Align and Pool for EEG Headset Domain Adaptation (ALPHA) to Facilitate Dry Electrode Based SSVEP-BCI

Bingchuan Liu, *Student Member, IEEE*, Xiaogang Chen, *Member, IEEE*, Xiang Li, *Student Member, IEEE*, Yijun Wang, *Member, IEEE*, Xiaorong Gao, *Member, IEEE*, and Shangkai Gao, *Fellow, IEEE*

***Abstract*—Objective: The steady-state visual evoked potential based brain-computer interface (SSVEP-BCI) implemented in dry electrodes is a promising paradigm for alternative and augmentative communication in real-world applications. To improve its performance and reduce the calibration effort for dry-electrode systems, we utilize cross-device transfer learning by exploiting auxiliary individual wet-electrode electroencephalogram (EEG). Methods: We proposed a novel transfer learning framework named ALign and Pool for EEG Headset domain Adaptation (ALPHA), which aligns the spatial pattern and the covariance for domain adaptation. To evaluate its efficacy, 75 subjects performed an experiment of 2 sessions involving a 12-target SSVEP-BCI task. Results: ALPHA significantly outperformed a baseline approach (canonical correlation analysis, CCA) and two competing transfer learning approaches (transfer template CCA, ttCCA and least square transformation, LST) in two transfer directions. When transferring from wet to dry EEG headsets, ALPHA significantly outperformed the fully-calibrated approach of task-related component analysis (TRCA). Conclusion: ALPHA advances the frontier of recalibration-free cross-device transfer learning for SSVEP-BCIs and boosts the performance of dry electrode based systems. Significance: ALPHA has methodological and practical implications and pushes the boundary of dry electrode based SSVEP-BCI toward real-world applications.**

***Index Terms*—Brain-computer interface, steady-state visual evoked potential, electroencephalography, transfer learning, domain adaptation, subspace learning.**

Manuscript received February 27, 2021; revised June 11, 2021; accepted August 4, 2021. Date of publication August 18, 2021; date of current version January 20, 2022. The work was supported in part by Key Research and Development Program of Guangdong Province under Grant 2018B030339001, in part by Doctoral Brain+X Seed Grant Program of Tsinghua University, in part by Strategic Priority Research Program of Chinese Academy of Science under Grant XDB32040200, in part by the Beijing Science and Technology Program under Grant Z201100004420015, in part by the National Key Research and Development Program of China under Grant 2017YFB1002505, and in part by the National Natural Science Foundation of China under Grant 61431007. *(Corresponding author: Xiaorong Gao and Yijun Wang.)*

Bingchuan Liu, Xiang Li, and Shangkai Gao are with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, China.

Xiaogang Chen is with the Institute of Biomedical Engineering, Chinese Academy of Medical Sciences and Peking Union Medical College, China.

Yijun Wang is with the State Key Laboratory on Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China (e-mail: wangyj@semi.ac.cn).

Xiaorong Gao is with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China (e-mail: gxr-dea@mail.tsinghua.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TBME.2021.3105331, provided by the authors.

Digital Object Identifier 10.1109/TBME.2021.3105331

I. Introduction

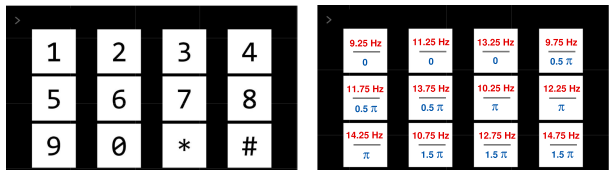
THE brain-computer interface (BCI) provides an alternative pathway between the brain and external devices, and this pathway obviates the need for the involvement of the peripheral nervous system by leveraging brain signals from sensation, perception and higher-level cognition [1], [2]. Compared with its invasive counterpart, noninvasive BCI boasts its safety and ease of use, which has the potential for wide applicability in able-bodied users. Among the noninvasive paradigms, steady-state visual evoked potential based BCI (SSVEP-BCI) has received increasing attention and constitutes one of the dominant paradigms to date [3]. Physiologically, SSVEP is elicited by periodic visual stimuli, e.g., flickers in a visual speller, and its frequency tagging attribute provides it with a relatively high signal-to-noise ratio (SNR). The merit of a high SNR makes it possible to develop a BCI system with a high information transfer rate (ITR) for practical applications. Over the past decades, continuous efforts have been devoted toward the goal, and one line of research is the development of frequency recognition methods [4]–[8]. For example, the training-free filter bank canonical correlation analysis (FBCCA) method [6] and the training-based task-related component analysis (TRCA) method [8] have led to remarkable improvements in ITR.

In parallel, another line of research works on the instrumentation and implementation of SSVEP-BCI for real-world applications, especially wearable systems based on dry EEG headsets [9]–[12]. For example, a single-channel dry EEG headset was developed by NeuroSky Inc. This consumer-grade user-friendly SSVEP-BCI system could identify 4 targets with an ITR of 34.3 bits per minute (bpm) [9]. In an industry monitoring task, a low-cost wearable and wireless SSVEP-BCI system was integrated into augmented reality (AR) glasses, which could achieve an average accuracy of 81.1% at 2 s using single-channel EEG [12]. In these applications, since it removes the hassle of the

gel setup by professionals and gel cleaning after use, the dry EEG headset would be better suited for the general public in daily use. However, probably due to its high impedance, unstable contact and susceptibility to artifacts, the SNR and ITR metrics from the dry EEG headset are generally lower than those from the wet EEG headset [10], [11]. In practical BCI applications, e.g., a BCI speller, the degradation in performance constitutes a standing problem for its widespread use. This raises the question of how to enhance the performance of dry electrode based SSVEP-BCI systems in practical applications.

Transfer learning [13], or specifically domain adaptation [13] has the potential to address the issue by boosting the system performance with auxiliary source data. As a special case of general transfer learning, domain adaptation transfers knowledge across different domains within the same task by means of labeled data in the source domain and unlabeled data in the target domain [13]. The concept of transfer learning has been explored in SSVEP-BCI in recent years [14]–[21]. In cross-subject transfer learning, subjects are assumed to share a common SSVEP template [14], embedding [17], [20], [21] or spatial filter [16]. SSVEPs from the source subjects are exploited in this fashion to enhance the performance of a target subject. For example, in the transfer template canonical correlation analysis (ttCCA) [14], the averaged EEG from the source subjects were used as the SSVEP template, but the difference between the source subjects and target subjects was neglected. This difference was taken into account in the least square transformation (LST) method, which minimized the error between the target SSVEP and weighted source SSVEPs [17], [20], [21]. In the cross-stimulus scenario, the spatial filter was transferred across stimuli [19], [20], or a target SSVEP template was manually synthesized [15] to reduce the number of calibration trials. In cross-device transfer learning, the SSVEP data from the source EEG headset are used to facilitate calibration for the target EEG headset [18]. To our knowledge, the existing literature in this area [18], [21] is relatively scarce so far compared with the other transfer learning scenarios. In the prior work of a recalibration-free setting, the LST method was employed to transform the source SSVEP to the target domain, but there still existed a large gap between the performance of the fully-calibrated method and the transfer learning method [18].

To bridge this gap, we utilize cross-device transfer learning to facilitate the performance of dry electrode based SSVEP-BCI by means of auxiliary data from the same subject. We use the wet-electrode EEG as the auxiliary data considering that it is a gold standard in the laboratory setting and it generally has a higher SNR [10], [11], which possibly yields more benefits in transfer learning. The goal of the SSVEP-BCI system is recalibration-free [18], where there is no need for subjects to recalibrate the system after wearing a dry EEG headset. To fulfill this goal, we propose a novel transfer learning framework in this work. Compared with the conventional pipeline of frequency recognition, the proposed framework features an intermediate step of subspace alignment for domain adaptation, which minimizes the discrepancy between the source domains and target domains. To harness the benefit of auxiliary data and avoid negative transfer [13], we propose new approaches



(A) (B)

Fig. 1. The virtual dialing keyboard for the 12-target SSVEP-BCI speller. (**A**) The layout of a dialing keyboard with 10 digits (0–9) and 2 punctuations (* and #). (**B**) The stimulus frequency and initial phase corresponding to each target.

intended for SSVEP-BCI to align the statistics of spatial patterns in tandem with covariance. A novel subspace pooling strategy is also proposed to aggregate the discriminative detection statistics from more than the conventional first subspace. Since it aligns and pools as its crucial steps, the proposed framework is named **AL**ign and **P**ool for EEG **H**eadset domain **A**daptation (ALPHA) in the present study. The feasibility and efficacy of ALPHA are demonstrated by experiments with a large number of subjects.

II. Materials and Methods

A. Boosting Protocol for a Dry Electrode Based System

1) Subjects: Seventy-five healthy volunteers (49 males and 26 females, mean age 30 ± 0.91 years) participated in this study. Only one subject was left-handed and the rest were right-handed. 49 subjects also took part in the study in [22]. All the subjects had a normal or corrected-to-normal vision. This study was approved by the institutional review board of Tsinghua University (NO.20200020), and subjects gave written informed consent before the experiment.

2) EEG Headsets: Two typical types of EEG headsets, i.e., EEG headsets with dry or wet electrodes were adopted in the study. The details of the dry multipin electrodes can be found in [11]. The wet-electrode headset used conventional Ag/AgCl electrodes. Both headsets were designated for SSVEP-BCI tasks, and the electrodes were located at the 8 channels of the international 10-20 system, i.e., POz/Oz, PO3/4, PO5/6, O1/O2, together with reference and ground electrodes on the forehead. A wireless amplifier (Neuracle, China) was mounted on both of the headsets and the data were sampled at 1000 Hz. Thus, the difference between the two EEG headset systems lies in the type of electrodes. Considering the long-term ease of use in real-world applications, the dry-electrode headset was used as a primary headset. The wet-electrode headset was used as an auxiliary headset to augment the performance of the dry-electrode headset. Thus, for each subject, the wet-electrode headset was employed prior to the dry-electrode headset. Before the experiment, the impedance was kept below 200 $k\Omega$ and 20 $k\Omega$ for dry and wet electrodes, respectively.

3) Experimental Procedure: This study designed a cued spelling SSVEP-BCI experiment. Twelve characters, including ten digits (0 – 9) and 2 punctuations (* and #), were presented on the screen to simulate a 12-target dialing keyboard (Fig. 1(**A**)). Each target on the keyboard had a dimension of 288×288 pixels

$(8.5^\circ \times 8.5^\circ)$ and it was encoded by joint frequency and phase modulation (JFPM) [7] to present visual flickers. As illustrated in Fig. 1(**B**), the stimulus frequency of the flicker ranged from 9.25 Hz to 14.75 Hz (frequency interval: 0.5 Hz) and the initial phase ranged from 0 to 1.5 π (phase interval: 0.5 π). Each stimulus was modulated by a sampled sinusoidal stimulation method [23] at a refresh rate of 60 Hz. The visual stimuli were presented on a 27-inch LCD monitor by MATLAB (MathWorks, Inc.) via Psychophysics Toolbox Version 3 [24].

In the SSVEP-BCI experiment, each subject took part in a session of the wet-electrode headset and a subsequent session of the dry-electrode headset on a single day. To eliminate confounding variables, the tasks of the two sessions were identical, and the electrode placement was maintained as close as possible when switching from wet-electrode headset to dry-electrode headset. In each session, subjects performed 10 blocks of the experiment, and there was a break between two consecutive blocks to mitigate visual fatigue. There were 12 trials in a block, in which each trial corresponded to a target, and the order of trials was randomized. At the beginning of each trial, a red square covering the target was prompted, and all targets started to flicker simultaneously. The subjects were asked to direct their attention to the prompted target and avoid eyeblink during flickering. The flickering of visual stimulation lasted 2 s and there was a 1-s rest time for gaze shift. Using the portable amplifier, the event triggers sent from the computer were synchronized to the EEG data as an event channel. The recorded data were resampled to 250 Hz and preprocessed with an infinite impulse response (IIR) notch filter to remove the power line interference.

B. Align and Pool for EEG Headset Domain Adaptation

1) Problem Definition: The notations and descriptions used in the paper are summarized in Table I. Given a labeled source domain $\mathcal{D}_s = \{\mathcal{X}_s, \mathcal{Y}_s\} = \{(\boldsymbol{X}_s^{(i)}, y_s^{(i)})\}, i = 1, \ldots, N_s$ and an unlabeled target domain $\mathcal{D}_t = \{\mathcal{X}_t\} = \{\boldsymbol{X}_t^{(i)}\}, i = 1, \ldots, N_t$, the problem of cross-device transfer learning is to exploit auxiliary EEG data from $\mathcal{X}_s$ to decode trials from $\mathcal{X}_t$. In this context, $\mathcal{D}_s$ and $\mathcal{D}_t$ come from two EEG headsets ($\mathcal{D}_s \neq \mathcal{D}_t$) and satisfy $\mathcal{X}_s \neq \mathcal{X}_t, \mathcal{Y}_s = \mathcal{Y}_t, |\mathcal{Y}| = N_c$ where $|\cdot|$ denotes the cardinality of a set. The data samples $\boldsymbol{X}_s \in \mathbb{R}^{N_{ch} \times N_p}, \boldsymbol{X}_t \in \mathbb{R}^{N_{ch} \times N_p}$ are multi-channel EEG trials from the same subject. For a practical SSVEP-BCI application, it is often the case that $N_s = N_c \times N_b, N_t = 1$ for the purpose of zero training in $\mathcal{D}_t$, in which the model learned from $\mathcal{D}_s$ is used to classify single-trial target data in a recalibration-free manner.

In the conventional scheme of SSVEP frequency recognition ($\mathcal{D}_s = \mathcal{D}_t$), a spatial filter $\boldsymbol{w}$ is constructed from the training data, and the grand average trials $\bar{\boldsymbol{X}}$ of a subject are used as SSVEP template. Both the SSVEP template $\bar{\boldsymbol{X}}$ and a test trial $\boldsymbol{X}_t$ are then weighted by $\boldsymbol{w}$ to yield the spatially filtered one-dimensional embeddings $\boldsymbol{w}^T\bar{\boldsymbol{X}}$ and $\boldsymbol{w}^T\boldsymbol{X}_t$. Finally a similarity measure of Pearson's correlation is performed on the embeddings to infer the stimulus frequencies with the largest correlation coefficient. For the problem of cross-device transfer learning, since $\boldsymbol{w}$ and $\bar{\boldsymbol{X}}$ lie in $\mathcal{D}_s$ and $\boldsymbol{X}_t$ lies in $\mathcal{D}_t$ ($\mathcal{D}_s \neq \mathcal{D}_t$), the existing domain gap would inevitably introduce bias in the similarity measure. As a result, the bias would deteriorate the decoding performance in classification accuracy. To address this issue, we propose the **AL**ign and **P**ool for **H**eadset domain **A**daptation (ALPHA) framework in the present study. As illustrated in Fig. 2, the proposed framework subsumes 3 key components, i.e., subspace decomposition, subspace alignment and subspace pooling. The details of each component are described as follows.

TABLE I
TABLE OF NOTATIONS

Notation	Description
N_b	Number of blocks
N_c	Number of classes
N_{ch}	Number of channels
N_h	Number of harmonics
N_k	Number of subspaces
$N_s(N_t)$	Number of trials in the source (target) domain
N_p	Number of sampling points
f_s	Sampling rate
$\boldsymbol{w} \in \mathbb{R}^{N_{ch} \times 1}$	The optimal projection direction for SSVEP data
$\boldsymbol{w}_\rho \in \mathbb{R}^{N_k \times 1}$	The optimal projection direction for correlation coefficients
$\boldsymbol{X}_t^{(i)} \in \mathbb{R}^{N_{ch} \times N_p}$	The i-th multi-channel test trial or SSVEP trial in the target domain
$\boldsymbol{X}_s^{(i)} \in \mathbb{R}^{N_{ch} \times N_p}$	The i-th multi-channel SSVEP trial in the source domain
$\boldsymbol{Y}_{sc} \in \mathbb{R}^{2N_h \times N_p}$	Sine-cosine reference
$\bar{\boldsymbol{X}}^{(i)} \in \mathbb{R}^{N_{ch} \times N_p}$	SSVEP template of the i-th trial's class by subjects-specific grand average
$\bar{\boldsymbol{X}}^{i} \in \mathbb{R}^{N_{ch} \times N_p}$	SSVEP template of the i-th class by subject-specific grand average
$\boldsymbol{W} \in \mathbb{R}^{N_{ch} \times N_{ch}}$	A set of spatial filters for SSVEP data
$\boldsymbol{W}_{AB}(\widetilde{\boldsymbol{W}}_{AB}) \in \mathbb{R}^{N_{ch} \times N_{ch}}$	A set of spatial filters by CCA (LDA) from $\boldsymbol{A}$ and $\boldsymbol{B}$ for SSVEP data
$\boldsymbol{A}_s(\boldsymbol{A}_t) \in \mathbb{R}^{N_{ch} \times N_{ch}}$	Spatial pattern of SSVEP in the source (target) domain
$\boldsymbol{P}(\boldsymbol{Q}) \in \mathbb{R}^{N_{ch} \times N_{ch}}$	Transformation matrix in the ASP (AC)
$\boldsymbol{S}_b(\boldsymbol{S}_w) \in \mathbb{R}^{N_{ch} \times N_{ch}}$	Between-class (within-class) scatter matrix
$\boldsymbol{R}(\boldsymbol{R}_{nt}) \in \mathbb{R}^{N_c \times N_k}$	Correlation coefficient matrix (corresponding to the nontarget)
$\mathbf{C}_n \in \mathbb{R}^{n \times n}$	Centering matrix
$\mathbf{H} \in \mathbb{R}^{N_b \times N_s}$	Selecting matrix for the target stimulus
$\mathcal{C}_i$	The i-th class
$\mathcal{D}_s(\mathcal{D}_t)$	Source (target) domain
$\mathcal{X}_s(\mathcal{X}_t)$	Sample in the source (target) domain
$\mathcal{Y}_s(\mathcal{Y}_t)$	Label in the source (target) domain

2) Subspace Decomposition: For the step of subspace decomposition in ALPHA, we perform the subspace decomposition in $\mathcal{D}_s$ and $\mathcal{D}_t$ simultaneously and obtain its associated spatial filters. Inspired by extended CCA [5] and m-extended CCA [7], we utilized canonical correlation analysis (CCA) [4], [25] as the main operation for feature extraction. Different from the conventional approaches that use only the largest subspace, here multiple subspaces are utilized after the CCA operation. The optimization objective of CCA is defined as [26]

$$\begin{aligned} \underset{\boldsymbol{U},\boldsymbol{V}}{\text{argmax}} \quad & tr\left(\boldsymbol{U}^T\boldsymbol{X}\boldsymbol{Y}^T\boldsymbol{V}\right) \\ \text{s.t.} \quad & \boldsymbol{U}^T\boldsymbol{X}\boldsymbol{Y}^T\boldsymbol{V} = \boldsymbol{\Lambda} \\ & \boldsymbol{U}^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{U} = \mathbf{I}, \quad \boldsymbol{V}^T\boldsymbol{Y}\boldsymbol{Y}^T\boldsymbol{V} = \mathbf{I} \end{aligned} \tag{1}$$

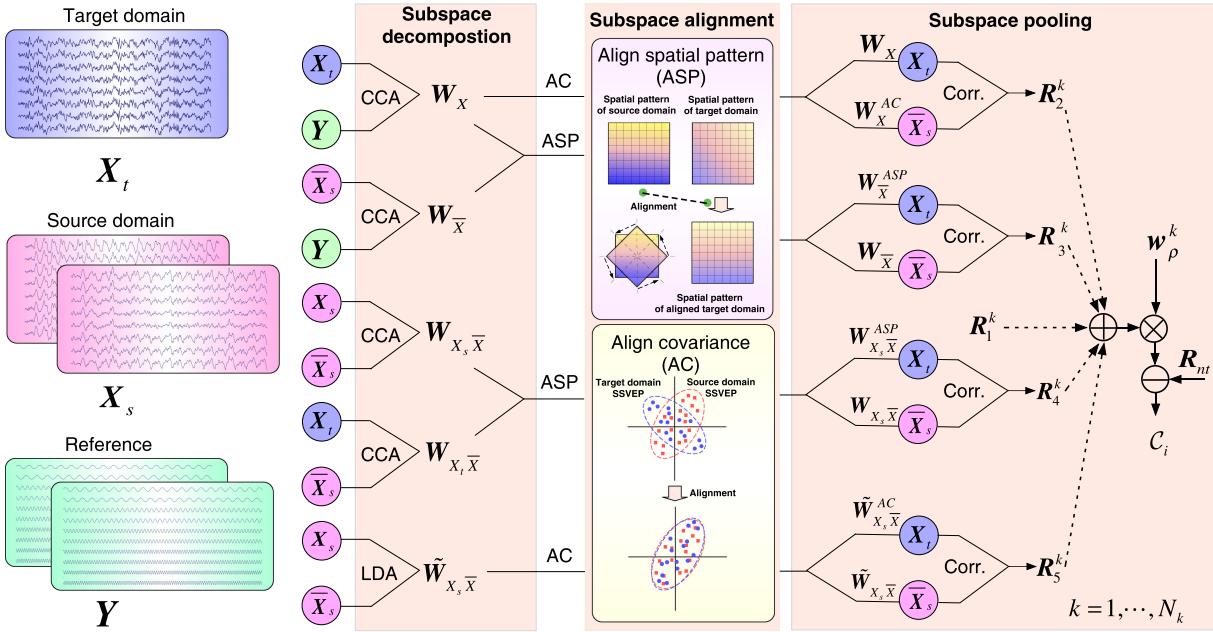


Fig. 2. A schematic diagram of the ALPHA framework. Each color in the circle represents a different type of data, e.g., purple denotes test data or samples from the target domain; magenta denotes samples from the source domain; and green denotes the sine-cosine reference signal. The ALPHA subsumes 3 key steps, i.e., subspace decomposition, subspace alignment and subspace pooling. Subspace decomposition is performed on each pair of input data and the spatial filters are then obtained. The subspace alignment encompasses two parts, i.e., the ASP (align spatial pattern) and the AC (align covariance). In the illustration of ASP, the spatial pattern of the target domain is aligned to that of the source domain by a mathematical rotation. In the illustration of AC, the covariance of the target domain is aligned to that of the source domain by whitening and coloring. Subspace pooling aggregates the correlation coefficient from each pair of embeddings and weights the correlation coefficients from multiple subspaces to form detection statistics. (Best view in color).

where $\boldsymbol{X}$ is the EEG data and $\boldsymbol{Y}$ is a reference signal corresponding to the query stimulus frequency f. The reference signal can be a sine-cosine reference signal $\boldsymbol{Y}_{sc}$ or a source SSVEP template $\bar{\boldsymbol{X}}_s$. The sine-cosine reference signal $\boldsymbol{Y}_{sc}$ is formed as [4], [25].

$$\boldsymbol{Y}_{sc} = \begin{bmatrix} \sin\left(2\pi f \boldsymbol{t}^T\right) \\ \cos\left(2\pi f \boldsymbol{t}^T\right) \\ \vdots \\ \sin\left(2\pi N_h f \boldsymbol{t}^T\right) \\ \cos\left(2\pi N_h f \boldsymbol{t}^T\right) \end{bmatrix}, \boldsymbol{t} = [1/f_s, \ldots, N_p/f_s]^T \quad (2)$$

In ALPHA, we use CCA subspace decomposition to obtain 4 sets of spatial filters. The first 2 sets of spatial filters are associated with the sine-cosine reference signal $\boldsymbol{Y}_{sc}$. In $\mathcal{D}_t$, an intersection subspace between the target trial $\boldsymbol{X}_t$ and $\boldsymbol{Y}_{sc}$ can be determined, i.e., $\boldsymbol{W}_X = \boldsymbol{U}$, when $\boldsymbol{X} = \boldsymbol{X}_t$ and $\boldsymbol{Y} = \boldsymbol{Y}_{sc}$ in Eq. (1). In $\mathcal{D}_s$, an intersection subspace between the source SSVEP template $\bar{\boldsymbol{X}}_s$ and $\boldsymbol{Y}_{sc}$ can be determined, i.e., $\boldsymbol{W}_{\bar{X}} = \boldsymbol{U}$, when $\boldsymbol{X} = \bar{\boldsymbol{X}}_s$ and $\boldsymbol{Y} = \boldsymbol{Y}_{sc}$. Note that the first two sets of spatial filters $\boldsymbol{W}_X$ and $\boldsymbol{W}_{\bar{X}}$ are derived from the same sine-cosine reference. Thus the difference between $\mathcal{D}_t$ and $\mathcal{D}_s$ may be reflected in the difference between $\boldsymbol{W}_X$ and $\boldsymbol{W}_{\bar{X}}$, which makes it possible to compensate for the domain gap introduced from switching EEG headsets.

In parallel, the 3rd and 4th sets of spatial filters are associated with the source SSVEP template $\bar{\boldsymbol{X}}_s$. Likewise, in $\mathcal{D}_s$ an intersection subspace between the single-trial source EEG data $\boldsymbol{X}_s$ and $\bar{\boldsymbol{X}}_s$ can be determined, i.e., $\boldsymbol{W}_{X_s\bar{X}} = \boldsymbol{U}$, when $\boldsymbol{X} = \boldsymbol{X}_s$ and $\boldsymbol{Y} = \bar{\boldsymbol{X}}_s$. In $\mathcal{D}_t$, an intersection subspace between the target EEG data $\boldsymbol{X}_t$ and $\bar{\boldsymbol{X}}_s$ can be determined, i.e., $\boldsymbol{W}_{X_t\bar{X}} = \boldsymbol{U}$, when $\boldsymbol{X} = \boldsymbol{X}_t$ and $\boldsymbol{Y} = \bar{\boldsymbol{X}}_s$. As indicated by a previous study [27], the 3rd set of spatial filters $\boldsymbol{W}_{X_s\bar{X}}$ is mathematically equivalent to the TRCA spatial filter [8].

Furthermore, we introduce a set of discriminant spatial filters to SSVEP-BCI frequency recognition using two-dimensional linear discriminant analysis (LDA) [28], [29]. The optimization objective of the spatial filter is known as the Fisher criterion

$$\underset{\boldsymbol{W}}{\text{maximize}} \quad \frac{tr(\boldsymbol{W}^T \boldsymbol{S}_b \boldsymbol{W})}{tr(\boldsymbol{W}^T \boldsymbol{S}_w \boldsymbol{W})} \quad (3)$$

The between-class scatter matrix $\boldsymbol{S}_b$ and the within-class scatter matrix $\boldsymbol{S}_w$ are defined as

$$\begin{aligned} \boldsymbol{S}_b &= \boldsymbol{H}_b \boldsymbol{H}_b^T \\ \boldsymbol{S}_w &= \boldsymbol{H}_w \boldsymbol{H}_w^T \end{aligned} \quad (4)$$

where $\boldsymbol{H}_b \in \mathbb{R}^{N_{ch} \times N_c N_p}$ and $\boldsymbol{H}_w \in \mathbb{R}^{N_{ch} \times N_s N_p}$ have the form

$$\begin{aligned} \boldsymbol{H}_b &= \frac{1}{\sqrt{N_c}}[\bar{\boldsymbol{X}}^1 - \bar{\boldsymbol{X}}^a, \ldots, \bar{\boldsymbol{X}}^{N_c} - \bar{\boldsymbol{X}}^a] \\ \boldsymbol{H}_w &= \frac{1}{\sqrt{N_s}}[\boldsymbol{X}^{(1)} - \bar{\boldsymbol{X}}^{(1)}, \ldots, \boldsymbol{X}^{(N_s)} - \bar{\boldsymbol{X}}^{(N_s)}] \end{aligned} \quad (5)$$

The superscript a denotes all classes. Note that the EEG data are from $\mathcal{D}_s$, and the subscript here is omitted. The set of spatial filters obtained from LDA is denoted as $\widetilde{\boldsymbol{W}}$.

3) Subspace Alignment: After the spatial filters are obtained, the next step in ALPHA is to perform subspace alignment. The goal of subspace alignment is to align $\mathcal{D}_s$ and $\mathcal{D}_t$ to the same subspace to bridge the gap between the two domains. Two types of subspace alignment are employed in the study.

Align spatial pattern (ASP). For SSVEP-BCI frequency recognition, we propose a novel alignment method to align the spatial patterns of SSVEP from $\mathcal{D}_s$ and $\mathcal{D}_t$. For the set of CCA spatial filters $\boldsymbol{W}_{\bar{X}}$ and $\boldsymbol{W}_X$, the corresponding spatial pattern in $\mathcal{D}_s$ and $\mathcal{D}_t$ can be obtained by [30]

$$\begin{aligned} \boldsymbol{A}_s &= \boldsymbol{W}_{\bar{X}}^{-T} \\ \boldsymbol{A}_t &= \boldsymbol{W}_X^{-T} \end{aligned} \tag{6}$$

Here, we assume that $\mathcal{D}_s$ and $\mathcal{D}_t$ should share the same underlying spatial pattern after a rotation. Thus, the ASP method seeks an orthogonal transformation matrix $\boldsymbol{P}$ to the estimated spatial pattern, which leads to the following optimization problem.

$$\begin{aligned} \underset{\boldsymbol{P}}{\text{minimize}} \quad & \left\| \boldsymbol{A}_s - \boldsymbol{A}_t \boldsymbol{P}^T \right\|_F^2 \\ \text{s.t.} \quad & \boldsymbol{P}\boldsymbol{P}^T = \mathbf{I} \end{aligned} \tag{7}$$

This is a typical Procrustes problem in mathematics [31] and the solution is obtained by

$$\boldsymbol{P} = \boldsymbol{U}\boldsymbol{V}^T \tag{8}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are the left and right singular vectors of $\boldsymbol{A}_s^T \boldsymbol{A}_t$. Then, the transformed spatial filters for $\boldsymbol{W}_X$ has the form

$$\boldsymbol{W}_{\bar{X}}^{ASP} = \boldsymbol{W}_X \boldsymbol{P}^T \tag{9}$$

In a similar vein, the spatial pattern corresponding to $\boldsymbol{W}_{X_s\bar{X}}$ and $\boldsymbol{W}_{X_t\bar{X}}$ can be aligned. In this study, the spatial pattern of $\boldsymbol{W}_{X_t\bar{X}}$ is rotated to that of $\boldsymbol{W}_{X_s\bar{X}}$ via the ASP, and the details are presented in the Supplementary Material.

Align covariance (AC). Apart from the shift in spatial pattern, we assume that the domain gap also lies in the covariance shift. To mitigate the covariance shift, we should align the covariance between the distribution of $\mathcal{D}_s$ and $\mathcal{D}_t$. Here, we used a correlation alignment (CORAL) [32] based method to fulfill the goal. CORAL finds a linear transformation $\boldsymbol{Q}$ to minimize the distance between the second-order statistics of $\mathcal{D}_s$ and $\mathcal{D}_t$ [32].

$$\underset{\boldsymbol{Q}}{\text{minimize}} \quad ||\boldsymbol{Q}^T \boldsymbol{C}_s \boldsymbol{Q} - \boldsymbol{C}_t||_F^2 \tag{10}$$

Different from the original CORAL, in ALPHA the covariance matrices $\boldsymbol{C}_s$ and $\boldsymbol{C}_t$ for $\mathcal{D}_s$ and $\mathcal{D}_t$ are obtained by

$$\begin{aligned} \boldsymbol{C}_s^i &= \frac{1}{N_b N_p - 1} \boldsymbol{X}_m^i \mathbf{C}_{N_b N_p} \boldsymbol{X}_m^{i^T}, \quad i = 1, \ldots, N_c \\ \boldsymbol{C}_t &= \frac{1}{N_p - 1} \boldsymbol{X}_t \mathbf{C}_{N_p} \boldsymbol{X}_t^T \end{aligned} \tag{11}$$

where $\mathbf{C}_n = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ is a centering matrix, and $\boldsymbol{X}_m^i$ is the concatenation of trials $\boldsymbol{X}_s$ in $\mathcal{C}_i$. Eq. (10) can be solved in a closed form [32], and it has an equivalent yet efficient solution

$$\boldsymbol{Q} = \boldsymbol{C}_s^{-\frac{1}{2}} \boldsymbol{C}_t^{\frac{1}{2}} \tag{12}$$

For the CCA spatial filter obtained from the last section, after AC it has the form

$$\boldsymbol{W}_X^{AC} = \boldsymbol{Q}\boldsymbol{W}_X \tag{13}$$

In a similar vein, by aligning $\mathcal{D}_t$ to $\mathcal{D}_s$ the LDA spatial filter is given by

$$\widetilde{\boldsymbol{W}}_{X_s\bar{X}}^{AC} = \boldsymbol{Q}^{-1} \widetilde{\boldsymbol{W}}_{X_s\bar{X}} \tag{14}$$

4) Subspace Pooling: After $\mathcal{D}_s$ and $\mathcal{D}_t$ are aligned to the same domain, a similarity measure procedure is performed to aggregate the similarities to a detection statistic. Different from the conventional approach, here we use more than one subspace, and the correlation coefficients of the subspace are pooled by weighting. Specifically, the correlation coefficients of the subspace are obtained by

$$\boldsymbol{R} = \sum_{i=1}^{5} \text{sign}(\boldsymbol{R}_i) \odot \boldsymbol{R}_i \odot \boldsymbol{R}_i \tag{15}$$

where $\odot$ is the Hadamard product. $\boldsymbol{R}_1$ is given by the first N_k canonical correlation coefficient in Eq. (1) when $\boldsymbol{X} = \boldsymbol{X}_t$ and $\boldsymbol{Y} = \boldsymbol{Y}_{sc}$, and the remaining is given by

$$\begin{aligned} \boldsymbol{R}_2^k &= corr(\boldsymbol{W}_{X[k]}^T \boldsymbol{X}_t, \boldsymbol{W}_{X[k]}^{AC^T} \bar{\boldsymbol{X}}_s), \quad k = 1, \ldots, N_k. \\ \boldsymbol{R}_3^k &= corr(\boldsymbol{W}_{\bar{X}[k]}^{ASP^T} \boldsymbol{X}_t, \boldsymbol{W}_{\bar{X}[k]}^T \bar{\boldsymbol{X}}_s), \quad k = 1, \ldots, N_k. \\ \boldsymbol{R}_4^k &= corr(\boldsymbol{W}_{X_s\bar{X}[k]}^{ASP^T} \boldsymbol{X}_t, \boldsymbol{W}_{X_s\bar{X}[k]}^T \bar{\boldsymbol{X}}_s), \quad k = 1, \ldots, N_k. \\ \boldsymbol{R}_5^k &= corr(\widetilde{\boldsymbol{W}}_{X_s\bar{X}[k]}^{AC^T} \boldsymbol{X}_t, \widetilde{\boldsymbol{W}}_{X_s\bar{X}[k]}^T \bar{\boldsymbol{X}}_s), \quad k = 1, \ldots, N_k. \end{aligned} \tag{16}$$

where N_k is the number of subspaces used for pooling and is set to 3 in the study. The first part and the second part in each correlation operation in Eq. (16) are the test embedding and template embedding, respectively.

We learn an optimal projection direction $\boldsymbol{w}_\rho \in \mathbb{R}^{N_k \times 1}$ from the data in $\mathcal{D}_s$. Specifically, a leave-one-block-out cross validation is conducted on the source domain EEG data, yielding a source $\boldsymbol{R}_s \in \mathbb{R}^{N_s \times N_k}$ for each class in Eq. (15). Then, for each class, $\boldsymbol{w}_\rho$ is obtained by maximizing the $\mathcal{L}_2$ norm of the target correlation coefficients, which can be formulated as

$$\begin{aligned} \underset{\boldsymbol{w}_\rho}{\text{maximize}} \quad & ||\mathbf{H}\boldsymbol{R}_s \boldsymbol{w}_\rho||_F^2 \\ \text{s.t.} \quad & \boldsymbol{w}_\rho^T \boldsymbol{w}_\rho = 1 \end{aligned} \tag{17}$$

where $\mathbf{H} \in \mathbb{R}^{N_b \times N_s}$ is a selecting matrix that satisfies

$$\mathbf{H}_{i,j} = \begin{cases} 1, & y_i \in \mathcal{C}_m, \quad j = N_c(i-1) + m \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

Eq. (17) can be cast into a Rayleigh quotient, and $\boldsymbol{w}_\rho$ is composed of the eigenvector corresponding to the largest eigenvalue of $\boldsymbol{R}_s^T \mathbf{H}^T \mathbf{H} \boldsymbol{R}_s$.

Inspired by the CCARV [33], the weighted correlation coefficients corresponding to the nontarget are averaged to form $\boldsymbol{R}_{nt}$.

Finally the decision criterion for frequency recognition is based on the statistics after weighting and variation reduction.

$$\mathcal{C}_i = \underset{i}{\operatorname{argmax}} \quad (\boldsymbol{R}^i - \boldsymbol{R}^i_{nt})\boldsymbol{w}^i_\rho \tag{19}$$

C. Evaluation

1) Data Evaluation: The signal profiles of the EEG signals from the two headsets were qualitatively and quantitatively compared. To visually display the profile in the temporal and spectral domains, a grand average was applied to the EEG data. The visual latency was estimated from the stimulus frequencies in the first row of the keyboard [34]. The SNR statistics, including the narrow-band SNR and wide-band SNR, were calculated [35]. In addition, we summarized the impedance value of the recording electrodes during the data acquisition.

2) Performance Evaluation: According to a previous study [18], two scenarios of cross-device transfer learning, i.e., from wet EEG headset to dry EEG headset (wet to dry), and from dry EEG headset to wet EEG headset (dry to wet), were considered. The performance of two competing transfer learning based methods (LST [18] and ttCCA [14]) and a baseline training-free method (CCA [4], [6]) were compared with ALPHA. The single-trial SSVEP data were latency-corrected by d ms [6]–[8], and the latency d was estimated based on the data [34]. Data lengths from 0.2 s to 2 s with an interval of 0.2 s were used for evaluation. The classification accuracy and simulated online ITR were calculated [1], [18]. A gaze shift time of 1 s was used to calculate the ITR. Note that ttCCA was designed for cross-subject transfer learning [14], and here, we evaluated it on cross-device transfer learning.

We then compared the performance of ALPHA with a state-of-the-art fully-calibrated method, i.e., task-related component analysis (TRCA) [8]. For TRCA the fully-calibrated setting, the recalibration-free setting, and the hybrid calibration setting were evaluated, while for ALPHA the recalibration-free setting and the hybrid calibrated setting were evaluated. In the fully-calibrated setting, a k-fold cross validation was performed by leaving one block out in each fold ($k = N_b$) in the target domain. In the recalibration-free setting, the model (ALPHA or TRCA) was trained from the source domain and tested on the target domain. In the hybrid calibration setting, a k-fold leave-one-block-out cross validation on the target domain was performed, in which we trained the model (ALPHA or TRCA) by simultaneously utilizing EEG from the source domain and the training set of the target domain and tested the model on the test set of the target domain. In implementation, we trained two models, i.e., the model from source domain EEG and from target domain EEG, and then the detection statistics (i.e., correlation coefficients) of each model were added. In each of the settings, the classification accuracy and ITR were obtained. For the wet-to-dry scenario, we further evaluated the performance of the two methods with varying parameters of calibration blocks. For the recalibration-free ALPHA, the number of source blocks varied from 1 to N_b. For the fully-calibrated TRCA, the number of training blocks varied from 1 to $N_b - 1$. For each parameter, we measured the performance by the maximum ITR across time after averaging the ITR values by block. The CCA was evaluated on the target domain, and its performance was used as a baseline.

We further validate the efficacy of ALPHA in a simulation manner. To simulate the scenario of exploiting high-SNR data (wet-electrode SSVEP) to augment the low-SNR data (dry-electrode SSVEP), we conducted a simulation experiment as follows. For each subject, we split the data from the 1st session (i.e., using a wet EEG headset) into source data (block 1 to 5) and target data (block 6 to 10). Then, we added drop-out instability and random noise with baseline drift [36] to the target data. This is in an effort to mimic the dry-electrode EEG, which is empirically found to suffer from contact instabilities and artifacts. Specifically, 2 of the 8 channels were zeroed out, and Gaussian white noise with a baseline drift $n(t) \sim \mathcal{N}(\mu, 1), \mu = 1$ was added to the remaining channels similar to [36]. To evaluate the performance of ALPHA on the contaminated simulation data, two transfer learning methods (ttCCA and LST) and a training-free baseline method (CCA) were compared.

In addition, we conducted a follow-up study to examine the effectiveness of ALPHA on the scenario of cross-day and cross-device transfer learning. The details of the experimental procedure are presented in the Supplementary Material.

3) Feature Evaluation: In this part, the efficacy of features in ALPHA was evaluated. First, the ITR obtained from using only 1 of the 5 correlation coefficients $\boldsymbol{R}$ in Eq. (15) was compared with the original result. Second, the ITR with a varying number of subspaces used in Eq. (16) was compared with the conventional approach of using a single largest subspace. Third, the effect of subspace alignment (ASP and AC) was explored. We initially compared the classification accuracy by removing the subspace alignment step in ALPHA. To characterize the change in the discriminability of SSVEP induced by subspace alignment, an R-squared statistic of the correlation coefficient was used [8], [37]. The R-squared of each correlation coefficient associated with the target and nontarget stimuli in Eq. (16) before and after subspace alignment was compared. Subsequently, a t-distributed stochastic neighbor embedding (t-SNE) [38] was applied to project $\boldsymbol{W}^T_{X[1]}\boldsymbol{X}_t, \boldsymbol{W}^{AC^T}_{X[1]}\bar{\boldsymbol{X}}_s$ and $\boldsymbol{W}^T_{X[1]}\bar{\boldsymbol{X}}_s$ in Eq. (16) to a low dimensional space to visualize the embeddings. The t-SNE was also applied to $\boldsymbol{W}^T_{\bar{X}[1]}\bar{\boldsymbol{X}}_s, \boldsymbol{W}^{ASP^T}_{\bar{X}[1]}\boldsymbol{X}_t$ and $\boldsymbol{W}^T_{\bar{X}[1]}\boldsymbol{X}_t$ in Eq. (16). Finally the wide-band SNRs of $\boldsymbol{W}^{ASP^T}_{\bar{X}[1]}\boldsymbol{X}_t$ and $\boldsymbol{W}^T_{\bar{X}[1]}\boldsymbol{X}_t$ were compared to unveil the effect of subspace alignment on the embedding's SNR.

4) Filter Bank Analysis: Filter bank analysis [6] is an effective temporal filtering method in SSVEP-BCI, which is generally orthogonal to most spatial filtering methods and could further enhance the performance. In the present study, each of the aforementioned methods was evaluated without the filter bank analysis (w/o FB) and with the filter bank analysis (w/ FB). The number of filter banks was set to 5 [8], [35], and the weight for each filter bank was set according to [6]. In ALPHA, the $\boldsymbol{R}$ values corresponding to different filter banks were weighted and summed before learning the coefficients in Eq. (17). Note that the ensemble method [8] was used for TRCA both w/ and w/o FB.

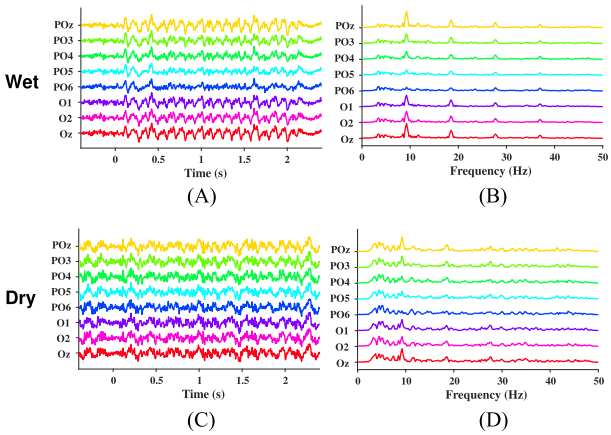


Fig. 3. Temporal (left panel) and spectral (right panel) profiles of SSVEP from wet (upper panel) and dry (lower panel) EEG headsets. Each color represents a channel. It can be observed that the dry-electrode EEG is noisier than the wet-electrode EEG. (Best view in color).

TABLE II
COMPARISON OF THE WET-EEG DATA AND THE DRY-EEG DATA

	Latency (ms)	Wide-band SNR (dB)	Narrow-band SNR (dB)	Impedance (kΩ)
Wet	150.24±5.02	-17.51±0.31	2.7±0.26	21.15±1.93
Dry	151.61±4.64	-18.33±0.29	2.26±0.25	160.09±10.27
p-value	.806	<.001	<.001	<.001

5) Statistical Analysis: For the comparison of multiple methods, a two-way repeated-measures analysis of variance (ANOVA) was conducted to determine whether there was a statistically significant interaction between two within-subjects factors of method and data length. Greenhouse-Geisser correction was applied in light of violations of sphericity assessed by Mauchly's test. When the main effect was significant ($p < .05$), post hoc pairwise comparisons of t-tests were performed with the adjustment of Bonferroni correction. For the comparison of two methods, a planned paired t-test was conducted to evaluate the statistical significance. Bonferroni adjustment was applied in the case of nonorthogonal contrast. Data were presented as the mean ± standard error of the mean (s.e.m) unless otherwise stated. The statistical procedures were conducted using SPSS Statistics 26 (IBM, Armonk, NY, USA).

III. RESULTS

1) Data Evaluation: Fig. 3 illustrates the temporal and spectral profiles of grand-average SSVEP for dry and wet EEG headsets. From visual inspection, it is evident that data from the dry EEG headset were noisier than the wet EEG headset in the temporal and spectral domains across channels. This is supported by the SNR statistics in Table II. Compared with the SSVEP from the wet EEG headset, the SSVEP from the dry EEG headset had a significantly lower wide-band SNR (dry: -18.33 ± 0.29 dB; wet: -17.51 ± 0.31 dB; $p < .001$) and narrow-band SNR (dry: 2.26 ± 0.25 dB; wet: 2.7 ± 0.26 dB; $p < .001$). The low-SNR attribute of the dry EEG headset is due in part to its higher impedance. The impedance of dry electrodes was approximately 8 times greater than that of wet electrodes (dry: 160.09 ± 10.27 kΩ; wet: 21.15 ± 1.93 kΩ; $p < .001$). In addition, the result of estimated latency showed that there was no significant difference between the two EEG headset systems (dry: 151.61 ± 4.64 ms; wet: 150.24 ± 5.02 ms; $p = .806$). The estimated latencies were comparable because the amplifier and stimulus computer were identical for the two sessions. Thus we choose $d = 150$ ms for latency correction in the following classification analysis.

2) Performance Evaluation: The performance of different transfer learning based methods as well as the training-free CCA method is illustrated in Fig. 4. In the upper panel the dry-electrode EEG was used as test data (wet-to-dry scenario), while in the lower panel the wet-electrode EEG was used as test data (dry-to-wet scenario). In the wet-to-dry scenario, the two-way repeated measures ANOVA revealed that there was a statistically significant interaction between the method and data length on accuracy, $F(2.847,\ 210.677) = 26.858, p < .001$ (w/o FB), $F(2.619,\ 193.835) = 27.740, p < .001$ (w/ FB), and on ITR, $F(2.526,\ 186.959) = 28.384, p < .001$ (w/o FB), $F(2.364.\ 174.909) = 30.825, p < .001$ (w/ FB). In the dry-to-wet scenario, the two-way repeated measures ANOVA revealed that there was a statistically significant interaction between the method and data length on accuracy, $F(4.406,\ 326.034) = 51.745, p < .001$ (w/o FB), $F(3.607,\ 266.941) = 46.146, p < .001$ (w/ FB), and on ITR, $F(3.793,\ 280.652) = 43.932, p < .001$ (w/o FB), $F(3.079,\ 227.882) = 42.045, p < .001$ (w/ FB). In both of the scenarios, the follow-up simple main effects were run and post-hoc pairwise comparison with a Bonferroni adjustment revealed that ALPHA outperformed other methods significantly ($p < .001$) for all data lengths, either on accuracy or ITR (w/ or w/o FB). The data length corresponding to the highest ITR varied between different methods and the details are summarized in Table III.

Fig. 5 illustrates the performance of TRCA compared with ALPHA in the fully-calibrated, the recalibration-free and the hybrid calibration settings. For the fully-calibrated TRCA (blue line), ALPHA in the recalibration-free setting (red line) significantly outperformed it in the wet-to-dry transfer direction on accuracy and ITR at all data lengths w/o FB, and at the data length of $0.4 - 2$s w/ FB, $p < .05$. For the recalibration-free TRCA (black line), ALPHA in this setting (red line) significantly outperformed it on accuracy and ITR regardless of the transfer directions, data lengths and the use of filter banks, $p < .05$. For TRCA in the hybrid calibration setting (cyan line), ALPHA in the hybrid calibration setting (magenta line) significantly ($p < .05$) outperformed it on accuracy and ITR for all scenarios (transfer directions, data lengths and the use of filter banks) except the dry-to-wet transfer direction w/ FB. When comparing the recalibration-free ALPHA (red line) and the TRCA in hybrid calibration setting (cyan line), in the wet-to-dry transfer direction, ALPHA significantly outperformed TRCA in the w/o FB scenario ($p < .001$) at all data lengths, and significantly surpassed TRCA in the w/ FB scenario ($p < .01$) at data lengths greater than 0.8 s.

When the number of training blocks was reduced from 9 to 1, the follow-up analysis in Fig. 6 shows that ALPHA significantly

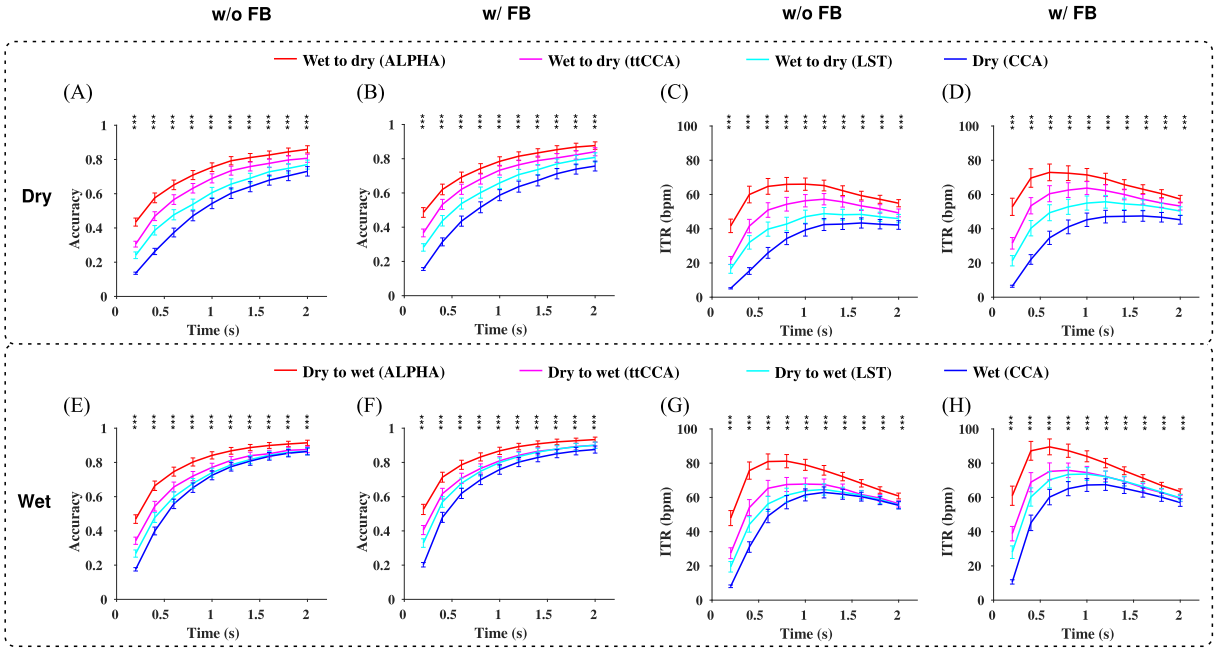


Fig. 4. The average classification accuracy (**A**, **B**, **E** and **F**) and ITR (**C**, **D**, **G** and **H**) for the transfer learning approaches (ttCCA and LST) and baseline approach (CCA). Data lengths of 0.2 s to 2 s with an interval of 0.2 s were used for evaluation. The upper panel illustrates the result when the test data or target domain is the dry-electrode EEG. "Wet to dry" indicates the transfer from wet-electrode EEG to dry-electrode EEG. The lower panel illustrates the result when the test data or target domain is the wet-electrode EEG. "Dry to wet" indicates the transfer from dry-electrode EEG to wet-electrode EEG. The w/o FB or w/ FB indicates the scenarios when all the approaches are applied without or with filter bank analysis, respectively. For instance, CCA in the scenario of w/ FB is FBCCA. The asterisks indicate a significant difference between ALPHA and ttCCA ($^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$, Bonferroni corrected). (Best view in color).

TABLE III
ITRs OF THE TRANSFER LEARNING BASED METHODS AND THE BASELINE METHOD

ITR (bpm) (Data length (s))[1]	ALPHA	ttCCA[14]	LST[18]	CCA[4], [6]
Wet to dry (w/o FB)	66.02 ± 3.58 (1.0)	57.24 ± 3.30 (1.2)	48.81 ± 3.58 (1.2)	43.36 ± 2.97 (1.6)
Wet to dry (w/ FB)	72.91 ± 4.87 (0.6)	63.73 ± 3.81 (1.0)	55.74 ± 3.71 (1.2)	47.44 ± 3.09 (1.6)
Dry to wet (w/o FB)	81.18 ± 3.86 (0.8)	67.79 ± 3.56 (1.0)	64.64 ± 3.31 (1.2)	62.92 ± 3.21 (1.2)
Dry to wet (w/ FB)	89.52 ± 4.64 (0.6)	75.82 ± 4.25 (0.8)	73.66 ± 3.74 (1.0)	67.46 ± 3.36 (1.2)

[1]The data length corresponding to the highest ITR.

outperformed TRCA on ITR either w/o or w/ FB ($p < .001$). It follows that ALPHA could achieve and surpass TRCA on the performance by leveraging a considerably fewer number of training blocks. More specifically, for the mean ITR in the w/ FB case, 1 training block in ALPHA (54.96 bpm) approximated 5 training blocks in TRCA (52.54 bpm). Two training blocks in ALPHA (59.66 bpm) approximated 7 training blocks in TRCA (59.54 bpm). Three training blocks in ALPHA (64.41 bpm) approximated 8 training blocks in TRCA (61.66 bpm). Four training blocks in ALPHA (65.68 bpm) approximated 9 training blocks in TRCA (64.58 bpm). Note that ALPHA yielded better performance than the baseline FBCCA regardless of the number of training blocks N_b, while TRCA was superior over the baseline when $N_b > 4$.

The classification accuracy and ITR on the simulation data are illustrated in Supplementary Fig. 1. Dovetailing with the result of the realistic data in Fig. 4, ALPHA outperformed ttCCA and LST as well as CCA substantially at all data lengths. As assessed by two-way repeated measures ANOVA, there was a statistically significant interaction between the method and data length on accuracy, $F(4.911,\ 363.423) = 36.145, p < .001$ (w/o FB), $F(4.080,\ 301.903) = 36.839, p < .001$ (w/ FB), and on ITR, $F(3.174,\ 234.852) = 23.972, p < .001$ (w/o FB), $F(2.982,\ 220.650) = 25.414, p < .001$ (w/ FB). Follow-up simple main effects and post-hoc pairwise comparison revealed that the differences between ALPHA and other methods were significant ($p < .001$, Bonferroni corrected) for all data lengths, either on accuracy or ITR (w/ or w/o FB).

For the cross-day and cross-device transfer scenario, the details of result are presented in the Supplementary Material.

3) Feature Evaluation: Fig. 7 depicts the maximum ITRs when only 1 correlation coefficient in Eq. (15) was used. In comparison with the original result, a significant drop in ITR was found ($p < .001$, Bonferroni corrected). This indicates that each feature played a prominent role in the feature fusion step of Eq. (15). In the subsequent step of Eq. (16), the effect of different numbers of subspaces N_k on the maximum ITR is illustrated in Supplementary Fig. 3. Compared with the conventional approach of $N_k = 1$, the introduction of more subspaces led to a tendency of an initial increase in the ITR and plateaued at N_k^{opt}

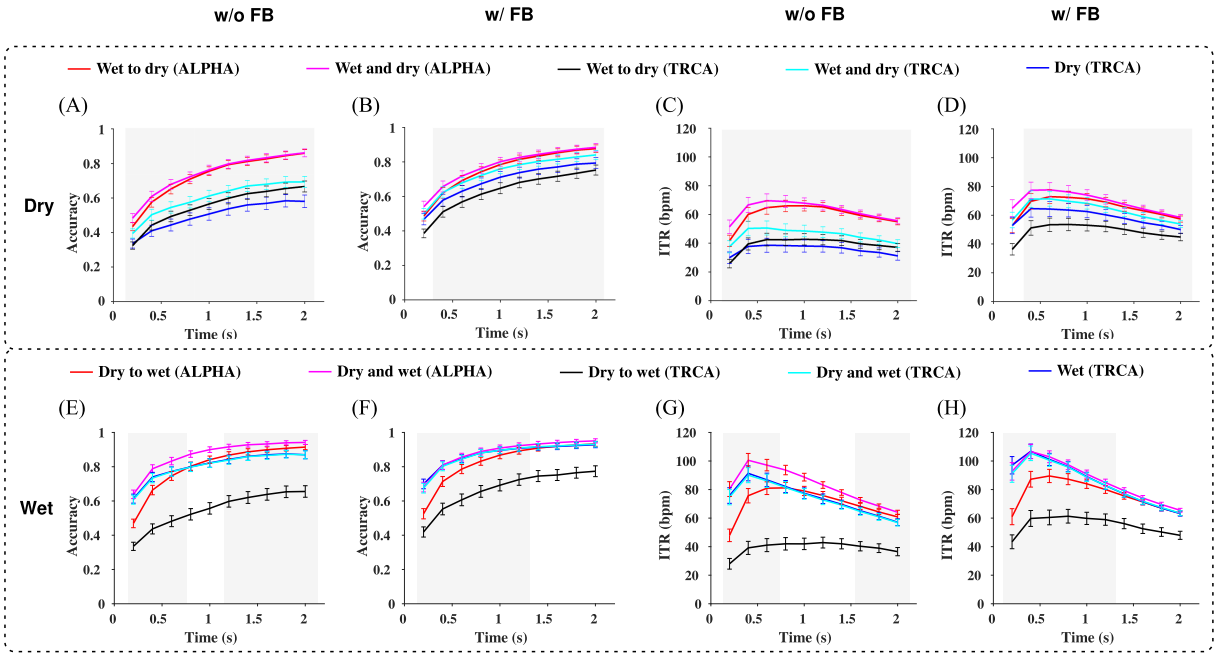


Fig. 5. The average classification accuracy (**A**, **B**, **E** and **F**) and ITR (**C**, **D**, **G** and **H**) for TRCA and ALPHA in various settings. The red line and black line ("wet to dry," or "dry to wet") indicate the performance of the recalibration-free setting, i.e., transfer from wet-electrode EEG to dry-electrode EEG, or vice versa. The magenta line and cyan line ("wet and dry," or "dry and wet") indicates the performance of the hybrid calibration setting, i.e., utilizing wet-electrode and dry-electrode EEG simultaneously to classify dry-electrode EEG or wet-electrode EEG by cross validation. The blue line indicates the performance of the fully-calibrated setting, i.e., evaluating the fully-calibrated TRCA in the target domain ("dry" or "wet") by cross validation. The upper panel illustrates the result when the test data or target domain is the dry-electrode EEG. The lower panel illustrates the result when the test data or target domain is the wet-electrode EEG. Data lengths of 0.2 s to 2 s with an interval of 0.2 s were used for evaluation. The w/o FB or w/ FB indicates the scenarios when all the approaches are applied without or with filter bank analysis, respectively. For instance, CCA in the scenario of w/ FB is FBCCA. The gray shaded areas indicate a significant difference between the recalibration-free ALPHA (red line) and the fully-calibrated TRCA (blue line) (Best view in color).

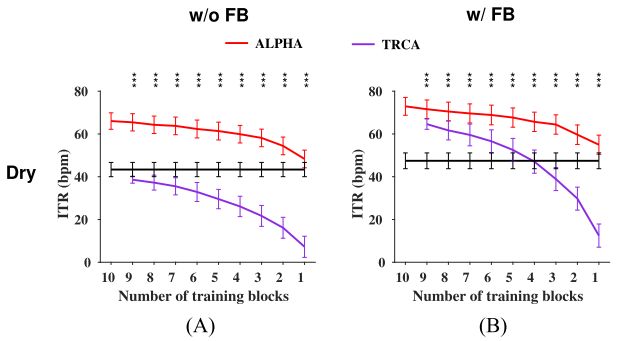


Fig. 6. The average ITR with a varying number of training blocks for ALPHA and TRCA when the test data or target domain is the dry-electrode EEG. The data length used for evaluation corresponds to the maximum average ITR of each method. The dark line denotes the baseline CCA method. Since CCA is training-free, it is not affected by the number of training blocks. The asterisks indicate a significant difference between ALPHA and TRCA (*$p < .05$, **$p < .01$, ***$p < .001$). (Best view in color).

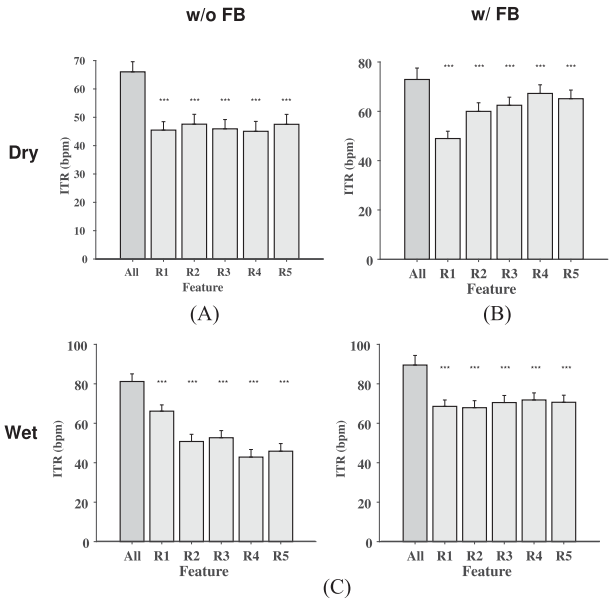


Fig. 7. The average ITR when only one of the correlation coefficients is used compared with the case when all correlation coefficients are used. The data length used for evaluation corresponds to the maximum average ITR of each method. The asterisks indicate a significant difference between the case of 1 correlation coefficient and the case of all correlation coefficients (*$p < .05$, **$p < .01$, ***$p < .001$, Bonferroni corrected).

(wet to dry, w/o FB: $p < .001, N_k^{opt} = 3$; dry to wet, w/o FB: $p < .001, N_k^{opt} = 3$; wet to dry, w/ FB: $p < .01, N_k^{opt} = 3$; dry to wet, w/ FB: $p = .45, N_k^{opt} = 2$; Bonferroni corrected).

The effect of the subspace alignment on the accuracy is illustrated in Fig. 8. For the scenarios of wet-to-dry transfer and dry-to-wet transfer (either w/o or w/ FB), the accuracy at each

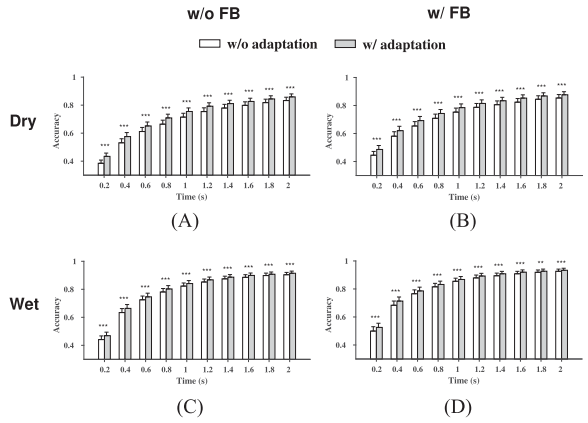


Fig. 8. The effect of subspace alignment on the average accuracy in various scenarios. Data lengths from 0.2 s to 2 s with an interval of 0.2 s were used for evaluation. The "adaptation" in the legend denotes the domain adaptation method in the study, i.e., subspace alignment. The asterisks indicate a significant difference between the pairs (*$p < .05$, **$p < .01$, ***$p < .001$).

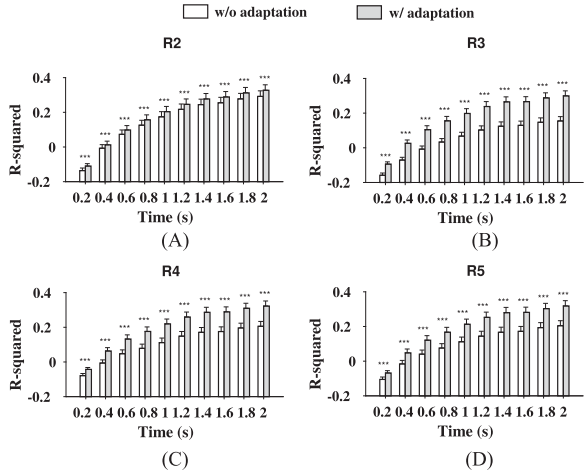


Fig. 9. The effect of subspace alignment on the R-squared statistics for various correlation coefficients. Data lengths from 0.2 s to 2 s with an interval of 0.2 s were used for evaluation. The asterisks indicate a significant difference between the pairs (*$p < .05$, **$p < .01$, ***$p < .001$).

data length was elevated significantly ($p < .001$) by aligning the source SSVEPs and target SSVEPs to the same domain via the proposed ASP and AC modules. This demonstrates the efficacy of the subspace alignment in ALPHA. Then, the following intermediate features were presented in an effort to delve into the cause and effect relationship between the subspace alignment and the improved accuracy.

As illustrated in Fig. 9, compared with the R-squared when no subspace alignment was applied, each of the R-squared for different data lengths was significantly boosted ($p < .001$) after the subspace alignment (ASP for $\boldsymbol{R}_3$ and $\boldsymbol{R}_4$; AC for $\boldsymbol{R}_2$ and $\boldsymbol{R}_5$). This is reasonable since in SSVEP-BCI, a higher R-squared of the correlation coefficient features leads to an improved accuracy [8], [37]. To further characterize the two embeddings before the correlation operation in Eq. (16), the t-SNE features for the embeddings before and after subspace alignment are diagrammed in Fig. 10 and Supplementary Fig. 4. In Fig. 10, the

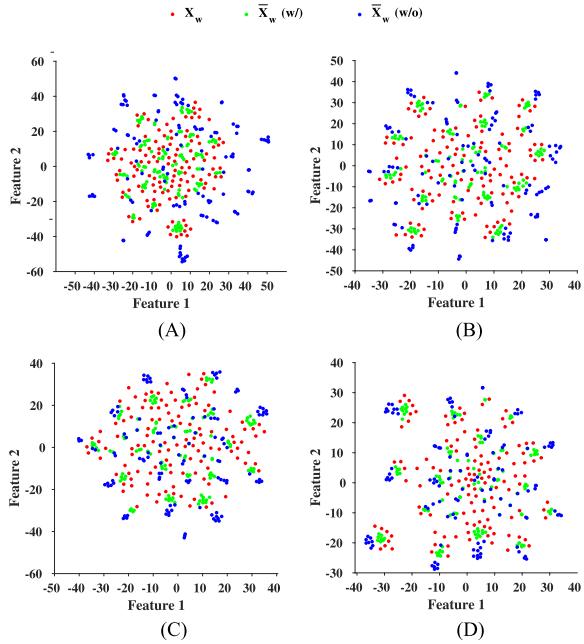


Fig. 10. The effect of subspace alignment (AC) on the low dimensional feature space projected by t-SNE. The data length used for evaluation is 2 s. The test embeddings ($\boldsymbol{W}_{X[1]}^T \boldsymbol{X}_t$) are represented by red dots. The template embeddings w/ ($\boldsymbol{W}_{X[1]}^{AC^T} \bar{\boldsymbol{X}}_s$) and w/o ($\boldsymbol{W}_{X[1]}^T \bar{\boldsymbol{X}}_s$) subspace alignment are represented by green dots and blue dots, respectively. Each subfigure illustrates the data from a subject, and each dot represents a trial. The template embeddings are moved toward the test embeddings under the influence of the subspace alignment. (Best view in color).

test embeddings ($\boldsymbol{W}_{X[1]}^T \boldsymbol{X}_t$, red dots) and template embeddings w/ ($\boldsymbol{W}_{X[1]}^{AC^T} \bar{\boldsymbol{X}}_s$, green dots) and w/o ($\boldsymbol{W}_{X[1]}^T \bar{\boldsymbol{X}}_s$, blue dots) subspace alignment are illustrated. It could be observed from the visualization that the template embeddings were moved toward the test embeddings by subspace alignment (AC) which aligned the source domain to the target domain. In a similar vein, the template embeddings ($\boldsymbol{W}_{\bar{X}[1]}^T \bar{\boldsymbol{X}}_s$, red dots) and test embeddings w/ ($\boldsymbol{W}_{\bar{X}[1]}^{ASP^T} \boldsymbol{X}_t$, green dots) and w/o ($\boldsymbol{W}_{\bar{X}[1]}^T \boldsymbol{X}_t$, blue dots) subspace alignment are illustrated in Supplementary Fig. 4. Since the template embedding here came from the wet EEG headset with a high SNR, the embeddings were densely clustered and each cluster could correspond to a stimulus frequency. The imposition of the subspace alignment (ASP) contributed to pushing the test embeddings toward the cluster of template embeddings. Specifically, the test embeddings of $\boldsymbol{W}_{\bar{X}[1]}^{ASP^T} \boldsymbol{X}_t$ and $\boldsymbol{W}_{\bar{X}[1]}^T \boldsymbol{X}_t$ were further evaluated. The comparison in Fig. 11 revealed that the subspace alignment (ASP) significantly enhanced the SNR of the test embeddings at all data lengths ($p < .001$).

IV. Discussion

Methodologically, it is noteworthy that ALPHA demonstrated its utility in both transfer directions, as shown in Fig. 4. The bidirectional boosting effect introduced by the auxiliary cross-device EEG data was in line with the prior study of LST [18].

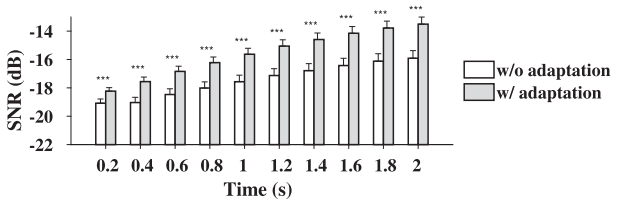


Fig. 11. The effect of subspace alignment on the SNR (wide-band SNR) for various data lengths. Data lengths from 0.2 s to 2 s with an interval of 0.2 s were used for evaluation. The asterisks indicate a significant difference between the pairs (*$p < .05$, **$p < .01$, ***$p < .001$).

However, ALPHA outperformed LST by a good margin in both transfer directions. It is of interest to note that ttCCA also showcased its efficacy in the cross-device transfer task, notwithstanding that it was originally designated for a cross-subject transfer task [14]. Thus, for a fair comparison, ttCCA was included in the study as a competitive method. Under the framework of ALPHA, i.e., subspace decomposition, subspace alignment and subspace pooling, ttCCA could be cast as a special case of ALPHA. The hallmark of ALPHA is that it provides an additional step of subspace alignment after obtaining spatial filters in the subspace decomposition. As such, Fig. 8 to Fig. 11 validated its efficacy and attempted to probe a causal role of the subspace alignment. As indicated by the findings, after aligning the source domain and target domain to a congruent subspace, the SNR of the spatially filtered SSVEP was markedly enhanced. In the feature space, the enhancement of SNR would exert a facilitative effect on reducing the distance between the test embedding and template embedding. Accordingly, the R-squared after the correlation operation was heightened, which led to improved classification accuracy and ITR. For subspace pooling, the ALPHA capitalized on more subspaces than the conventional first largest subspace. For most of the cases, the introduction of more subspaces brought about improvement in ITR. This is reasonable since the additional subspaces may contain discriminative information for classification, which has received less attention in conventional frequency recognition approaches [39], [40].

From a perspective of practical application, the utility of ALPHA in the transfer direction of wet-to-dry EEG headsets merits extra attention. In this scenario, with zero training for the dry EEG headset, ALPHA has been shown to outperform other transfer learning based methods (ttCCA and LST) as well as the fully-calibrated method (TRCA) as evidenced in Fig. 4, Fig. 5, and the result of simulation data in Supplementary Fig. 1. By leveraging additional data from wet EEG headset as hybrid calibration, the performance of ALPHA could be further improved and it significantly outperformed the TRCA method in the same setting. For the fully-calibrated method, we also compared the extended CCA (eCCA) [5] and found that ALPHA significantly outperformed eCCA ($p < .001$) under various conditions (w/o FB, w/ FB and w/ different number of training blocks) in the wet-to-dry scenario. Thus the results indicate the effectiveness of ALPHA and imply that the conventional fully-calibrated protocol may not be best suited for the dry electrode based SSVEP-BCI system. To boost the performance of the system, an alternative is to use ALPHA to exploit the calibration data from a high-SNR source, e.g., the gold-standard wet EEG headset. With the aid of auxiliary data from the high-SNR source domain, the performance of the dry-electrode system (w/o FB: 66.02 ± 3.58 bpm, w/ FB: 72.91 ± 4.87 bpm) could surpass the training-free baseline of the wet-electrode system (w/o FB: 62.92 ± 3.21 bpm, w/ FB: 67.46 ± 3.36 bpm), as shown in Table III and Fig. 4. Moreover, the number of calibration blocks required could be reduced. As suggested by Fig. 6, using only 4 blocks of wet-electrode data in the alternative protocol via ALPHA could yield performance comparable to 9 blocks of dry-electrode data in the conventional fully-calibrated protocol. Quantitatively, measured by the data utility $\eta = \frac{ITR - ITR^{FBCCA}}{N_b}$ (a metric modified from the cost-performance ratio (CPR) in [20]), the alternative protocol via ALPHA has a significantly higher data utility than the conventional protocol ($\eta^{ALPHA} = 5.52 \pm 0.32$, $\eta^{TRCA} = 2.44 \pm 0.27$, $p < .001$). Compared with other BCI paradigms, the reduction in the calibration block is practically more important. This is because the SSVEP-BCI has a downside of eliciting visual fatigue, and the fatigue deteriorates the performance and user experience. Thus, the effort toward fewer blocks of calibration data is indispensable in practical settings, and ALPHA can lend insight to it. For real-world applications, the alternative protocol via ALPHA can improve the usefulness of dry electrode based SSVEP-BCI systems. In other practical scenarios, for instance, when we are concerned that the gel of a wet EEG headset is prone to dry out in a test session, the alternative protocol can also be applied to augment the efficacy of a dry EEG headset. In addition, for computational cost in practical application, the single-trial EEG (e.g., 2-s data length in w/ FB scenario) could be recognized by ALPHA with less than 90-ms computation overhead on a desktop computer with a 3.60 GHz CPU (16 GB RAM), which indicates the feasibility of online implementation.

A large number of subjects (i.e., 75) in the study helps to reduce bias in the evaluation of the algorithm. Nevertheless, the present study has limitations, and the following issues await further investigation. The prospect of our work is to develop a performant dry-electrode SSVEP-BCI system for daily applications with the assistance of wet-electrode EEG. Besides cross-device transfer learning, the present study reveals promising result on the cross-day transfer learning. In future studies, other confounding variables, e.g., cross-subject variability etc. may be taken into account to realize a long-term plug-and-play transfer learning system. Thus, more methodological improvements and engineering applications are necessitated in our future work. Recent advances in deep transfer learning could be introduced in considering the large data volume in the study. It should be noted that ALPHA can also be tailored for other tasks, e.g., cross-subject transfer learning and training-based frequency recognition. Its components, such as the subspace pooling could probably yield greater ITR gains considering the short data length and gaze shift time used in the training-based methods. The efficacy of ALPHA is orthogonal to existing feature enhancement methods [6], [19], and it could be applied with filter bank analysis [6] and other techniques [19].

V. Concluding Remarks

To conclude, we proposed a recalibration-free cross-device transfer learning framework, i.e., the **AL**ign and **P**ool for EEG **H**eadset domain **A**daptation (ALPHA) to boost the performance of dry electrode based SSVEP-BCI. Domain adaptation is introduced to the SSVEP-BCI and is accomplished by aligning the spatial pattern and covariance of the source domain and target domains. The efficacy of ALPHA is corroborated under various conditions, and the impact on the features exerted by the subspace alignment is characterized. The result indicates that ALPHA achieves state-of-the-art performance among the transfer learning methods, whether transferring from wet to dry EEG headsets or from dry to wet EEG headsets. From the viewpoint of practical utility, ALPHA outperforms the fully-calibrated method of TRCA when transferring from wet to dry EEG headsets. By augmenting the dry-electrode EEG with wet-electrode EEG, the need for calibration blocks could be lessened. Seen in this light, the present study opens an avenue for boosting the performance of dry-electrode SSVEP-BCI by high-fidelity auxiliary data and proposes potential solutions to expedite its real-world application.

Acknowledgment

We would like to thank X. Liu, H. Zhao, F. Zhu and S. Tian for their assistance in data collection, and thank L. Jiang for the help in the manuscript preparation. We would like to thank the reviewers for their constructive advice on the manuscript.

References

[1] J. R. Wolpaw *et al.*, "Brain-computer interfaces for communication and control," *Clin. Neuriophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.

[2] X. Gao *et al.*, "Interface, interaction, and intelligence in generalized brain-computer interfaces," *Trends Cogn. Sci.*, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364661321000966

[3] Y. Zhang *et al.*, "Data analytics in steady-state visual evoked potential-based brain-computer interface: A review," *IEEE Sensors J.*, vol. 21, no. 2, pp. 1124–1138, Jan. 2021.

[4] G. Bin *et al.*, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *J. Neural Eng.*, vol. 6, no. 4, 2009, Art. no. 046002.

[5] M. Nakanishi *et al.*, "A high-speed brain speller using steady-state visual evoked potentials," *Int. J. Neural Syst.*, vol. 24, no. 6, 2014, Art. no. 1450019.

[6] X. Chen *et al.*, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface," *J. Neural Eng.*, vol. 12, no. 4, 2015, Art. no. 046008.

[7] X. Chen *et al.*, "High-speed spelling with a noninvasive brain-computer interface," *Proc. the Nat. Acad. Sci. United States America*, vol. 112, no. 44, 2015, Art. no. 201508080.

[8] M. Nakanishi *et al.*, "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 104–112, Jan. 2018.

[9] A. Luo and T. J. Sullivan, "A user-friendly ssvep-based brain-computer interface using a time-domain classifier," *J. Neural Eng.*, vol. 7, no. 2, 2010, Art. no. 026010.

[10] V. Mihajlović, G. Garcia-Molina, and J. Peuscher, "Dry and water-based eeg electrodes in SSVEP-based BCI applications," in *Proc. Int. Joint Conf. Biomed. Eng. Syst. Technol.*. Springer, 2012, pp. 23–40.

[11] X. Xing *et al.*, "A high-speed SSVEP-based BCI using dry eeg electrodes," *Sci. Rep.*, vol. 8, no. 1, pp. 1–10, 2018.

[12] L. Angrisani *et al.*, "A wearable brain-computer interface instrument for augmented reality-based inspection in industry 4.0," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 4, pp. 1530–1539, Apr. 2020.

[13] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[14] P. Yuan *et al.*, "Enhancing performances of SSVEP-based brain-computer interfaces via exploiting inter-subject information," *J. Neural Eng.*, vol. 12, no. 4, 2015, Art. no. 046006.

[15] K. Suefusa and T. Tanaka, "Reduced calibration by efficient transformation of templates for high speed hybrid coded SSVEP brain-computer interfaces," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2017, pp. 929–933.

[16] W. Liu *et al.*, "A cross-subject SSVEP-BCI based on task related component analysis," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2019, pp. 3022–3025.

[17] K.-J. Chiang *et al.*, "Cross-subject transfer learning improves the practicality of real-world applications of brain-computer interfaces," in *Proc. 9th Int. IEEE/EMBS Conf. Neural Eng.*, 2019, pp. 424–427.

[18] M. Nakanishi *et al.*, "Facilitating calibration in high-speed BCI spellers via leveraging cross-device shared latent responses," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 4, pp. 1105–1113, 2019.

[19] C. M. Wong *et al.*, "Learning across multi-stimulus enhances target recognition methods in SSVEP-based BCIs," *J. Neural Eng.*, vol. 17, no. 1, 2019, Art. no. 016026.

[20] C. M. Wong *et al.*, "Inter-and intra-subject transfer reduces calibration effort for high-speed SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 10, pp. 2123–2135, Oct. 2020.

[21] K.-J. Chiang *et al.*, "Boosting template-based SSVEP decoding by cross-domain transfer learning," *J. Neural Eng.*, vol. 18, no. 1, 2020, Art. no. 016002.

[22] F. Zhu *et al.*, "An open dataset for wearable SSVEP-based brain-computer interfaces," *Sensors*, vol. 21, no. 4, 2021, Art. no. 1256.

[23] X. Chen *et al.*, "A high-itr SSVEP-based BCI speller," *Brain-Comput. Interfaces*, vol. 1, no. 3-4, pp. 181–191, 2014.

[24] D. H. Brainard, "The psychophysics toolbox," *Spatial Vis.*, vol. 10, no. 4, pp. 433–436, 1997.

[25] Z. Lin *et al.*, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, Jun. 2007.

[26] C. M. Wong *et al.*, "Spatial filtering in SSVEP-based BCIs: Unified framework and new improvements," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 11, pp. 3057–3072, Nov. 2020.

[27] Q. Wei *et al.*, "A training data-driven canonical correlation analysis algorithm for designing spatial filters to enhance performance of SSVEP-based BCIs," *Int. J. Neural Syst.*, vol. 30, no. 5, 2020, Art. no. 2050020.

[28] J. Ye, R. Janardan, and Q. Li, "Two-dimensional linear discriminant analysis," *Adv. Neural Inf. Process. Syst.*, vol. 17, pp. 1569–1576, 2004.

[29] Y. Zhang *et al.*, "Spatial-temporal discriminant analysis for ERP-based brain-computer interface," *IEEE Trans. neural Syst. Rehabil. Eng.*, vol. 21, no. 2, pp. 233–243, Mar. 2013.

[30] S. Haufe *et al.*, "On the interpretation of weight vectors of linear models in multivariate neuroimaging," *Neuroimage*, vol. 87, pp. 96–110, 2014.

[31] P. H. Schönemann, "A generalized solution of the orthogonal procrustes problem," *Psychometrika*, vol. 31, no. 1, pp. 1–10, 1966.

[32] B. Sun, J. Feng, and K. Saenko, "Correlation alignment for unsupervised domain adaptation," *Domain Adapt. Comput. Vis. Appl.*, Springer, 2017, pp. 153–171.

[33] E. Yin *et al.*, "A dynamically optimized SSVEP brain-computer interface (BCI) speller," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1447–1456, 2014.

[34] J. Pan *et al.*, "Enhancing the classification accuracy of steady-state visual evoked potential-based brain-computer interfaces using phase constrained canonical correlation analysis," *J. Neural Eng.*, vol. 8, no. 3, 2011, Art. no. 036027.

[35] B. Liu *et al.*, "Beta: A large benchmark database toward SSVEP-BCI application," *Front. Neurosci.*, vol. 14, p. 627, 2020.

[36] A. D. Degenhart *et al.*, "Stabilization of a brain-computer interface via the alignment of low-dimensional spaces of neural activity," *Nature Biomed. Eng.*, vol. 4, no. 7, pp. 672–685, 2020.

[37] Y. Wang *et al.*, "Enhancing detection of steady-state visual evoked potentials using individual training data," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2014, pp. 3037–3040.

[38] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.

[39] Y. Zhang *et al.*, "Hierarchical feature fusion framework for frequency recognition in SSVEP-based BCIs," *Neural Netw.*, vol. 119, pp. 1–9, 2019.

[40] K.-J. Chiang, M. Nakanishi, and T.-P. Jung, "Statistically optimized spatial filtering in decoding steady-state visual evoked potentials based on task-related component analysis," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2020, pp. 3070–3073.

I. Supplementary Material

A. Align spatial pattern (ASP) for the 3rd and 4th set of filters

For the 3rd set of spatial filters, the spatial filters are obtained by calculating the source SSVEP template and all the single-trial source EEG data in Eq. (1). Thus the 3rd set of spatial filters lie in the source domain, similar to the 2nd set of spatial filters. And the 4th set of spatial filters can be regarded to lie in the target domain, similar to the 1st set of spatial filters. Since the 3rd and 4th set of spatial filters are derived from the same reference, i.e., source SSVEP template, the difference between source domain and target domain can be estimated and compensated by the subspace alignment of ASP. More formally, the details are the reusing of Eq. (6) to (9) as follows.

For the 3rd set of spatial filters $\boldsymbol{W}_{X_s\bar{X}}$ and and the 4th set of spatial filters $\boldsymbol{W}_{X_t\bar{X}}$, the corresponding spatial pattern in $\mathcal{D}_s$ and $\mathcal{D}_t$ have the form

$$\begin{aligned} \boldsymbol{A}_s &= \boldsymbol{W}_{X_s\bar{X}}^{-T} \\ \boldsymbol{A}_t &= \boldsymbol{W}_{X_t\bar{X}}^{-T} \end{aligned} \tag{S1}$$

By assuming $\mathcal{D}_s$ and $\mathcal{D}_t$ share underlying spatial pattern after affine transformation, we can solve the following optimization objective

$$\begin{aligned} \underset{\boldsymbol{P}}{\text{minimize}} \quad & \left\|\boldsymbol{A}_s - \boldsymbol{A}_t\boldsymbol{P}^T\right\|_F^2 \\ \text{s.t.} \quad & \boldsymbol{P}\boldsymbol{P}^T = \mathbf{I} \end{aligned} \tag{S2}$$

The solution to the Procrustes problem can be written as

$$\boldsymbol{P} = \boldsymbol{U}\boldsymbol{V}^T \tag{S3}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are the left and right singular vectors of $\boldsymbol{A}_s^T\boldsymbol{A}_t$, i.e.,

$$\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T = \boldsymbol{A}_s^T\boldsymbol{A}_t \tag{S4}$$

$\boldsymbol{S}$ is a diagonal matrix in the singular value decomposition. Then the transformed spatial filter for $\boldsymbol{W}_{X_t\bar{X}}$ can be written as

$$\boldsymbol{W}_{X_t\bar{X}}^{ASP} = \boldsymbol{W}_{X_t\bar{X}}\boldsymbol{P}^T \tag{S5}$$

B. The pilot study of cross-day and cross-device transfer learning

For a user in practical applications, especially for the patients (e.g., with amyotrophic lateral sclerosis, ALS), there exists a scenario of long-term use for the BCI system that provides assistance and communication. For this scenario, the cross-day variability should be taken into account, besides the cross-device variability. To this end, one of our followed-up works applies the proposed method to the problem of cross-day and cross-device transfer learning.

In our pilot study, 5 subjects participated in the experiment of a 40-target SSVEP-BCI spelling task. Each subject performed the spelling task on two different days (day 1 and day 2). The interval of the two days lasted at least one week. Two different EEG headsets, i.e., wet-electrode headset and dry-electrode headset were used in each day, and the wet-electrode headset was applied prior to the dry-electrode headset. In each session of the wet or dry EEG headset, 6 blocks of SSVEP-BCI task were performed, and each block comprised 40 trials. Each trial consisted of 4-s visual stimulation tagged by a stimulus frequency, which ranged from 8 Hz to 15.8 Hz with an interval of 0.2 Hz. The EEG recording and preprocessing were in line with the 12-target experiment.

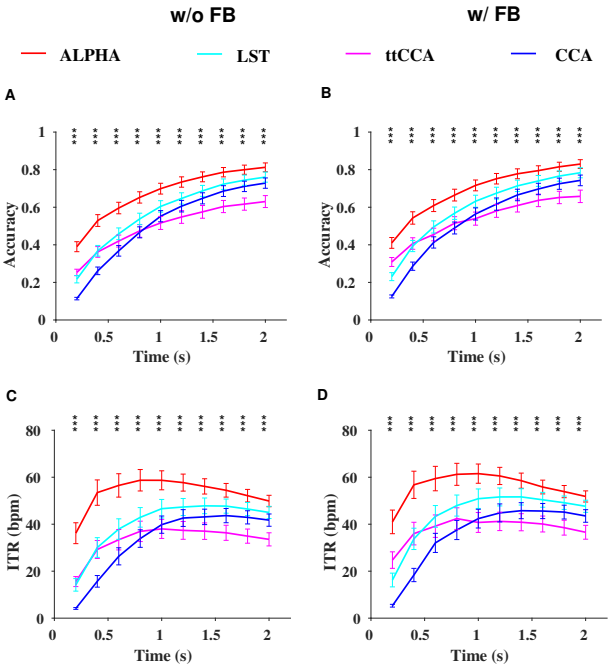


Fig. 1. (Supplementary Figure 1) The average classification accuracy (**A**, **B**) and ITR (**C**, **D**) for the transfer learning approaches (ttCCA and LST) and baseline approach (CCA) evaluated on the contaminated simulation data. The simulated transfer direction is from the wet EEG headset to the dry EEG headset. Data lengths from 0.2 s to 2 s with an interval of 0.2 s were used for evaluation. The w/o FB or w/ FB indicates the scenarios when all the approaches are applied without or with filter bank analysis, respectively. For instance, CCA in the scenario of w/ FB is FBCCA. The asterisks indicate a significant difference between ALPHA and LST ($*p < .05$, $**p < .01$, $***p < .001$, Bonferroni corrected). (Best view in color.)

In performance evaluation, we compared 3 transfer learning methods (ALPHA, ttCCA and LST), a baseline training-free method (CCA/FBCCA), and a fully-calibrated method (TRCA) in different cross-day and cross-device scenarios. The post-hoc cross-day analysis includes the transfer from day 1 to day 2 (1to2), day 2 to day 1 (2to1), day 1 to day 1 (1to1), day 2 to day 2 (2to2). The cross-device scenarios include wet to dry, dry to wet, as well as dry to dry, and wet to wet. For all the scenarios, EEG data from the source domain were used as training data and the EEG from the target domain were used as test data.

When the test data is dry-electrode EEG, the result for the w/ FB case is illustrated in Supplementary Figure 2. For the cross-day and cross-device transfer, ALPHA (red line) outperformed other methods in various scenarios. The average maximum ITRs for the within-day wet to dry (1to1 and 2to2), the cross-day wet to dry (1to2 and 2to1), the cross-day dry to dry (1to2 and 2to1), and the fully-calibrated TRCA (day 1

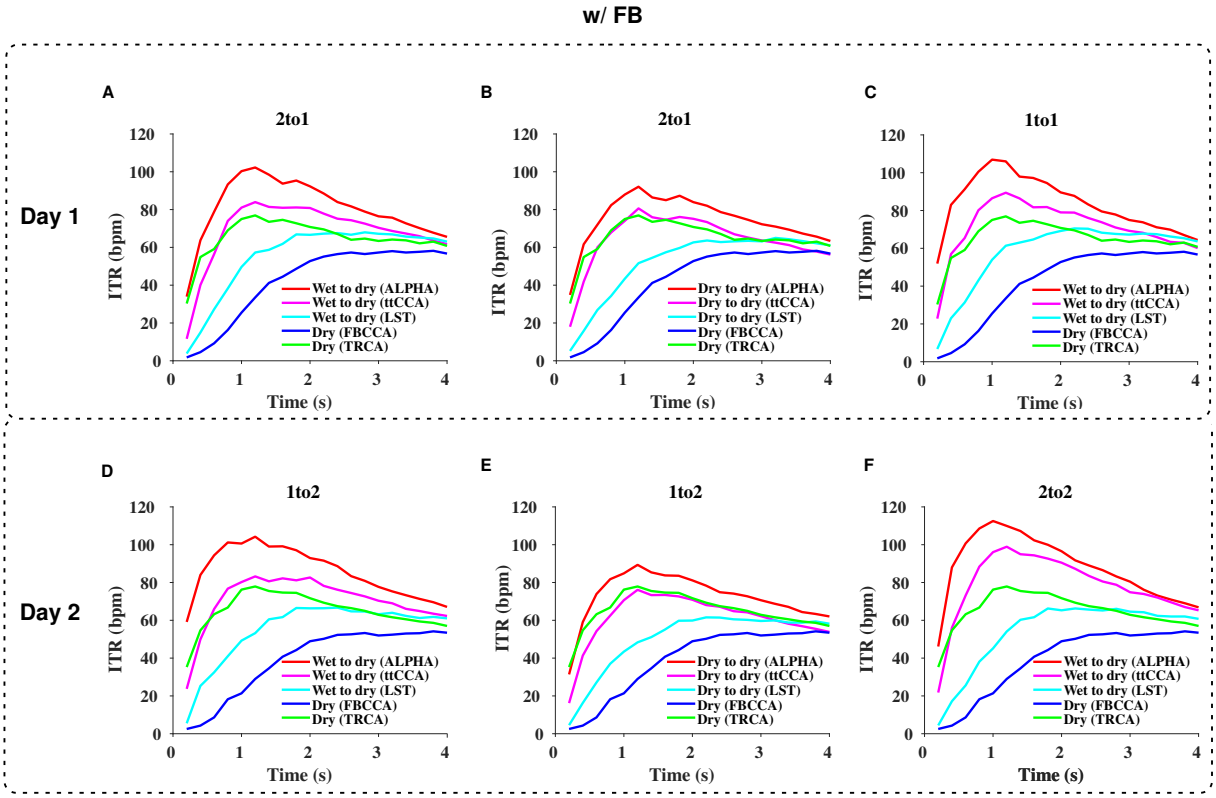


Fig. 2. (Supplementary Figure 2) The average information transfer rate (ITR) for the w/ FB scenario in cross-day cross-device transfer learning scenario. The upper panel denotes the transfer direction when day 1 is the target domain or test data. The lower panel denotes the transfer direction when day 2 is the target domain or test data. "AtoB" denotes transfer from day A to day B. Data lengths of 0.2 s to 4 s with an interval of 0.2 s were used for evaluation.

and day 2) were in a descending order, i.e., 109.73 bpm at 1 s, 103.24 bpm at 1.2 s, 90.71 bpm at 1.2 s, 77.40 at 1.2 s, respectively. For each of the 5 subjects, ALPHA consistently yielded better performance than LST and ttCCA at the data length of 1 s in various transfer directions.

The pilot study validates the feasibility of the proposed method (ALPHA) in the cross-day and cross-device transfer learning circumstance. And it implies that the within-day cross-device ALPHA could achieve the best performance in various transfer directions. Nevertheless, the cross-day wet-to-dry transfer via ALPHA is a competitive protocol for it could achieve better performance than other methods. The pilot study holds promise for the long-term use of SSVEP-BCI in a plug-and-play fashion, where the training data are recorded from wet-electrode headset on the first day and users can use the dry-electrode SSVEP-BCI system with zero training in the future days.

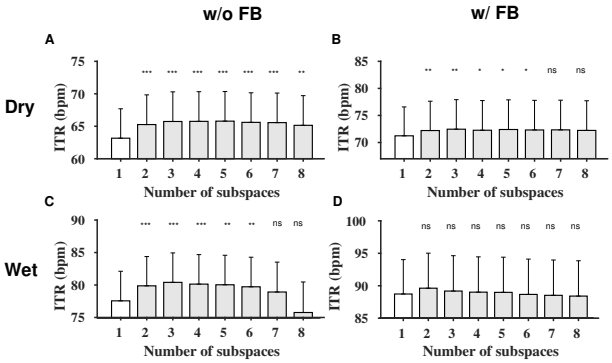


Fig. 3. (Supplementary Figure 3) The average ITR with a varying number of subspaces utilized in the subspace pooling. Data length used for evaluation corresponds to the maximum average ITR of each method. The asterisks indicate a significant difference between the case of one subspace and the case of more subspaces (*$p < .05$, **$p < .01$, ***$p < .001$, Bonferroni corrected).

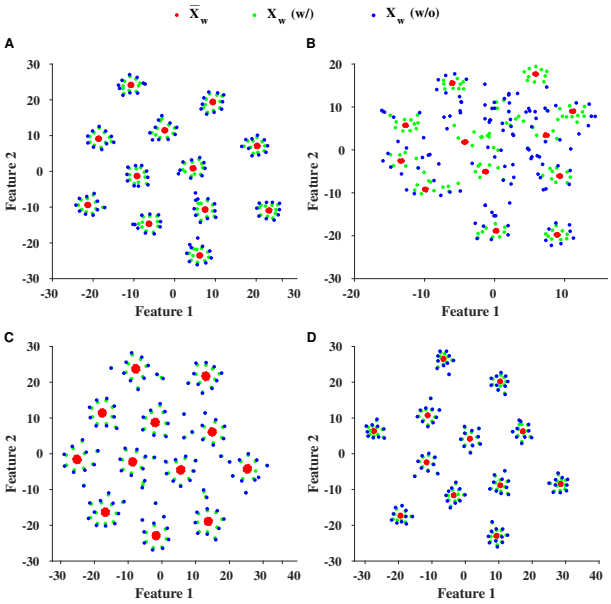


Fig. 4. (Supplementary Figure 4) The effect of subspace alignment (ASP) on the low dimensional feature space projected by t-SNE. The data length used for evaluation is 2s. The template embeddings ($\boldsymbol{W}_{\bar{X}[1]}^{T}\bar{\boldsymbol{X}}_s$) are represented by red dots. The test embeddings w/ ($\boldsymbol{W}_{\bar{X}[1]}^{ASP^T}\boldsymbol{X}_t$) and w/o ($\boldsymbol{W}_{\bar{X}[1]}^{T}\boldsymbol{X}_t$) subspace alignment are represented by green dots and blue dots, respectively. Each subfigure illustrates the data from a subject. Each dot represents a trial and each cluster formed by the dots represents a stimulus frequency or a class. The red dots are dense since the template embeddings have a high SNR in the source domain (wet-electrode EEG). The test embeddings are moved toward the template embeddings under the influence of the subspace alignment. (Best view in color.)