

PAPER • OPEN ACCESS

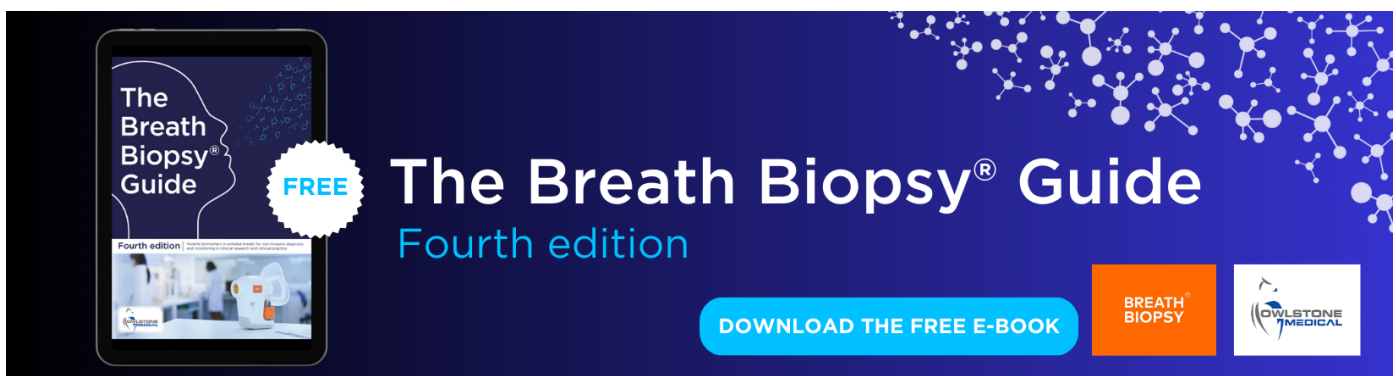
## Improving EEG-based decoding of the locus of auditory attention through domain adaptation<sup>\*</sup>

To cite this article: Johanna Wilroth *et al* 2023 *J. Neural Eng.* **20** 066022

View the [article online](#) for updates and enhancements.

### You may also like

- [Numerical calculation of the turbulent flow past a surface mounted cube with assimilation of PIV data](#)  
Konstantinos Kellaris, Nikolaos Petros Pallas and Demetri Bouris
- [Multi-channel EEG-based sleep staging using brain functional connectivity and domain adaptation](#)  
Wenhao Yuan, Wentao Xiang, Kaiyue Si et al.
- [Relating EEG to continuous speech using deep neural networks: a review](#)  
Corentin Puffay, Bernd Accou, Lies Bollens et al.



**The Breath Biopsy® Guide**  
Fourth edition

DOWNLOAD THE FREE E-BOOK

BREATH BIOPSY

OWLSTONE MEDICAL



## PAPER

## OPEN ACCESS

RECEIVED  
29 June 2022

REVISED  
1 November 2023

ACCEPTED FOR PUBLICATION  
21 November 2023

PUBLISHED  
1 December 2023

Original Content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).

Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.



# Improving EEG-based decoding of the locus of auditory attention through domain adaptation\*

Johanna Wilroth<sup>1,\*\*</sup> , Bo Bernhardsson<sup>3</sup> , Frida Heskebeck<sup>3</sup> , Martin A Skoglund<sup>1,2</sup> ,  
Carolina Bergeling<sup>4,5</sup> and Emina Alickovic<sup>1,2,5</sup>

<sup>1</sup> Department of Electrical Engineering, Linköping University, Linköping, Sweden

<sup>2</sup> Eriksholm Research Centre, Oticon A/S, Snekersten, Denmark

<sup>3</sup> Department of Automatic Control, Lund University, Lund, Sweden

<sup>4</sup> Department of Mathematics and Natural Sciences, Blekinge Institute of Technology, Karlskrona, Sweden

<sup>5</sup> Joint senior authorship.

\*\*Author to whom any correspondence should be addressed.

E-mail: [johanna.wilroth@liu.se](mailto:johanna.wilroth@liu.se)

**Keywords:** auditory attention classification, parallel transport, domain adaptation, EEG, transfer learning, locus of attention

## Abstract

**Objective.** This paper presents a novel domain adaptation (DA) framework to enhance the accuracy of electroencephalography (EEG)-based auditory attention classification, specifically for classifying the direction (left or right) of attended speech. The framework aims to improve the performances for subjects with initially low classification accuracy, overcoming challenges posed by instrumental and human factors. Limited dataset size, variations in EEG data quality due to factors such as noise, electrode misplacement or subjects, and the need for generalization across different trials, conditions and subjects necessitate the use of DA methods. By leveraging DA methods, the framework can learn from one EEG dataset and adapt to another, potentially resulting in more reliable and robust classification models. **Approach.** This paper focuses on investigating a DA method, based on parallel transport, for addressing the auditory attention classification problem. The EEG data utilized in this study originates from an experiment where subjects were instructed to selectively attend to one of the two spatially separated voices presented simultaneously. **Main results.** Significant improvement in classification accuracy was observed when poor data from one subject was transported to the domain of good data from different subjects, as compared to the baseline. The mean classification accuracy for subjects with poor data increased from 45.84% to 67.92%. Specifically, the highest achieved classification accuracy from one subject reached 83.33%, a substantial increase from the baseline accuracy of 43.33%. **Significance.** The findings of our study demonstrate the improved classification performances achieved through the implementation of DA methods. This brings us a step closer to leveraging EEG in neuro-steered hearing devices.

## 1. Introduction

Lacking the capacity to select and enhance a specific sound source of choice and suppress the background, hearing aids generally amplify the volume of everyone in the environment. This presents a significant challenge, as developing computational models that replicate the human brain's ability in effectively suppressing unwanted sounds in noisy environments would

be extremely advanced. Over the last few years, intelligent hearing aids have become better at suppressing background noises (Andersen *et al* 2021). However, the problem of knowing which speaker to enhance, famously known as the *cocktail party problem* (Cherry 1953), is unsolved and most people with hearing aids still experience discomfort in noisy environments (Han *et al* 2019). One solution to this problem is to classify (i.e. decode) attended and unattended sounds from the brain signals, using a group of methods referred to as *auditory attention decoding* (AAD) (Alickovic *et al* 2019, Geirnaert *et al* 2021b).

\* This paper is based on a Master's Thesis work (Wilroth 2020).

Using AAD methods, it was reported that there is a significant difference in the cortical speech representation depending on whether the sound source was attended or unattended (Ding and Simon 2012, Mesgarani and Chang 2012, O'Sullivan *et al* 2015, Mirkovic *et al* 2016). The cortical activity can be detected by several methods, such as invasive intracranial electroencephalography (iEEG) as in (Mesgarani and Chang 2012, Golumbic *et al* 2013) or noninvasive magnetoencephalography as in (Ding and Simon 2012, Akram *et al* 2016) and electroencephalography (EEG; wholes-scalp EEG as in O'Sullivan *et al* 2015, Etard *et al* 2019, Aroudi and Doclo 2020 and in/around-ear EEG as in Mirkovic *et al* 2016, Fiedler *et al* 2017, Nogueira *et al* 2019, Hölle *et al* 2021). However, many advantages of EEG make it the most prevalent method for AAD. EEG is a cheap and widely available technique where the signals are picked up by several small electrodes placed on the head. Unlike some other methods, an EEG recording is able to capture both the radial and the tangential components of the signal, which makes the method effective and accurate. There are however some disadvantages with EEG, mainly that it has limited spatial resolution and low signal-to-noise ratio. The EEG-based AAD has become an important tool in the research for tackling the cocktail party problem in hearing aids (Alickovic *et al* 2020, 2021, Lunner *et al* 2020).

The EEG dataset used in this paper reflects complex, real life situations. Subjects were instructed to attend to one of two different fictional stories, narrated by one female and one male voice, and presented simultaneously. We considered so-called locus-of-attention (LoA) classification, i.e. decoding whether the attended speech was coming from the listener's left or right side from EEG signals. A common approach for distinguishing between attended and unattended sound sources in EEG data is classification using machine learning (ML). Given a training set, the ML algorithm first learns how to classify the data, and it can thereafter be used on new data. Naturally, a large training dataset increases the accuracy of the model and decreases the risk of overfitting. One major issue with classifiers created from EEG recordings is the data shifts occurring between trials and sessions for the same subject. The measurements are also subject dependent, meaning that data shifts also occur when comparing EEG-measurements from different subjects. These shifts are due to factors such as instrument imperfectness (e.g. jitter) or human factors (e.g. misplacement of electrodes), and a classification made for subject A might not work well on data from subject B. Hence, the classification algorithm would need to be constructed from scratch for each new subject. This is time-consuming and not realistic in real-time situations.

The overfitting problem is common to all model estimation methods. ML methods are however often

over-parameterized and thus more sensitive in this regard, and the algorithms tend to learn patterns in the data which could be due to measurement variations or noise. This is common for small training datasets, and it often results in overfitted models that does not generalize well over other datasets (Mutasa *et al* 2020). When working with EEG measurements, the dataset size is constrained by the number of trials each participant is able to carry out without losing focus or altering the measurement setup. Thus, there is a risk of overfitting the model, which compromises its reliability.

Another aspect of a limited dataset is that good data cannot always be obtained from all subjects. The signals can be noisy, electrodes might be misplaced, or the subjects may lack sufficient focus during the attention task. The main focus of this paper is to investigate whether combining poor data from one subject with good data from other subjects can enhance the auditory attention classification performances.

This paper focuses on transfer learning, more specifically a domain adaptation (DA) approach, to answer the research question above. DA is a specific field in ML where the source data distribution (subject A) is different from the target data distribution (subject B) (Weiss *et al* 2016). It has previously been used on EEG data from i.a. motor imagery tasks such as movements of hands, feet and tongue (Yair *et al* 2019, 2020); emotion recognition (Bao *et al* 2021) and working memory (Chen *et al* 2021). However, DA has not yet been used on EEG data to decode (i.e. classify) auditory attention. We primarily focused on the parallel transport (PT) method (Yair *et al* 2019) which is based on covariance matrix computations on the Riemannian manifold of positive definite matrices. The full MATLAB-code for the results is available at GitHub<sup>6</sup>. The main objective of this paper is to evaluate whether a DA method that solely relies on EEG data can be used for AAD to further improve decoding (i.e. classification) performances. The presented problem is LoA with the aim to improve the performances for subjects with initially low classification accuracy, overcoming challenges posed by instrumental and human factors. The findings of our study demonstrate the improved classification performances achieved through the implementation of DA methods.

### 1.1. Paper outline

This paper is structured as follows: section 2 explains Riemannian geometry, and section 3 describes LoA classification with EEG. DA and PT are explained in section 4. The experimental setup, preprocessing of the EEG data, model evaluation method and statistical analysis are discussed in section 5. Results

<sup>6</sup> <https://github.com/JohannaWil/EEG-AAD-DomainAdaptation>

and discussions are presented in section 6. Lastly, section 7 sums up the paper with conclusions. A pipeline from data to classification accuracy is provided in appendix A and a more in depth explanation of PT is presented in appendix B.

## 2. Riemannian geometry

The DA method used in this work relies on covariance matrices and Riemannian geometry. By definition, stated in section 2.1, covariance matrices are symmetric positive definite (SPD) and capture linear relations in the data. The relative easiness with which they can be computed have caught researchers interest when working with complex, high-dimensional datasets such as EEG recordings (Vidyaratne and Iftekharuddin 2017, Yair et al 2019, 2020).

The use of Riemannian geometry will be motivated through a simple example where two points are plotted in  $\mathbb{R}^3$ . Naturally, the minimal Euclidean distance between these two points is easy to compute as a straight line. Now, let each point be a  $2 \times 2$  covariance matrix. Computing the minimal distance between two such SPD matrices has been proven to yield problems when utilizing Euclidean geometry (Yair et al 2019). One drawback is referred to as the swelling effect, where the original determinants are smaller than the Euclidean average determinant (Arsigny et al 2006, Yger et al 2017, Lin 2019). Another problem is due to the non-complete space for SPD matrices when using Euclidean geometry (Fletcher et al 2004). A third issue is the computational approximations, where traditional algorithms for large datasets relying on Euclidean geometry tend to yield unreliable results (Sommer et al 2010).

Mentioned drawbacks might be alleviated using Riemannian geometry. It explores the shapes on a curved space, such as the surface of a cylinder, sphere or cone, and has demonstrated effectiveness when working with SPD matrices such as covariance matrices (Fletcher et al 2004, Arsigny et al 2006, Sommer et al 2010). Again referring to the two covariance matrices plotted as points in  $\mathbb{R}^3$ . Their positivity constraints span a cone manifold in which both points lie strictly inside (Yger et al 2017, Mahadevan et al 2019, Yair et al 2019). The Riemannian distance between the points, further explained in section 2.2, is curved, which has the benefit of reducing the impact of the swelling effect (Arsigny et al 2006, Yger et al 2017).

The effectiveness of using covariance matrices with Riemannian geometry has been established in EEG analysis (Congedo et al 2017, Kalaganis et al 2022) and has previously been applied with success to classify the directional focus of auditory attention in the LoA classification problem (Geirnaert et al 2021a).

### 2.1. Covariance matrices

The covariance matrix  $\mathbf{P}_{s,i} \in \mathbb{R}^{d \times d}$  is defined as:

$$\mathbf{P}_{s,i} = \mathbb{E} \left[ (\mathbf{x}_{s,i}(t) - \mathbb{E}[\mathbf{x}_{s,i}(t)]) (\mathbf{x}_{s,i}(t) - \mathbb{E}[\mathbf{x}_{s,i}(t)])^T \right] \quad (1)$$

where  $\mathbf{x}_{s,i}(t)$  is the recorded EEG time series for subject  $s$  and trial  $i$ . Each trial in the EEG dataset used in this study is structured in a  $t \times d$  matrix, where  $t$  is the number of samples and  $d$  is the number of EEG channels. Each element in the  $d \times d$  covariance matrix  $\mathbf{P}_{s,i}$  describes the covariance between the corresponding channels. Preprocessing of the EEG data is further explained in section 5.2.

### 2.2. Riemannian distance

One way to describe the curvature of the Riemannian manifold  $\mathcal{M}$  is through a so-called sectional curvature. The manifold has a tangent space  $\mathcal{T}_{\mathbf{P}}\mathcal{M}$  at the point  $\mathbf{P} \in \mathcal{M}$ , and the sectional curvature is defined by the point  $\mathbf{P}$  and a two-dimensional subspace of the tangent space. Hence, the sectional curvature depends on two linear independent tangent vectors, and it is therefore possible to view the (symmetric) covariance matrices on the Riemannian manifold as vectors in the Euclidean tangent space. These vectors are used as features in the classification, further explained in section 3. In this paper, the point  $\mathbf{P}$  is a covariance matrix and  $\mathbf{S}$  is the vector in Euclidean space. The shortest path between two covariance matrices  $\mathbf{P}_1, \mathbf{P}_2 \in \mathcal{M}$  on a Riemannian manifold is called a geodesic curve, and it is given by;

$$\varphi(t) = \mathbf{P}_1^{\frac{1}{2}} \left( \mathbf{P}_1^{-\frac{1}{2}} \mathbf{P}_2 \mathbf{P}_1^{-\frac{1}{2}} \right)^t \mathbf{P}_1^{\frac{1}{2}}, \quad 0 \leq t \leq 1. \quad (2)$$

The length of the curve above, referred to as the Riemannian distance, is unique and is given by (Yair et al 2019);

$$\begin{aligned} d_R^2(\mathbf{P}_1, \mathbf{P}_2) &= \left\| \log \left( \mathbf{P}_1^{-\frac{1}{2}} \mathbf{P}_2 \mathbf{P}_1^{-\frac{1}{2}} \right) \right\|_F^2 \\ &= \sum_{i=1}^n \log^2 \left( \lambda_i \left( \mathbf{P}_1^{-\frac{1}{2}} \mathbf{P}_2 \mathbf{P}_1^{-\frac{1}{2}} \right) \right) \end{aligned} \quad (3)$$

where  $\|\cdot\|_F$  is the Frobenius norm,  $\log(\mathbf{P})$  is the matrix logarithm and  $\lambda_i(\mathbf{P})$  is the  $i$ th eigenvalue of  $\mathbf{P}$ .

### 2.3. Riemannian mean

The Riemannian mean  $\mathbf{M}_s$  for subject  $s$ , also referred to as the Fréchet mean, is the sum of distances to the point on the manifold that is closest to all the subject points (Yger et al 2017, Yair et al 2019):

$$\mathbf{M}_s \triangleq \arg \min_{\mathbf{P}_s \in \mathcal{M}} \sum_i d_R^2(\mathbf{P}_s, \mathbf{P}_{s,i}) \quad (4)$$

where  $d_R^2(\mathbf{P}_s, \mathbf{P}_{s,i})$  is the Riemannian distance defined above. In this paper, each subject point is a covariance matrix and the Riemannian mean is therefore a symmetric matrix, which can be interpreted as finding

the center of mass in a high-dimensional Riemannian geometric figure. The Riemannian mean of two covariance matrices  $\mathbf{P}_1, \mathbf{P}_2 \in \mathcal{M}$  is the midpoint  $\varphi(1/2)$  in equation (2). The Riemannian mean for more than two covariance matrices can be computed by an iterative algorithm 1 developed by Barachant *et al*, where  $\text{Log}_{\mathbf{M}}(\mathbf{P}_i)$  and  $\text{Exp}_{\mathbf{M}}(\mathbf{S})$  are defined in (B.1) and (B.2).

---

**Algorithm 1.** The Riemannian mean iterative algorithm for more than two SPD matrices (Barachant *et al* 2013).

---

**Input:** a set of SPD matrices  $\{\mathbf{P}_i \in \mathcal{M}\}_{i=1}^n$  where  $n$  is the number of trials

**Output:** the Riemannian mean matrix  $\mathbf{M}$

1. Compute the initial term  $\mathbf{M} = \frac{1}{n} \sum_{i=1}^n \mathbf{P}_i$
  2. **while**  $\|\mathbf{S}\|_{\text{F}} > \epsilon$ 
    - 2.1. Compute the Euclidean mean in the tangent space  $\mathbf{S} = \frac{1}{n} \sum_{i=1}^n \text{Log}_{\mathbf{M}}(\mathbf{P}_i)$
    - 2.2. Update  $\mathbf{M} = \text{Exp}_{\mathbf{M}}(\mathbf{S})$
  - end while**
- 

### 3. Locus-of-attention classification

The main goal of this study is to accurately decode (i.e. classify) the direction (left or right) of attended speech using EEG data, a technique known as LoA classification. Previous research (Geirnaert *et al* 2020, Cai *et al* 2021, Li *et al* 2021, Vandecappelle *et al* 2021, Su *et al* 2022, Puffay *et al* 2023) has shown promising results in this area. Unlike stimulus reconstruction (SR) methods that use regression techniques (Mesgarani and Chang 2012, O’Sullivan *et al* 2015, Alickovic *et al* 2019, Geirnaert *et al* 2021b) to reconstruct and classify sound stimuli, LoA methods focus on classifying the direction of the attended sound. In this study, we propose using DA to align less reliable data with more reliable data by transporting poor data from one subject to the domain of good data from different subjects, aiming to enhance LoA classification performances.

#### 3.1. Classification methods

This paper focuses on EEG-based locus of auditory attention classification (left vs. right). The classification method involves four steps: (1) computing the covariance matrix, (2) projecting it onto the tangent plane of the Riemannian manifold, (3) vectorizing the covariance matrix to create a feature vector, and (4) employing a linear support vector machine (SVM) for classification. Although we also explored four alternative classification methods—k-nearest neighbor, regression tree, decision tree, and neural network with various configurations—the regression and decision trees yielded unsatisfactory results, and the neural network exhibited a similar accuracy to SVM but with considerably longer training time.

While conventional methods like temporal response functions (TRFs), canonical correlation

analysis (CCA) or match-mismatch are commonly used for classification, our paper explores an alternative approach using vectorized covariance matrices as the data features in the SVM classifier. This choice offers several benefits. First, our approach is audio-free, addressing challenges when audio data is not available. Second, computing covariance matrices is relatively straightforward as they are based on and capture linear relations. Third, past research (Yair *et al* 2019, 2020) demonstrates the effectiveness of SPD matrices (such as covariance matrices), particularly with simpler classifiers like SVMs, which are less complex and computationally more efficient compared to deep neural networks. Covariance matrices have been successfully used as features in diverse fields such as medical imaging, ML and computer vision (Tuzel *et al* 2008, Sra and Hosseini 2013, Freifeld *et al* 2014, Bergmann *et al* 2018). In the context of physiological signal analysis and medical imaging, Riemannian geometry of covariance matrices has been utilized (Pennec *et al* 2004, Barachant *et al* 2013).

### 4. Domain adaptation

In this section, we shortly introduce transfer learning, define domain adaptation (DA) and explain parallel transport (PT).

#### 4.1. Notations

The training data  $D_S$  comes from a *source domain*  $\mathcal{D}_S$  and its predictive learning task is denoted  $\mathcal{T}_S$ . The testing data  $D_T$ , as well as all other data used after the training, comes from a *target domain*  $\mathcal{D}_T$  and its predictive learning task is denoted  $\mathcal{T}_T$ . All notations are presented in table 1 (Weiss *et al* 2016).

#### 4.2. Transportation

In classical ML, the domains and the tasks of the source and target are the same, hence  $\mathcal{D}_S = \mathcal{D}_T$  and  $\mathcal{T}_S = \mathcal{T}_T$ . This is the same as fulfilling the conditions:

$$\mathcal{X}_S = \mathcal{X}_T \text{ same feature space}$$

$$\mathcal{Y}_S = \mathcal{Y}_T \text{ same label space}$$

$$P(\mathbf{X}_S) = P(\mathbf{X}_T) \text{ same marginal distribution}$$

$$P(\mathbf{Y}_S|\mathbf{X}_S) = P(\mathbf{Y}_T|\mathbf{X}_T) \text{ same conditional distribution.}$$

When one or more of these conditions are not satisfied, generalization methods built on a learning-to-learn principle need to be used. This is most commonly referred to as *transfer learning*. The specific case when  $\mathcal{X}_S = \mathcal{X}_T$ ,  $\mathcal{Y}_S = \mathcal{Y}_T$  and, the mismatch between the source and target solely comes from the probability distributions is called DA. This is the situation studied in this paper (Kouw and Loog 2018). The definition of DA presented by (Weiss *et al* 2016) is:



**Table 1.** Notations used in this section (Weiss et al 2016).

Notation	Description	Notation	Description
$\mathcal{X}$	Input feature space	$D_S$	Source domain data
$\mathcal{Y}$	Label space	$D_T$	Target domain data
$\mathcal{T}$	Predictive learning task	$P(X)$	Marginal distribution
Subscript S	Denotes source	$P(Y X)$	Conditional distribution
Subscript T	Denotes target	$X$	Particular learning sample
$\mathcal{D}_S$	Source domain	$x_i$	Feature vector $i$
$\mathcal{D}_T$	Target domain	$y_i$	Class label $i$

**Definition 4.1 (DA).** ‘Given a source feature space  $\mathcal{X}_S$  with corresponding source label space  $\mathcal{Y}_S$  and a target feature space  $\mathcal{X}_T$  with corresponding target label space  $\mathcal{Y}_T$ , DA is the specific case of transfer learning when  $\mathcal{X}_S = \mathcal{X}_T$ ,  $\mathcal{Y}_S = \mathcal{Y}_T$  and the mismatch between source and target comes from  $P(X_S) \neq P(X_T)$  and/or  $P(Y_S|X_S) \neq P(Y_T|X_T)$ .’

After the transportation, the marginal and conditional distributions are merged into the joint distribution  $P(X,Y) = P(Y|X)P(X)$ . In DA, this joint distribution can be broken down into two different cases (Kouw and Loog 2018):

$$\begin{aligned}
 P(X_S) &= P(X_T) \quad \text{and} \\
 P(Y_S|X_S) &\neq P(Y_T|X_T) \quad \text{concept shift} \\
 P(X_S) &\neq P(X_T) \quad \text{and} \\
 P(Y_S|X_S) &= P(Y_T|X_T) \quad \text{covariate shift.}
 \end{aligned}$$

The brain’s electrical pulses, which can be detected by EEG electrodes, are captured as time series data. These time series are typically non-stationary due to factors such as electrode placements, fluctuations in attention levels, environmental influences, eye and motor movements. As a consequence, comparing EEG signals across different trials or sessions often reveals covariate shifts (Razaa et al 2019). In the context of comparing subjects (as in this study), anatomical variations among individuals may introduce differences in the conditional distribution, leading to concept shifts (Albuquerque et al 2019).

### 4.3. Parallel transport

The DA method proposed in this study is grounded in the concept of PT (Yair et al 2019). It specifically addresses the challenge of covariance matrices residing in different regions of the manifold, which commonly arises when data is gathered from multiple subjects and/or sessions. Previous applications of this method in the field of EEG include emotion recognition (Wang et al 2021) and seizure detection and prediction (Peng et al 2022).

PT uses Riemannian geometry to:

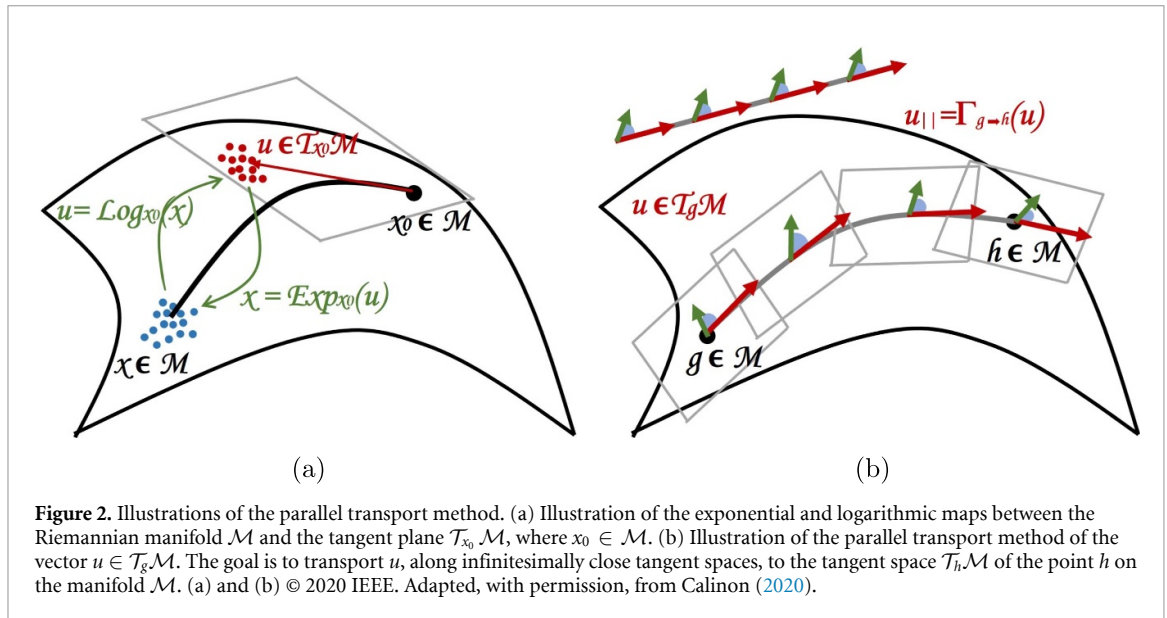
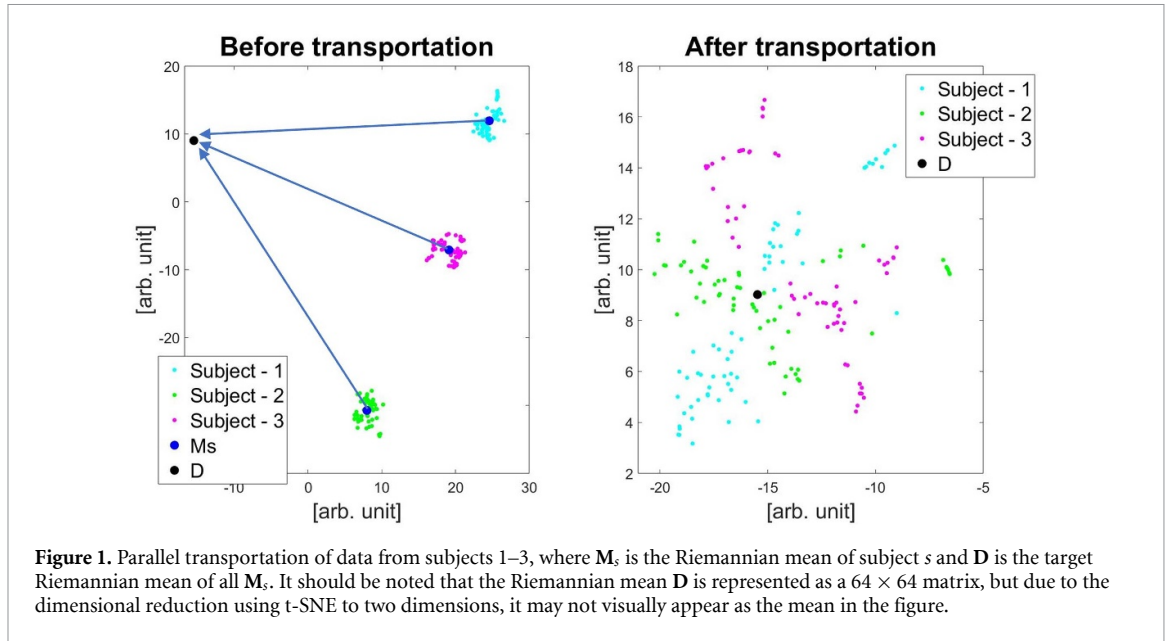
- (i) Compute the Riemannian mean  $\mathbf{M}_s$  of all covariance matrices for each subject  $s$ .
- (ii) Compute the Riemannian mean  $\mathbf{D}$  of all  $\mathbf{M}_s$  from step (i).

- (iii) Project all the covariance matrices from the Riemannian manifold onto a Riemannian tangent plane at  $\mathbf{M}_s$  for each subject  $s$ . This is done with a *logarithm map* which is illustrated in figure 2(a) and further explained in appendix B.
- (iv) Move all the data to  $\mathbf{D}$  using PT, which is illustrated in figure 2(b) and further explained in appendix B.
- (v) Project the covariance matrices from the Riemannian tangent plane back to the Riemannian manifold. This is done with an *exponential map*, which is illustrated in figure 2(a) and further explained in appendix B.
- (vi) Project the covariance matrices to the Euclidean tangent space for classification and plotting purposes. This step is also computed for the baseline algorithms (explained in sections 5.3 and 5.4) which do not use DA.

Visualization plays a significant role in comprehending high-dimensional data, yet condensing such data into two or three dimensions for visualization purposes may result in the loss of significant information. A dimension reduction technique called t-distributed stochastic neighbor embedding (t-SNE) (van der Maaten and Hinton 2008) is used in this paper. The accompanying figure 1 provides an illustration of PT. It should be noted that the Riemannian mean  $\mathbf{D}$  is represented as a  $64 \times 64$  matrix, but due to the dimensional reduction using t-SNE to two dimensions, it may not visually appear as the mean in the figure.

The exponential and logarithm maps are illustrated in figure 2(a). The gray/black line between the two points  $x_0$  and  $x$  on the Riemannian manifold  $\mathcal{M}$  is the minimum length curve.  $u$  is a vector on the tangent plane of  $x_0$ . The exponential map projects  $u$  to a point  $x \in \mathcal{M}$  in the direction of  $u$ . The logarithmic map is the inverse, where the point  $x \in \mathcal{M}$  is projected to the tangent plane  $u \in \mathcal{T}_{x_0}\mathcal{M}$  (Calinon 2020).

Figure 2(b) shows an illustration of the PT method of the vector  $u \in \mathcal{T}_g\mathcal{M}$ . The goal is to transport  $u$ , along infinitesimally close tangent spaces, to the tangent space  $\mathcal{T}_h\mathcal{M}$  of the point  $h$  on the manifold  $\mathcal{M}$ . The black vectors show the direction of the transportation in each tangent space. Using infinitesimally close tangent spaces gives a smooth



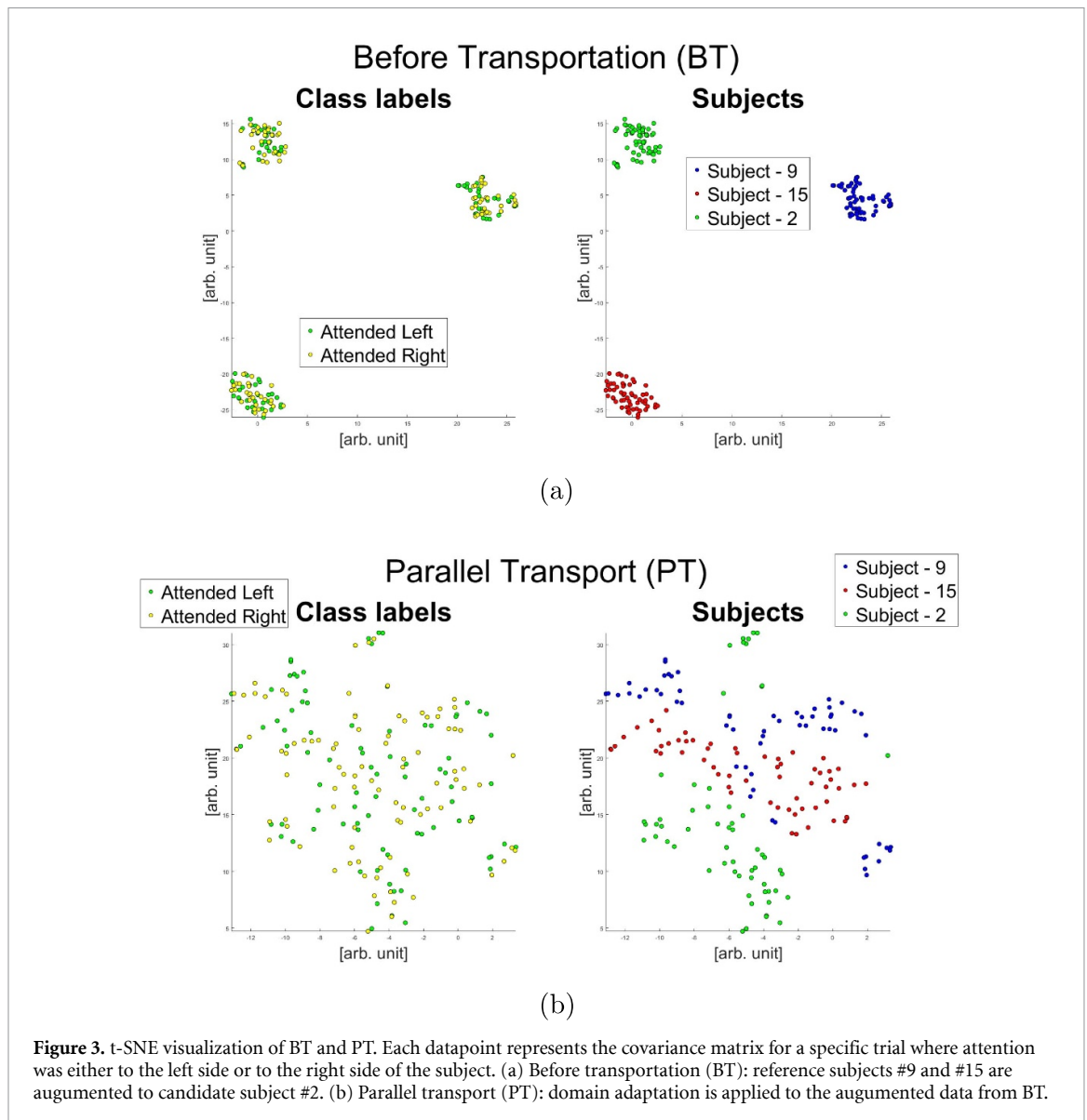
transportation with preserved features of the vector  $u$  (Calinon 2020).

## 5. Experiments

The dataset used in this paper was presented in (Fuglsang et al 2017) and was made publicly available by the authors (Fuglsang et al 2018). The authors state that the written consent according to the Declaration of Helsinki was collected from all subjects and that the Science Ethics Committee for the Capital Region of Denmark has approved the protocol (Wong et al 2018, Fuglsang et al 2017).

### 5.1. Setup

EEG data were collected from a group of 19 participants who were native Danish speakers with normal hearing, aged between 19 and 30 years, and had no reported neurological disorders. The data were recorded using  $d = 64$  scalp electrodes during 60 trials, with a sampling rate of 512 Hz. In each trial  $i$ , participants listened to a pair of competing speech stimuli (one male, one female) presented at 65 dB. The speech streams were narrated by storytellers and recorded at 44.1 kHz in an anechoic chamber. To reflect real life situations, some of these recordings were simulated in a mildly reverberant room and some in a highly reverberant room using the Odeon



**Figure 3.** t-SNE visualization of BT and PT. Each datapoint represents the covariance matrix for a specific trial where attention was either to the left side or to the right side of the subject. (a) Before transportation (BT): reference subjects #9 and #15 are augmented to candidate subject #2. (b) Parallel transport (PT): domain adaptation is applied to the augmented data from BT.

room acoustic modeling software. The three environment scenarios included anechoic, mildly reverberant and highly reverberant. Following each trial, participants answered multiple-choice questions related to the content of the attended speech.

## 5.2. Preprocessing of EEG data

The EEG data was preprocessed through the FieldTrip (Oostenveld *et al* 2011) and COCOHA (Wong *et al* 2018) toolboxes in MATLAB. The procedure included removal of artifacts (eye blinks, muscle movements, heart beats etc), filter out 50 Hz line noise and down-sampling to 64 Hz. The script *preproc\_data.m* which was used for the preprocessing can be downloaded from [zenodo.org](https://zenodo.org) (Wong *et al* 2018, (Fuglsang *et al* 2017)).

In scenarios with several talkers, an increase in oscillatory alpha (frequency band 7–14 Hz) power due to the brain activity of ignoring the unattended speakers (Paul *et al* 2020) has been reported.

This alpha activity is an epiphenomenon of spatial attention, and an additional Butterworth bandpass filter of order 6 with frequency band [1–30] Hz was therefore applied to the EEG data.

## 5.3. Classification strategy

Three different classification tasks are performed; baseline classification, before transportation (BT) and after PT. The last two tasks are illustrated in figure 3. Each datapoint in the figure represents the covariance matrix for a specific trial where attention was either to the left side or to the right side of the subject. All three classification tasks are made on the tangent plane of the Riemannian manifold, hence step 6 in the PT-pipeline (section 4.3) and table B2 is also computed for baseline and BT classification.

The baseline LoA classification accuracy is computed for each subject using leave-one-out cross-validation (LOO CV) with an SVM classifier. This first classification task gives a LOO CV accuracy for



each subject and helps identify *candidate subjects* who exhibit low classification accuracy and may benefit from DA. *Reference subjects*, on the other hand, are identified as the subjects with the highest classification accuracy. Detailed information on the selection of candidate and reference subjects are provided in section 6.1.

Candidate subject data is augmented with data from the reference subjects. Two different approaches are used to evaluate whether data augmentation improves the classification performance. The first approach, referred to as BT, involves temporal concatenation of the reference and candidate subject datasets. This is visualized in figure 3(a), where the candidate subject #2 data is augmented with the data from reference subjects #9 and #15. The second approach utilizes DA via PT. This is used to evaluate the benefits of transfer learning on classification accuracy of the candidate subjects, and is illustrated in figure 3(b). As in the baseline condition, LOO CV with SVM is used for evaluating both BT and PT. Further details on these approaches are provided in section 6.2.

#### 5.4. Model evaluation

The experiment involved 60 trials with two competing talkers (one male and one female) positioned at spatially-separated angles ( $\pm 60^\circ$ ). Each trial lasted for 50 s and correct classification rate was computed using LOO CV.

##### 5.4.1. Baseline classification

For single subject classification (before data augmentation), one trial was designated for testing, while the remaining trials were used for training. This procedure is repeated 60 times, ensuring that each candidate trial was tested once. The final classification accuracy was then determined as the mean of these 60 accuracies.

##### 5.4.2. BT classification

A similar procedure was followed after augmenting the data from reference subjects to the candidate subject. The reference subjects' data served as the training set, and LOO CV was performed exclusively on the candidate subject's data. For instance, augmenting the data of two reference subjects with one candidate subject resulted in a total of 180 trials. Out of these, the 120 trials from the reference subjects were always used as training data, whereas LOO CV was performed on the 60 trials from the candidate subject. As a result, LOO CV was executed 60 times, where each round encompassed one test trial from the candidate subject and 179 training trials derived from both the reference subjects and the candidate subject. The final classification accuracy was then determined by averaging these 60 accuracies.

##### 5.4.3. PT classification

PT was performed with data from all subjects, hence the Riemannian mean was computed for both the reference subjects and the candidate subject. Thereafter, the same classification procedure as for BT was implemented with reference subjects as training data and LOO CV was performed on the 60 trials from the candidate subject. The final classification accuracy was then determined by averaging these 60 accuracies.

#### 5.5. Statistical analysis

To derive the statistically significant threshold (i.e. empirical chance level), the binomial inverse cumulative distribution function (Combrissona and Jerbi 2015) was computed by using the MATLAB function  $x = \text{binoinv}(y, n, p)$ , where  $y = 0.95$  is the significance level,  $n$  is total number of trials and  $p = 0.50$  is the theoretical chance level for binary classification problems. Hence, the result  $x$  is the smallest integer such that the binomial cumulative distribution function (computed at  $x$ ) is equal or greater than  $y$ . Dividing  $p_c = x/n$  gives the level of chance for binary classification problems with  $n$  trials at the 95% confidence interval. The last step is to assign the variable  $f$  to 0 (statistically significant) or to 1 (not statistically significant) by comparing the computed classification accuracy  $Acc$  with  $p_c$ .

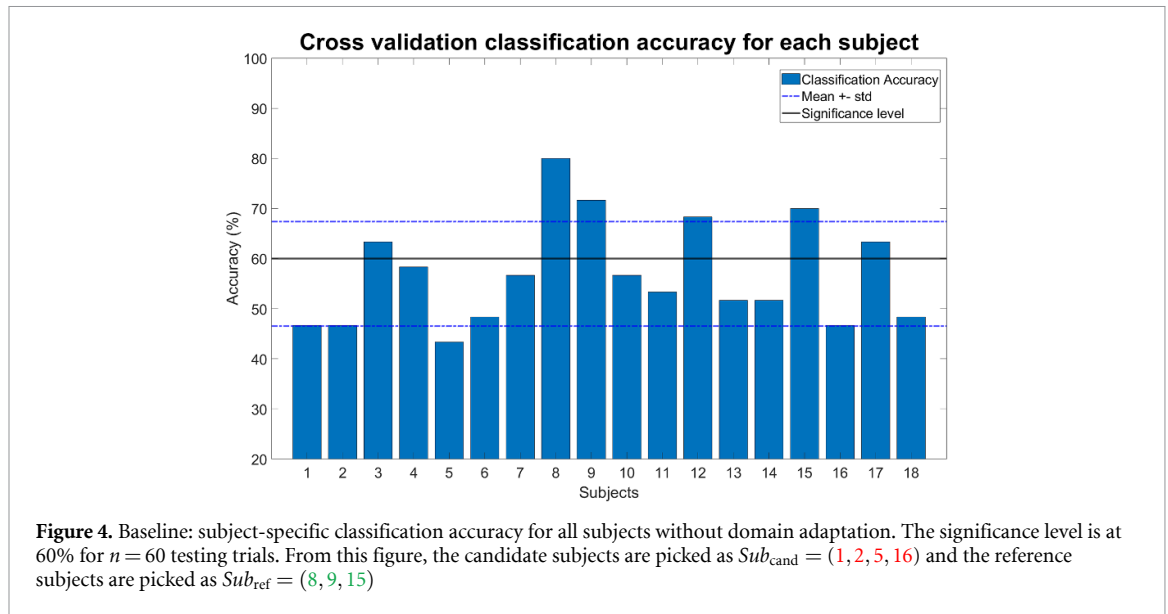
## 6. Results and discussions

First, the baseline classification accuracy for each subject is presented and discussed. This results in the selection of candidate and reference subjects which is used for BT and PT evaluation in section 6.2. Here, the classification accuracies are presented, discussed and the two best combinations of reference subjects are determined. In section 6.3, these two combinations are augmented to data on 12 more subjects to strengthen our results that the use of PT with reference subjects increase the classification accuracy compared to augment data without DA. Finally, potential applications and future work are discussed in section 6.4.

### 6.1. Selection of reference and candidate subjects

The main results in this section show that PT increases the classification accuracy compared to both the baseline and BT. It indicates that only adding more subjects (adding more data) is not enough but a DA method, such as PT, is needed to reach statistically significant results.

Figure 4 shows the LoA classification results for each subject. The level of chance is 60% for a two-classes classification problem with  $n = 60$  trials and a significance level of  $p = 0.05$ . This is computed with the binomial cumulative distribution function (Combrissona and Jerbi 2015). As seen, there is a large variability between subjects and only six of them have a classification accuracy above the significance level,



while the highest value reached 80%. The dashed lines indicate the  $mean \pm standard\ deviation\ (std) = 56.11 \pm 10.92$  of the classification accuracy across all subjects.

The candidate subjects  $Sub_{cand} = (1, 2, 5, 16)$  are selected as subjects with a classification accuracy below or equal to round ( $mean - std$ ). The three subjects with highest classification accuracy are chosen as reference subjects  $Sub_{ref} = (8, 9, 15)$ . The first goal is to increase the classification accuracy of the candidate subjects with data from different combinations of reference subjects, investigated in section 6.2.

Thereafter, we conducted an additional analysis to explore the practical task of classifying LoA in subjects whose EEG data the model had not seen. To tackle this, a two-step analysis was made. First, the best combination of reference subjects was selected. Then, BT and PT classification accuracy was performed for each subject in the dataset and the mean was computed. The second goal of this paper is therefore to investigate if augmentation with the best combination of reference subjects and PT increases the mean classification accuracy over all other subjects. This is investigated and discussed in section 6.3.

## 6.2. Classification performances after reference data augmentation

The classification accuracy rates for the candidate subjects  $Sub_{cand} = (1, 2, 5, 16)$  with different combinations of reference subjects  $Sub_{ref} = (8, 9, 15)$  are shown in figure 5. The solid black line shows the candidate subject baseline.

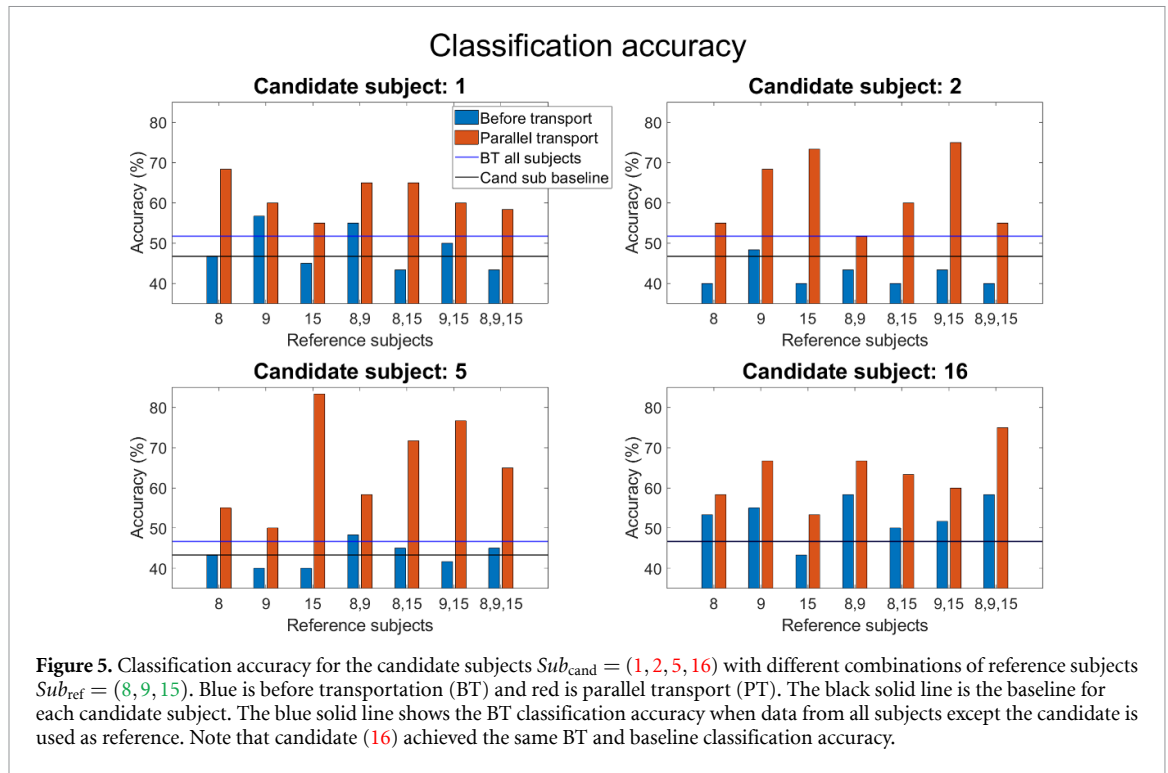
The blue solid line shows the BT classification accuracy when data from all subjects except the candidate is used as reference. Hence, this line shows the result with data from 17 reference subjects and one candidate subject. Candidate (16) achieved the same BT classification accuracy as the subject-specific

baseline classification, resulting in one black line representing both accuracies in the subfigure. The classification accuracy around 50% indicates that only adding data from more subjects does not give any statistically significant results. All four subfigures show the benefit of PT from fewer reference subjects compared to the blue solid line.

In general, BT marginally increased the classification accuracy compared to baseline, indicating that augmenting the data this way has a small impact on the accuracy. This is further outlined with the mean over the candidates shown in table 2, where  $mean_{BT}$  slightly increased the accuracy in all cases except two when compared to  $mean_{baseline} = 45.84\%$ . However, no reference combinations with  $mean_{BT}$  reached above the level of chance. Both figure 5 and table 2 show the benefit of DA. All combinations increased the classification accuracy rates compared to both baseline and BT, although the optimal set of reference subjects highly varies across candidate subjects. In particular, references (15) and (9, 15) significantly benefited from PT, since their mean accuracy increased from baseline of 45.84% to 66.25% and 67.92% respectively with PT. It is important to acknowledge that the obtained accuracy of around 70% for AAD with DA using 50 s trials might seem less favorable when compared to other methodologies such as SR results. While we have omitted the inclusion of other methodologies in this paper, our focus remains on demonstrating how EEG data from subjects with lower classification accuracy can be enhanced to yield improved results, thereby enhancing the overall performance of AAD.

## 6.3. Evaluation of reference subject combinations

Two metrics were studied to evaluate which combination of references induced the best improvement of the candidates:



**Table 2.** The mean of the classification accuracy for the candidates shown in figure 5. The baseline mean for the candidates is 45.84%.

Reference subjects	8	9	15	8,9	8,15	9,15	8,9,15
Mean BT	45.83	50.00	42.08	51.25	44.58	46.67	46.67
Mean PT	59.17	61.25	66.25	60.42	65.00	67.92	63.33

- (i)  $mean(Acc_{PT} - Acc_{BT})$ : the mean over all candidates of the accuracy difference between PT and BT. The desired outcome would be a large positive value, indicating a large benefit with PT.
- (ii)  $mean(Acc_{PT})$ : the mean over all candidates of the PT accuracy, also presented in table 2. This is used to assess the significance of the results.

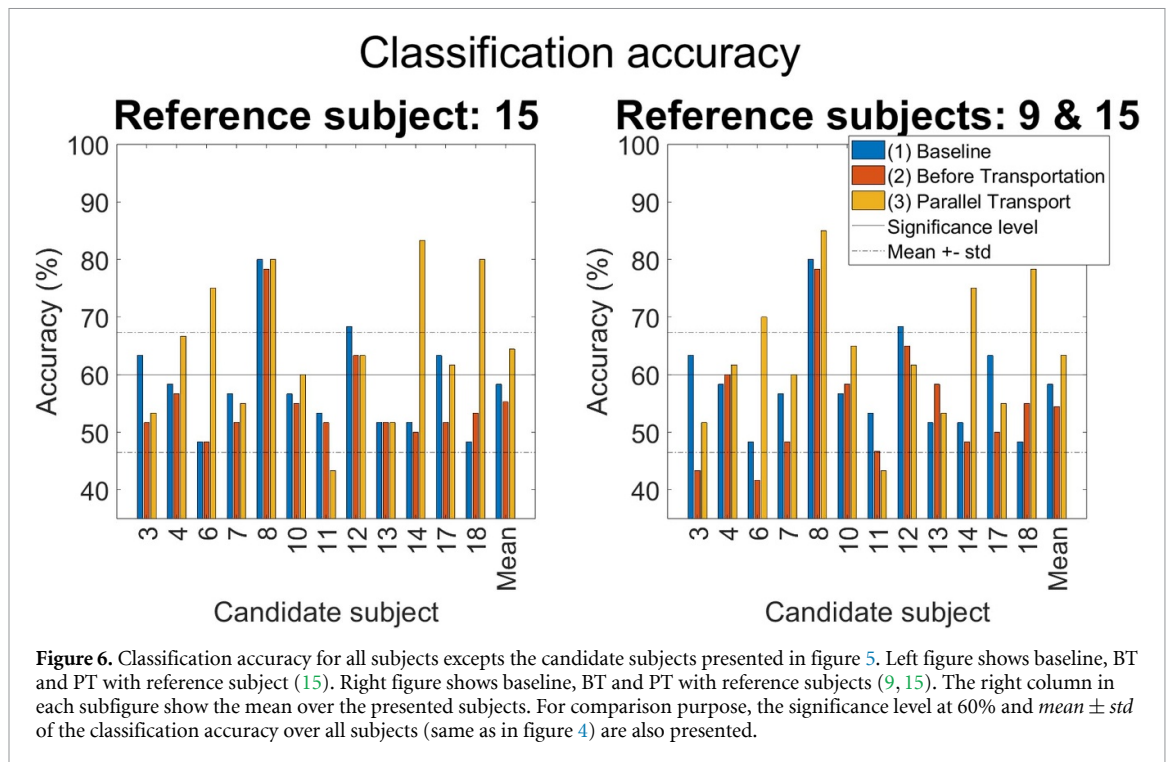
These metrics resulted in (15) and (9,15) as the two best reference combinations. The classification accuracy for these two combinations together with subjects  $\#(3, 4, 6, 7, 8, 10, 11, 12, 13, 14, 17, 18)$  are shown in figure 6. Hence, all subjects in the dataset except the reference subjects and the candidate subjects in figure 5 are presented. Note that also subject #8, which was a reference subject in figure 5, is included. For the sake of comparison, the significance level at 60% and  $mean \pm std$  of the classification accuracy over all subjects (same as in figure 4) are also presented. Subjects #3 and #11 stand out, as the classification accuracy decreases significantly when adding PT. Furthermore, subjects #12 and #17 slightly decrease the classification accuracy compared to baseline. Interestingly, subjects  $\#(3, 12, 17)$  all achieved a baseline accuracy above the significance level and performed worse when adding PT, indicating that higher performing subjects might not benefit as much as low performing subjects when applying

DA. Nevertheless, the high performing subject #8 performed equally well with PT as their baseline classification.

However, eight out of 12 subjects reached an accuracy above or equal to the significance level after PT. In particular, subjects #6, #14 and #18 could greatly benefit from PT as the classification accuracy for these subjects increased from  $48.33\%_{baseline} \rightarrow 75.00\%_{PT,(15)}$ ,  $51.67\%_{baseline} \rightarrow 83.33\%_{PT,(15)}$  and  $51.67\%_{baseline} \rightarrow 80.00\%_{PT,(15)}$  respectively. Again addressing the issue of new label-free datasets, the most interesting metric is the mean over these twelve subjects, which is shown in the last column in each subfigure. The two reference combinations with PT resulted in mean classification accuracy rates of 64.44% and, 63.33%, respectively, both above the significance level.

#### 6.4. Potential applications and future work

EEG holds promising potential as a mean to objectively assess hearing abilities and evaluate the effectiveness of hearing devices. Dealing with subjects who exhibit low classification rates and performance poses significant challenges. However, enhancing EEG data by augmenting suboptimal recordings from one test participant with high-quality EEG data from a different test participant, thereby refining the signal quality. This process may enable the use of EEG



as a reliable tool for objective assessment, overcoming the challenges encountered in such cases.

In order to integrate EEG into hearing devices, there are several challenges that need to be addressed. Among these challenges are the reduction of the number of electrodes and specific investigation of the electrodes close to the ear. If similar results can be obtained with fewer electrodes near the ear, a practical approach could involve training the model using data from well-performing reference subjects. Subsequently, the trained model could be utilized to augment data from candidate subjects with low classification performance. This could then serve as feedback to the hearing aid, enabling real-time adjustment of noise reduction algorithms implemented in hearing devices.

In summary, the practical implication of our results lies in the potential of DA to bring us closer to using EEG in neuro-steered hearing devices. It is crucial to emphasize that significant efforts are required to actualize the development of brain-controlled hearing aids, but the benefits make it an attractive pursuit.

## 7. Conclusions

This study investigated decoding (i.e. classification) of the locus of auditory attention (target on the right or left side) classification from EEG signals using DA based on PT. The primary objective was to mitigate the subject-dependency of classification performances by combining data from candidate subjects, who initially exhibited low classification results, with data from reference subjects, who initially exhibited

high classification results. The results demonstrate that the implementation of PT led to a significant improvement in classification accuracy for the majority of subjects, resulting in a noteworthy increase in mean classification accuracy from 45.84% to 67.92% across the four subjects with poor data. Moreover, the best result achieved by one subject reached 83.33%, marking a notable improvement from the baseline accuracy of 43.33%.

In conclusion, the classification of the locus of auditory attention classification presents a complex problem. The quality of the EEG data is influenced by both human and instrumental factors, including noise, electrode misplacement and subjects not concentrating on the auditory attention task. As a result, subject-dependent datasets are created, where EEG data from one subject may not be sufficient to achieve high classification results. This paper has demonstrated the potential to enhance classification results through DA, wherein data from candidate subjects is augmented with data from reference subjects. Such an approach holds the potential to advance utilization of EEG in neuro-steered hearing devices, bringing us closer to their practical implementation.

## Funding

This work was supported in part by the ELLIIT Strategic Research Area.

## Author contributions

J W, E A, C B and B B contributed to study concept, hypothesis generation, and design of the experiment.

J W analyzed the data. J W, E A, C B, B B and F H interpreted the data. J W drafted the manuscript. J W, C B, F H, E A, B B and M S read the manuscript and provided critical revision. All authors contributed to the article and approved the submitted version.

## Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://doi.org/10.5281/zenodo.1199011>.

## Appendix A. Pipeline from data to classification accuracy

The experimental setup and EEG preprocessing is described in section 5. In summary, EEG data was recorded from 18 subjects and consisted of 60 trials of 50 s. The process from data to classification accuracy utilized in this work can be described through the following pipeline, each step of which will be described in more detail in the following sections.

- (i) Compute the covariance matrices of EEG data for each trial, resulting in 60 covariance matrices per subject, section 2.1. These covariance matrices are used as feature data in the classifier, a strategy which have previously been used in other works (Barachant et al 2013, Yair et al 2019, 2020).
- (ii) Project the covariance matrices from the Riemannian manifold to the Euclidean tangent plane (step 6 in table B2). Compute the subject-specific leave-one-out cross-validation (LOO CV) accuracy with a simple classifier. This classification accuracy will be referred to as the *baseline*. The classification method is described in section 3.1, the model evaluation method in section 5.4 and the statistical analysis in section 5.5.
- (iii) Categorize poor data from subjects with a classification accuracy below or equal to the  $\text{round}(\text{mean} - \text{std})$  of the classification accuracy over all subjects. Std is the standard deviation. These subjects will be referred to as *candidates*, section 6.1.
- (iv) Choose the three subjects with the highest classification accuracy as *references*, section 6.1.
- (v) Augment the reference subjects to each candidate. Project the covariance matrices from the Riemannian manifold to the Euclidean tangent plane (step 6 in table B2). Note that this tangent space is different than the tangent space in (ii) due to the augmented reference subjects. Compute the LOO CV accuracy without domain adaptation. This step will be referred to as *before transportation (BT)*.
- (vi) Apply domain adaptation for each candidate with the same combinations of reference subjects used for BT, section 4.3. Project the covariance matrices from the Riemannian manifold to the Euclidean tangent plane (step 6 in table B2). Note that this tangent space is different than the tangent spaces in (ii) and (v) due to the domain adaptation method. Compute the LOO CV accuracy. This step will be referred to as *parallel transport (PT)*.
- (vii) Evaluate which combinations of references that give the best improvement with PT compared to BT, section 6.3. This is done through two metrics:
  - $\text{mean}(\text{Acc}_{\text{PT}} - \text{Acc}_{\text{BT}})$ : The mean over all candidates of the accuracy difference between PT and BT. The desired outcome would be a large positive value, indicating a large benefit with PT.
  - $\text{mean}(\text{Acc}_{\text{PT}})$ : The mean over all candidates of the PT accuracy. This is used to assess the significance of the results.
- (viii) Evaluate if it is possible to increase the mean accuracy of all non-candidate subjects. These subjects were augmented with the two best reference combinations given in step (vii). The LOO CV accuracy was computed after augmentation (BT) and after parallel transport (PT).

The full MATLAB-code for the results is available at GitHub<sup>6</sup>.

## Appendix B. Parallel transport

Table B1 shows all notations used in sections 2 and 4.3.

### B.1. Parallel transport

This section gives some more in depth explanation of parallel transport, and more information of the method can be found in the article by Yair et al (2019).

The logarithm map  $\mathbf{S}_{s,i}$  and exponential map  $\mathbf{P}_{s,i}$  for subject  $s$  and trial  $i$  is defined as:

$$\mathbf{S}_{s,i} = \text{Log}_{\mathbf{P}}(\mathbf{P}_{s,i}) = \mathbf{P}^{\frac{1}{2}} \log\left(\mathbf{P}^{-\frac{1}{2}} \mathbf{P}_{s,i} \mathbf{P}^{-\frac{1}{2}}\right) \mathbf{P}^{\frac{1}{2}} \in \mathcal{T}_{\mathbf{P}} \mathcal{M} \quad (\text{B.1})$$

$$\mathbf{P}_{s,i} = \text{Exp}_{\mathbf{P}}(\mathbf{S}_{s,i}) = \mathbf{P}^{\frac{1}{2}} \exp\left(\mathbf{P}^{-\frac{1}{2}} \mathbf{S}_{s,i} \mathbf{P}^{-\frac{1}{2}}\right) \mathbf{P}^{\frac{1}{2}} \in \mathcal{M} \quad (\text{B.2})$$

where  $\mathbf{S}_{s,i}$  is a symmetric matrix. Table B2 and algorithm 2 shows the MATLAB pseudocode for the six steps of the parallel transport method outlined in section 4.3.



**Table B1.** Notations used in sections 2 and 4.3 (Yair *et al* 2019).

Notation	Description
$n$	Number of trials
$d$	Number of EEG channels and audio-files
$N$	Number of subjects
$\mathcal{M}$	The Riemannian manifold
$\mathbf{P}_{s,i}$	Covariance matrix $\in \mathcal{M}$ for trial $i$ and subject $s$
$\mathcal{T}_{\mathbf{P}}\mathcal{M}$	The tangent plane of the manifold $\mathcal{M}$ at the symmetric matrix $\mathbf{P}$
$\mathbf{M}_s$	Riemannian mean of $\mathbf{P}_{s,i}, \forall i$ , for subject $s$
$\mathbf{D}$	Riemannian mean of $\mathbf{M}_s, \forall s$
$\mathbf{S}_{s,i}$	Symmetric matrix $\in \mathcal{T}_{\mathbf{P}}\mathcal{M}$ for trial $i$ and subject $s$

**Table B2.** The parallel transport steps.

Step	Description	MATLAB
1.	Compute the Riemannian means $\mathbf{M}_s, \forall s$	Algorithm 1, section 2
2.	Compute the Riemannian mean $\mathbf{D}$ of all $\mathbf{M}_s$	Algorithm 1, section 2
3.	Project all the covariance matrices $\mathbf{P}_{s,i}$ from the manifold $\mathcal{M}$ to the tangent plane $\mathcal{T}_{\mathbf{M}_s}\mathcal{M}$	$\mathbf{S}_{s,i}^{\mathbf{M}_s} = \text{Log}_{\mathbf{M}_s}(\mathbf{P}_{s,i}^{\mathbf{M}_s})$ , (B.1)
4.	Move all the data to $\mathbf{D}$	$\mathbf{S}_i^{\mathbf{D}} = \Gamma_{\mathbf{M}_s \rightarrow \mathbf{D}}(\mathbf{S}_{s,i}^{\mathbf{M}_s})$
5.	Project the symmetric matrices $\mathbf{S}_{s,i}$ back to the manifold $\mathcal{M}$	$\mathbf{P}_i^{\mathbf{D}} = \text{Exp}_{\mathbf{D}}(\mathbf{S}_i^{\mathbf{D}})$ , (B.2)
6.	Project the covariance matrices to the Euclidean tangent space	$\tilde{\mathbf{S}}_i^{\mathbf{D}} = \log(\mathbf{D}^{-\frac{1}{2}} \mathbf{P}_i^{\mathbf{D}} \mathbf{D}^{-\frac{1}{2}})$

**Algorithm 2.** Domain adaptation using parallel transport for SPD matrices (Yair *et al* 2019).

**Input:**  $\{\mathbf{P}_{1,i}\}_{i=1}^n, \dots, \{\mathbf{P}_{s,i}\}_{i=1}^n, \dots, \{\mathbf{P}_{N,i}\}_{i=1}^n$  where  $\mathbf{P}_{s,i}$  is the covariance matrix for subject  $s$  and trial  $i$ .

**Output:**  $\{\tilde{\mathbf{S}}_{s,i}\}_{i=1}^n, \dots, \{\tilde{\mathbf{S}}_{s,i}\}_{i=1}^n, \dots, \{\tilde{\mathbf{S}}_{N,i}\}_{i=1}^n$  where  $\tilde{\mathbf{S}}_{s,i}$  is the new representation of  $\mathbf{P}_{s,i}$  in a Euclidean space.

1. For each  $i \in \{1, 2, \dots, n\}$ , compute the Riemannian mean  $\mathbf{M}_s$  of the subset  $\{\mathbf{P}_{s,i}\}$
2. Compute  $\mathbf{D}$ , the Riemannian mean of  $\{\mathbf{M}_s\}_{s=1}^N$
- 3–5. For all  $s$  and  $i$ , apply projection and parallel transport using equation (B.3):

$$\mathbf{P}_{s,i}^{\mathbf{D}} = \mathbf{E} \mathbf{P}_{s,i}^{\mathbf{M}_s} \mathbf{E}^T, \quad \mathbf{E} = (\mathbf{D} \mathbf{M}_s^{-1})^{\frac{1}{2}}$$

6. For all  $i$ , project the transported matrix to the tangent space via:

$$\tilde{\mathbf{S}}_{s,i}^{\mathbf{D}} = \log(\mathbf{D}^{-\frac{1}{2}} \mathbf{P}_{s,i}^{\mathbf{D}} \mathbf{D}^{-\frac{1}{2}})$$

Steps 3–5 can be combined, which gives the projection to the tangent plane, transportation along the tangent planes and projection back to the manifold in one equation:

$$\mathbf{P}_{s,i}^{\mathbf{D}} = \text{Exp}_{\mathbf{D}}\left(\Gamma_{\mathbf{M}_s \rightarrow \mathbf{D}}\left(\text{Log}_{\mathbf{D}}\left(\mathbf{P}_{s,i}^{\mathbf{M}_s}\right)\right)\right) = \mathbf{E} \mathbf{P}_{s,i}^{\mathbf{M}_s} \mathbf{E}^T, \quad (\text{B.3})$$

where  $\mathbf{E} = (\mathbf{D} \mathbf{M}_s^{-1})^{\frac{1}{2}}$  (Yair *et al* 2019). Step 6 is also computed for the baseline algorithms (explained

in sections 5.3 and 5.4) which do not use domain adaptation.

## ORCID iDs

Johanna Wilroth  <https://orcid.org/0000-0002-1762-7454>

Bo Bernhardsson  <https://orcid.org/0000-0002-6629-1889>

Frida Heskebeck  <https://orcid.org/0000-0002-6831-5119>

Martin A Skoglund  <https://orcid.org/0000-0001-9183-3427>

Carolina Bergeling  <https://orcid.org/0000-0002-8660-0263>

Emina Alickovic  <https://orcid.org/0000-0002-4655-9112>

## References

- Akram S, Presacco A, Simon J Z, Shamma S A and Babadi B 2016 Robust decoding of selective auditory attention from MEG in a competing-speaker environment via state-space modeling *NeuroImage* **124** 906–17
- Albuquerque I, Monteiro J, Rosanne O, Tiwari A, Gagnon J F and Falk T H 2019 Cross-subject statistical shift estimation for generalized electroencephalography-based mental workload assessment (arXiv:1906.08823)
- Alickovic E, Lunner T, Gustafsson F and Ljung L 2019 A tutorial on auditory attention identification methods *Front. Neurosci.* **13** 153
- Alickovic E, Lunner T, Wendt D, Fiedler L, Hietkamp R, Ng E H N and Graversen C 2020 Neural representation enhanced for

- speech and reduced for background noise with a hearing aid noise reduction scheme during a selective attention task *Front. Neurosci.* **14** 846
- Alickovic E, Ng E H N, Fiedler L, Santurette S, Innes-Brown H and Graversen C 2021 Effects of hearing aid noise reduction on early and late cortical representations of competing talkers in noise *Front. Neurosci.* **15** 636060
- Andersen A H, Santurette S, Pedersen M S, Alickovic E, Fiedler L, Jensen J and Behrens T 2021 Creating clarity in noisy environments by using deep learning in hearing aids *Seminars in Hearing* (Thieme Medical Publishers, Inc.) pp 260–81
- Aroudi A and Doclo S 2020 Cognitive-driven binaural beamforming using EEG-based auditory attention decoding *IEEE/ACM Trans. Audio Speech Lang. Process.* **28** 862–75
- Arsigny V, Fillard P, Pennec X and Ayache N 2006 Geometric means in a novel vector space structure on symmetric positive-definite matrices *SIAM J. Matrix Anal. Appl.* **29** 328–47
- Bao G, Zhuang N, Tong L, Yan B, Shu J, Wang L, Zeng Y and Shen Z 2021 Two-level domain adaptation neural network for EEG-based emotion recognition *Front. Hum. Neurosci.* **14** 605246
- Barachant A, Bonnet S, Congedo M and Jutten C 2013 Classification of covariance matrices using a Riemannian-based kernel for BCI applications *Neurocomputing* **112** 172–8
- Bergmann R, Fitschen J, Persch J and Steidl G 2018 Priors with coupled first and second order differences for manifold-valued image processing *J. Math. Imaging Vis.* **60** 1459–81
- Cai S, Li P, Su E and Xie L 2021 Auditory attention detection via cross-modal attention *Front. Neurosci.* **15** 652058
- Calinon S 2020 Gaussians on Riemannian manifolds: applications for robot learning and adaptive control *IEEE Robot. Autom. Mag.* **27** 33–45
- Chen J, Chen Y and Bi W 2021 Cross-subject domain adaptation for multi-frame EEG images (arXiv:2106.06769)
- Cherry E C 1953 Some experiments on the recognition of speech, with one and with two ears *J. Acoust. Soc. Am.* **25** 975–9
- Combrissona E and Jerbi K 2015 Exceeding chance level by chance: the caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy *J. Neurosci. Methods* **250** 126–36
- Congedo M, Barachant A and Bhatia R 2017 Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review *Brain-Comput. Interfaces* **4** 155–74
- Ding N and Simon J Z 2012 Emergence of neural encoding of auditory objects while listening to competing speakers *Proc. Natl Acad. Sci.* **109** 11854–9
- Etard O, Kegler M, Braiman C, Forte A E and Reichenbach T 2019 Decoding of selective attention to continuous speech from the human auditory brainstem response *NeuroImage* **200** 1–11
- Fiedler L, Wöstmann M, Graversen C, Brandmeyer A, Lunner T and Obleser J 2017 Single-channel in-ear-EEG detects the focus of auditory attention to concurrent tone streams and mixed speech *J. Neural Eng.* **14** 036020
- Fletcher P, Lu C, Pizer S and Joshi S 2004 Principal geodesic analysis for the study of nonlinear statistics of shape *IEEE Trans. Med. Imaging* **23** 995–1005
- Freifeld O, Hauberg S and Black M 2014 Model transport: towards scalable transfer learning on manifolds *2014 IEEE Conf. on Computer Vision and Pattern Recognition* pp 1378–85
- Fuglsang S A, Dau T and Hjortkjær J 2017 Noise-robust cortical tracking of attended speech in real-world acoustic *NeuroImage* **156** 435–44
- Fuglsang S R A, Wong D and Hjortkjær J 2018 EEG and audio dataset for auditory attention decoding *Zenodo* <https://doi.org/10.5281/zenodo.1199011>
- Geirnaert S, Francart T and Bertrand A 2020 Fast EEG-based decoding of the directional focus of auditory attention using common spatial patterns *IEEE Trans. Biomed. Eng.* **68** 1557–68
- Geirnaert S, Francart T and Bertrand A 2021a Riemannian geometry-based decoding of the directional focus of auditory attention using EEG *ICASSP 2021–2021 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE) pp 1115–9
- Geirnaert S, Vandecappelle S, Alickovic E, de Cheveigne A, Lalor E, Meyer B T, Miran S, Francart T and Bertrand A 2021b Electroencephalography-based auditory attention decoding: toward neurosteered hearing devices *IEEE Signal Process. Mag.* **38** 89–102
- Golumbic E M Z et al 2013 Mechanisms underlying selective neuronal tracking of attended speech at a “cocktail party” *Neuron* **77** 980–91
- Han C, O’Sullivan J, Luo Y, Herrero J, Mehta A D and Mesgarani N 2019 Speaker-independent auditory attention decoding without access to clean speech sources *Sci. Adv.* **5** eaav6134
- Hölle D, Meekes J and Bleichner M G 2021 Mobile ear-EEG to study auditory attention in everyday life *Behav. Res. Methods* **53** 2025–36
- Kalaganis F P, Laskaris N A, Oikonomou V P, Nikopolopoulos S and Kompatsiaris I 2022 Revisiting Riemannian geometry-based EEG decoding through approximate joint diagonalization *J. Neural Eng.* **19** 066030
- Kouw W M and Loog M 2018 Technical report: an introduction to domain adaptation and transfer learning (arXiv:1812.11806)
- Li P, Cai S, Su E and Xie L 2021 A biologically inspired attention network for EEG-based auditory attention detection *IEEE Signal Process. Lett.* **29** 284–8
- Lin Z 2019 Riemannian geometry of symmetric positive definite matrices via Cholesky decomposition *SIAM J. Matrix Anal. Appl.* **40** 1353–70
- Lunner T, Alickovic E, Graversen C, Ng E H N, Wendt D and Keidser G 2020 Three new outcome measures that tap into cognitive processes required for real-life communication *Ear Hear.* **41** 395
- Mahadevan S, Mishra B and Ghosh S 2019 A unified framework for domain adaptation using metric learning on manifolds *Machine Learning and Knowledge Discovery in Databases* ed M Berlingerio, F Bonchi, T Gärtner, N Hurley and G Ifrim (Springer International Publishing) pp 843–60
- Mesgarani N and Chang E F 2012 Selective cortical representation of attended speaker in multi-talker speech perception *Nature* **485** 233–6
- Mirkovic B, Bleichner M G, De Vos M and Debener S 2016 Target speaker detection with concealed EEG around the ear *Front. Neurosci.* **10** 349
- Mutasa S, Sun S and Ha R 2020 Understanding artificial intelligence based radiology studies: what is overfitting? *Clin. Imaging* **65** 96–99
- Nogueira W, Dolhopiatenko H, Schierholz I, Büchner A, Mirkovic B, Bleichner M G and Debener S 2019 Decoding selective attention in normal hearing listeners and bilateral cochlear implant users with concealed ear EEG *Front. Neurosci.* **13** 720
- O’Sullivan J A, Power A J, Mesgarani N, Rajaram S, Foxe J J, Shinn-Cunningham B G, Slaney M, Shamma S A and Lalor E C 2015 Attentional selection in a cocktail party environment can be decoded from single-trial EEG *Cereb. Cortex* **25** 1697–706
- Oostenveld R, Fries P, Maris E and Schoffelen J-M 2011 FieldTrip: open source software for advanced analysis of MEG, EEG and invasive electrophysiological data *Comput. Intell. Neurosci.* **2011** 1–9
- Paul B T, Uzelac M, Chan E and Dimitrijevic A 2020 Poor early cortical differentiation of speech predicts perceptual difficulties of severely hearing-impaired listeners in multi-talker environments *Sci. Rep.* **10** 6141
- Peng P, Xie L, Zhang K, Zhang J, Yang L and Wei H 2022 Domain adaptation for epileptic EEG classification using adversarial

- learning and riemannian manifold *Biomed. Signal Process. Control* **75** 103555
- Pennec X, Fillard P, Ayache N and Epidaure P 2004 A Riemannian framework for tensor computing *Int. J. Comput. Vis.* **66** 41–66
- Puffay C et al 2023 Relating EEG to continuous speech using deep neural networks: a review (arXiv:2302.01736)
- Razaa H, Ratheeb D, Zhouc S-M, Cecottid H and Prasadb G 2019 Covariate shift estimation based adaptive ensemble learning for handling non-stationarity in motor imagery related EEG-based brain-computer interface *Neurocomputing* **343** 154–66
- Sommer S, Lauze F, Hauberg S and Nielsen M 2010 Manifold valued statistics, exact principal geodesic analysis and the effect of linear approximations *Computer Vision—ECCV 2010* pp 43–56
- Sra S and Hosseini R 2013 Conic geometric optimization on the manifold of positive definite matrices *SIAM J. Optim.* **25** 713–39
- Su E, Cai S, Xie L, Li H and Schultz T 2022 STAnet: a spatiotemporal attention network for decoding auditory spatial attention from EEG *IEEE Trans. Biomed. Eng.* **69** 2233–42
- Tuzel O, Porikli F and Meer P 2008 Pedestrian detection via classification on Riemannian manifolds *IEEE Trans. Pattern Anal. Mach. Intell.* **30** 1713–27
- van der Maaten L and Hinton G 2008 Visualizing data using t-SNE *J. Mach. Learn. Res.* **9** 2579–605 (available at: <https://jmlr.org/papers/v9/vandermaaten08a.html>)
- Vandecappelle S, Deckers L, Das N, Ansari A H, Bertrand A and Francart T 2021 EEG-based detection of the locus of auditory attention with convolutional neural networks *eLife* **10** e56481
- Vidyaratne L S and Iftekharuddin K M 2017 Real-time epileptic seizure detection using EEG *IEEE Trans. Neural Syst. Rehabil. Eng.* **25** 2146–56
- Wang Y, Qiu S, Ma X and He H 2021 A prototype-based SPD matrix network for domain adaptation EEG emotion recognition *Pattern Recognit.* **110** 107626
- Weiss K, Khoshgoftaar T M and Wang D 2016 A survey of transfer learning *J. Big Data* **3** 9
- Wilroth J 2020 Domain adaptation for attention steering *Master's Thesis* (available at: <http://lup.lub.lu.se/student-papers/record/9024799>)
- Wong D, Fuglsang S A, Hjortkjær J, Ceolini E, Slaney M and de Cheveigné A 2018 A comparison of temporal response function estimation methods for auditory attention decoding *Front. Neurosci.* **12** 531
- Yair O, Ben-Chen M and Talmon R 2019 Parallel transport on the cone manifold of SPD matrices for domain adaptation *IEEE Trans. Signal Process.* **67** 1797–811
- Yair O, Dietrich F, Talmon R and Kevrekidis I G 2020 Optimal transport on the manifold of SPD matrices for domain adaptation (arXiv:1906.00616 [cs.LG])
- Yger F, Berar M and Lotte F 2017 Riemannian approaches in brain-computer interfaces: a review *IEEE Trans. Neural Syst. Rehabil. Eng.* **25** 1753–62