz, and segmented into 3-second epochs. For mental stress, signals were processed with a 30 Hz low-pass filter and a 50 Hz notch filter, resampled to 500 Hz, and segmented into 4-second epochs.

D.2 IMPLEMENTATION DETAILS

All models were implemented using official repositories, maintaining the default hyperparameters. Each session is treated as a domain. We evaluate using either leave-one-subject-out (subject count ≤ 10) or 10-fold cross-validation, with subjects serving as the grouping variable for cross-subject adaptation. Euclidean frameworks use the Adam optimizer (Kingma & Ba, 2014), while geometric frameworks use Riemannian Adam (Bécigneul & Ganea, 2018), both with default PyTorch hyperparameters. Source-domain data are split into 80% training and 20% validation sets, stratified by domain and label, and trained for 50 epochs with cross-entropy loss.

Baseline methods were implemented using publicly available Python code; custom approaches were developed using PyTorch (Paszke et al., 2019), scikit-learn (Pedregosa et al., 2011), and Geoopt (Kochurov et al., 2020). All experiments were performed on standard computation PCs equipped with 32-core CPUs, 128 GB of RAM, and a single GPU.

E SPDALIGN DETAILS

E.1 MODEL PARAMETERS

For the feature extractor f_θ, we adopt the architecture from Kobler et al. (2022a), which includes two linear convolutional layers, covariance pooling (Acharya et al., 2018), and BiMap and ReEig layers (Huang & Gool, 2017). Specifically, the BiMap layer performs a linear subspace projection, and the ReEig layer thresholds the eigenvalues of symmetric matrices with the default threshold of 10^{-4}; it was never activated.

Table 3: SPDAlign parameters

Block	Output (dim)	Parameter (dim)
TempConv	$4 \times P \times T$	$4 \times 1 \times 1 \times 25$
SpatConv	$40 \times 1 \times T$	$40 \times 4 \times P \times 1$
CovPool	40×40	–
BiMap	20×20	40×20
ReEig	20×20	–
SPDAlign	20×20	1
LogEig	210	–
Linear	C	$211 \times C$

E.2 SPDALIGN SOURCE DOMAIN TRAINING

Algorithm 1 SPDAlign source training

Input: batch $\mathcal{B}_k = \{(\mathbf{x}_i, y_i, d(i))\}_{i=1}^{M}$, where $d(i)$ denotes the associated source domain; domain-specific running Fréchet means $\{\bar{\mathbf{C}}_{k-1}^d\}_{d\in\mathcal{J}_s}$, with $\bar{\mathbf{C}}_0^d = \mathbf{I}$; training momentum η_k; model parameters: feature extractor f_θ and classifier g_ψ
Output: normalized representations $\{\widetilde{\mathbf{Z}}_i\}$, updated running means $\{\bar{\mathbf{C}}_k^d\}$, and updated θ, ψ
$\mathbf{C}_i = f_\theta(\mathbf{x}_i)$ {extract SPD features}
for each domain d in $\mathcal{B}_k$ **do**
  $\mathbf{G}_k^d = \arg\min_{\mathbf{G}\in\mathcal{S}_D^+} \frac{1}{|\mathcal{B}_k^d|} \sum_{i:d(i)=d} \delta^2(\mathbf{G}, \mathbf{C}_i)$ {domain-specific batch mean, Eq. (14)}
  $\bar{\mathbf{C}}_k^d = \bar{\mathbf{C}}_{k-1}^d \#_{\eta_k} \mathbf{G}_k^d$ {update running mean, Eq. (15)}
  $\mathbf{Z}_i = \log\Big[\big(\bar{\mathbf{C}}_k^d\big)^{-\frac{1}{2}} \mathbf{C}_i \big(\bar{\mathbf{C}}_k^d\big)^{-\frac{1}{2}}\Big], \quad d(i) = d$ {domain-specific recentering, Eq. (11)}
end for
$\hat{y}_i = g_\psi\left(\mathrm{upper}(\mathbf{Z}_i)\right)$
Update θ and ψ by minimizing $\sum_i \ell(\hat{y}_i, y_i)$
Rotation alignment stage:
Freeze θ and randomly select reference domain $r \in \mathcal{J}_s$
for each source domain $d \neq r$ **do**
  Construct class-constrained correspondences between domains d and r
  Downsample unequal class-wise sample counts
  $\widehat{\mathbf{Q}}^d = \mathrm{Procrustes}(\mathbf{Z}^d, \mathbf{Z}^r)$ {estimate source rotation, Eq. (7)}
  $\widetilde{\mathbf{Z}}_i = (\widehat{\mathbf{Q}}^d)^\top \mathbf{Z}_i \widehat{\mathbf{Q}}^d = \log(E_i), \quad d(i) = d$ {normalized representations, Eq. (12)}
end for
$\widehat{\mathbf{Q}}^r = \mathbf{I}$
Fine-tune ψ using $\{\widetilde{\mathbf{Z}}_i, y_i\}$, while keeping θ fixed

E.3 WASSERSTEIN PROCRUSTES IN SPDALIGN

Algorithm 2 Wasserstein Procrustes in SPDAlign

Input: Source data $\mathbf{X} \in \mathbb{R}^{n_a \times d}$, Target data $\mathbf{Y} \in \mathbb{R}^{n_b \times d}$, learning rate α, iterations T
Initialize $\mathbf{Q}_0 \in \mathcal{O}_d$ (orthogonal matrix)
Set uniform weights $\mathbf{a} = \frac{1}{n_a}\mathbf{1}$, $\mathbf{b} = \frac{1}{n_b}\mathbf{1}$
for $t = 1$ **to** T **do**
    Compute the optimal transport coupling given $\mathbf{Q}_t$:

$$\boldsymbol{\pi}_t = \arg\min_{\boldsymbol{\pi} \in \Pi(\mathbf{a},\mathbf{b})} \sum_{i,j} \boldsymbol{\pi}_{ij} \left\| \mathbf{Q}_t \mathbf{x}_i - \mathbf{y}_j \right\|_2^2$$

    Compute the gradient with respect to $\mathbf{Q}$:

$$\mathbf{G}_t = -2\mathbf{X}^\top \boldsymbol{\pi}_t \mathbf{Y}$$

    Projected Gradient Descent on the orthogonal group:

$$\mathbf{Q}_{t+1} = \Pi_{\mathcal{O}_d}(\mathbf{Q}_t - \alpha \mathbf{G}_t)$$

end for
The projection $\Pi_{\mathcal{O}_d}(\mathbf{Q})$ is given by $\mathbf{U}\mathbf{V}^\top$, where $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top$ is the SVD of $\mathbf{Q}$.
Output: $\mathbf{Q} = \mathbf{U}\mathbf{V}^\top$

E.4 INTERPRETABILITY

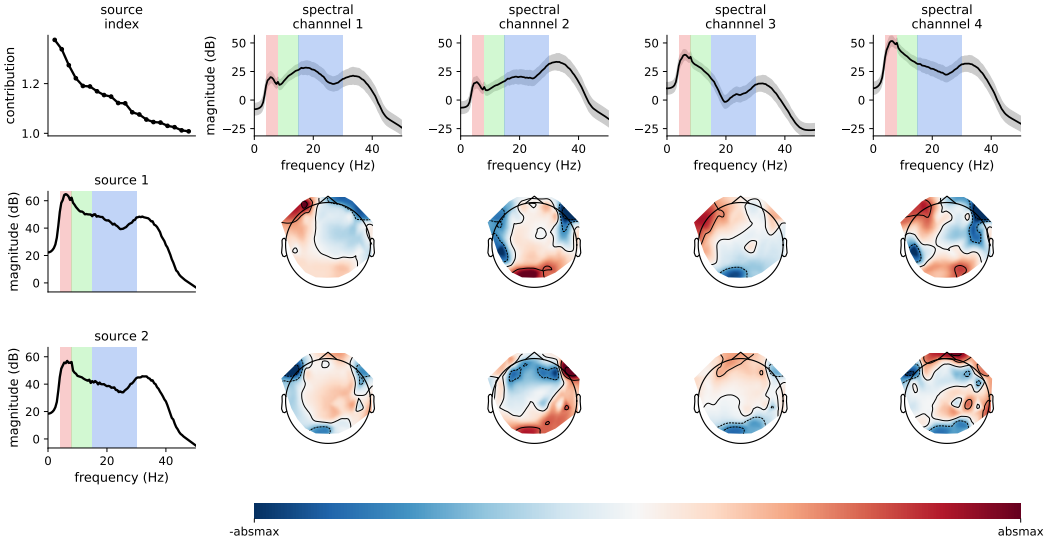

Figure 12: SPDAlign interpretation results. Patterns extracted from the workload dataset and class high (cross-subject setting).