Exploiting Approximate Joint Diagonalization for Covariance Estimation in Imagined Speech Decoding *

Fotis P. Kalaganis[1][0000−0002−5474−9098], Kostas Georgiadis[1][0000−0001−9116−4729], Vangelis P. Oikonomou[1][0000−0001−7092−8486], Spiros Nikolopoulos[1][0000−0002−1367−5133], Nikos A. Laskaris[2][0000−0002−1960−394X], and Ioannis Kompatsiaris[1][0000−0001−6447−9020]

[1] Centre for Research & Technology Hellas, Information Technologies Institute (ITI), Thermi-Thessaloniki, Greece
[2] AIIA-Lab, Informatics Dept, AUTH, NeuroInformatics.Group, Thessaloniki, Greece
kalaganis@csd.auth.gr

Abstract. Recently, imagined speech has become a subject of study due to its potential as an intuitive communication system. It involves registering neural responses generated by mental speaking without moving the articulators. Although it may not perform as well as other paradigms, it has multiclass scalability, making it suitable for building extensible BCI systems. Hence, our study revolves around this intuitive paradigm that decodes human speech imagery from EEG signals using Riemannian geometry and a recently introduced covariance estimation method that is based on the concept of Approximate Joint Diagonalization (AJD). The employed methodological framework approach sets its grounds on neuroscientifically sound theories and is being validated on a competition dataset consisting of multichannel EEG trials from five different imagined prompts. Despite its simplicity, the presented methodology achieves over 70% accuracy in some classes, which is on par with State-of-the-Art performance on the dataset. Our methodology performs significantly better in monosyllabic prompts (i.e., 'yes' and 'stop') which may constitute it more appropriate in immediate-response critical BCI applications. Moreover, the conducted preliminary analysis that was used for sensor selection and onset detection sheds light into the understudied neural phenomena of imagined speech as captured in EEG signals.

Keywords: Riemannian Geometry · Imagined Speech · Speech Imagery · Covariance Estimator · Electroencephalography (EEG) · Approximate Joint Diagonalization (AJD) · Brain-Computer Interfaces.

1   Introduction

We, as human beings, keep talking within us most of the time. We rehearse over and over again how to manage a particular difficult situation, what to say to a prospective customer, how to answer certain critical questions in an interview, and so on. This speech, unlike the overt speech in a conversation with another person, is imagined and hence, there is no movement of the articulators. Thus, imagined speech is defined as the internal process of the voluntary imagination of speaking without actually moving any of the articulators.

Decoding and communicating human thought into the outside world, so as called, 'reading the mind' (i.e. to interpret internally-generated speech) has been a long-held ambition of Brain-Computer Interfaces (BCIs). Imagined speech has recently been studied as an intuitive paradigm [12], with the goal of decoding the neural responses generated via imagining pronunciation. This paradigm is particularly suitable for building communication systems and restoring communication for individuals that have lost the ability to speak (e.g. Stroke patients) as their ability to actively think or imagine speaking remains intact. While this paradigm currently lacks in terms of performance compared to other paradigms, it has multiclass scalability [8], thus showing the possibility of building extensible BCI systems [17] characterized by higher degrees of freedom. In this direction, imagined speech BCIs can be employed to mitigate the shortcomings of typical BCI paradigms (e.g., motor imagery, SSVEP, ERPs), like the limited number of distinct prompts/commands or the difficulty in training someone to use these systems [1], since they allow users to convey their intentions in a natural manner. Out of the neuroimaging methods currently available, Electroencephalography (EEG), is the one encountered the most in imagined speech BCIs as it is the least invasive and the most cost-effective. Although EEG may lack in terms of spatial resolution, compared to other neuroimaging technologies, it can reliably capture brain activity changes over shorter timescales (i.e. high temporal resolution).

Despite the recent efforts of the neuroscientific community [18], imagined speech recognition has proven to be a difficult task to achieve within an acceptable range of classification accuracy [14]. Concerning the particular case of EEG-based decoding schemes, a recent review study [20] uncovered that conventional Machine Learning algorithms (e.g., random forests, LDA, SVMs, etc.) are typically employed. Contemporarily, the aforementioned Machine Learning approaches are combined either with statistical or wavelet-based features. It was only until very recently that imagined speech decoding methods took advantage of modern Machine Learning schemes such as Deep Learning (mostly CNN architectures) [4, 7]. In the same direction, Riemannian geometry, which exhibits impressive decoding capabilities in other BCI paradigms, like motor imagery [5], neuromarketing [11] and P300 detection [19], has not been studied sufficiently in imagined speech. This fact indicates that this particular paradigm is still in its infancy and there is plenty of room for active research and improvement.

Hence, our study revolves around the intuitive paradigm of speech imagery as captured in EEG signals and aims to explore the potential of Affine Invariant Riemannian Metric (AIRM)-induced Riemannian geometry when combined

with a more delicate covariance estimation. Despite some recent efforts [22, 2], to the best of our knowledge this is the first study, in the context of imagined speech, that takes advantage of the equivalence between the sensor and source space as enabled by AIRM that is typically employed on Riemannian geometry settings for EEG decoding tasks. The employed covariance estimation is based on the Approximate Joint Diagonalization (AJD) of the spatial covariance matrices, which in essence, finds a common unitary matrix that allows us to approximately diagonalize all the spatial covariance matrices (e.g. each covariance matrix corresponds to a different trial which took place in the same recording session) simultaneously. Then, the obtained dominantly diagonal matrices are used, by considering only their diagonal entries and discarding the almost zero off-diagonal elements, to reconstruct the covariance matrices under a common mixing model (expressed by the common unitary matrix uncovered through AJD) [16]. Intuitively, this procedure serves as a denoising procedure on the spatial covariance matrices and leads to more robust estimates of the spatial covariance pattern.

The resulting covariance estimation is based on a preliminary analysis that was conducted in an effort to determine the most appropriate sensors, frequency ranges and time segments that the imagined speech phenomenon is taking place. This preliminary analysis takes advantage of the continuous wavelet transform and is performed in an unsupervised manner. By doing so, we were able to obtain a neuroscientifically informed idea about the imagined speech phenomena that remain understudied [24], as captured by means of EEG.

Our approach sets its ground on well-developed neuroscientific theories and is validated on a competition dataset (2020 International BCI Competition; Task 3 [14]) that contains multichannel EEG trials from five distinct imagined prompts. The employed methodological framework is characterized by simplicity (both in terms of the covariance estimation and the demployed machine learning scheme which is based on k-nearest neighbours) while its performance exceeds 70% accuracy in some classes which is on par to State-of-the-Art (SotA) performance on this particular dataset. As presented in the Results section the employed methodology performs significantly better in monosyllabic prompts (i.e., 'yes' and 'stop') which may constitute it more appropriate in immediate-response critical BCI applications. Moreover, the conducted preliminary analysis that was used for sensor selection and onset detection sheds light into the underlying phenomena of imagined speech. Overall, our approach ranks 2nd among the published competition results [14], a fact that further demonstrates its efficiency and neuroscientific validity.

2   Dataset

The original dataset (DOI: 10.17605/OSF.IO/PQ7VB) consists of 15 participants (S1-S15), aged between 20-30 years, who were instructed to silently imagine pronouncing five different words/phrases ("hello," "help me," "stop," "thank you," and "yes") without moving their mouth or making any sound. The par-

ticipants were seated comfortably in front of a 24-inch LCD monitor screen and were asked to avoid any other mental activity except for the task at hand. The imagination trials were conducted with a black screen to eliminate any external stimuli. An auditory cue for one of the five words/phrases was randomly presented for 2 seconds, followed by a presentation of a cross mark lasting between 0.8-1.2 seconds. The participants were instructed to begin their imagined speech immediately after the cross mark disappeared, and this cycle was repeated four times for each cue. After the four cycles, there was a 3-second relaxation phase to prepare for the next cue. Fig. 1 depicts the aforementioned experimental procedure's timeline.
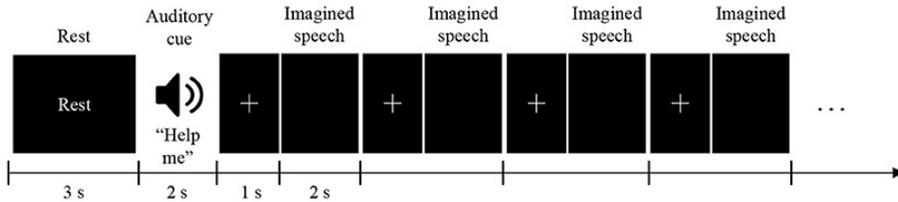
Fig. 1. Timeline of the experimental procedure followed during the imagined speech recording sessions. Image Source: [14]

A total of 400 trials per participant (80 trials per class) were recorded, out of which 60 trials per class are provided for training and 10 trials per class for validation purposes. The test data consist of 10 trials per class too. The train-validation-test split is provided in a subject-specific manner and all trials are belonging to a single recording session.

3   Methodology

3.1   Riemannian Geometry

Let us denote by $\mathbf{X}_i \in \mathbb{R}^{E \times T}, i = 1, \ldots, n$ a multichannel EEG trial, with E denoting the number of electrodes, T the number of samples in time and n the number of available trials. Each trial (assuming zero mean signals) can also be described by the corresponding spatial covariance matrix $\mathbf{C}_i = \frac{1}{T-1}\mathbf{X}_i\mathbf{X}_i^\top \in \mathbb{R}^{E \times E}$. Under a sufficiently large T value to guarantee a full rank covariance matrix, spatial covariance matrices are Symmetric and Positive Definite (SPD) that lie on a Riemannian manifold.

When dealing with EEG data, the manifold of SPD matrices denoted by $Sym_E^+ = \{\mathbf{C} \in \mathbb{R}^{E \times E} : \mathbf{x}^\top\mathbf{C}\mathbf{x} > 0, \text{ for all non-zero } \mathbf{x} \in \mathbb{R}^E\}$, is typically studied when it is endowed with the AIRM [23],

$$\langle \mathbf{A}, \mathbf{B} \rangle_\mathbf{P} \triangleq Trace(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}^{-1}\mathbf{B}) \tag{1}$$

for $\mathbf{P} \in Sym_E^+$ and $\mathbf{A}, \mathbf{B} \in T_E^+(\mathbf{P})$, where $T_E^+(\mathbf{P})$ denotes the tangent space of Sym_E^+ at $\mathbf{P}$. Then, the following geodesic distance is induced

$$\delta(\mathbf{C}_i, \mathbf{C}_j) = \left\| logm(\mathbf{C}_i^{-1/2}\mathbf{C}_j\mathbf{C}_i^{-1/2}) \right\|_F = \sqrt{\sum_{q=1}^{E} log^2\lambda_q} \qquad (2)$$

with $logm(\cdot)$ denoting the matrix logarithm operator and λ_q the eigenvalues of $\mathbf{C}_i^{-1/2}\mathbf{C}_j\mathbf{C}_i^{-1/2}$ or similarly of the matrix $\mathbf{C}_i^{-1}\mathbf{C}_j$. These two matrices hold the same eigenvalues while the indices i and j can be permuted.

As its name states, δ is affine-invariant for non singular matrices $\mathbf{W}$, i.e. $\delta(\mathbf{W}\mathbf{C}_i\mathbf{W}^\top, \mathbf{W}\mathbf{C}_j\mathbf{W}^\top) = \delta(\mathbf{C}_i, \mathbf{C}_j)$. This is an important property in EEG signal processing since it provides equivalence between the sensor and the source space [6]. According to the prevailing EEG model, the recorded activity is well approximated by a linear mixture of source signals. Hence, $\mathbf{X}_i = \mathbf{M}\mathbf{S}_i$ with $\mathbf{M}$ denoting the mixing matrix and $\mathbf{S}_i$ the source signals. Then, by substituting the observed signal with the equivalent mixing of sources, one may obtain the following covariance matrix, $\mathbf{C}_i = \frac{1}{T-1}\mathbf{M}\mathbf{S}_i\mathbf{S}_i^\top\mathbf{M}_i^\top$. Therefore, the mixing procedure in the time domain results in a congruent transformation in the corresponding covariance matrices. It becomes obvious that since δ is invariant to such transformations, the two spaces can be equivalently treated under the AIRM.

3.2   AJD-based Covariance Estimation

The mixing matrix, denoted as $\mathbf{M}$, is determined by the position and orientation of dipoles in the brain, the physical characteristics of the head, and the placement of electrodes on the scalp. It is therefore reasonable to assume that $\mathbf{M}$ remains constant for a certain period, such as during a single recording session. Assuming that sources are independent and the associated activity (i.e., source signals) are uncorrelated, the spatial covariance matrices of the sources are diagonal.

The process of estimating the mixing matrix, denoted as $\mathbf{M}$, from the observed sensor signals is an ill-posed problem known as Blind Source Separation (BSS) [21]. Two approaches are commonly used to tackle the BSS problem: The first approach is Independent Component Analysis (ICA), which aims to transform the data so that the components become as independent as possible [13]. An alternative approach involves using the diagonality of certain characteristic matrices derived from the data to approximate $\mathbf{M}^{-1}$ through the concept of AJD [26]. This involves finding an orthonormal change of basis denoted as $\mathbf{U}$, which makes the set of symmetric square matrices as diagonal as possible. This, second approach, intuitively uncovers the 'average eigenspace' of matrices that are approximately jointly diagonalizable [3].

Following the notation of section 3.1, we denote by $\mathbf{C}_i$ covariance matrix that corresponds to the EEG trial, $\mathbf{X}_i$. Let $\mathbf{U}$ be the orthonormal matrix calculated by AJD over the set of $\mathbf{C}_i$ with $i = 1, \ldots, n$ that estimates the mixing matrix $\mathbf{M}$. Then, each $\mathbf{C}_i$ can be transformed to a dominantly diagonal matrix through $\mathbf{U}^\top\mathbf{C}_i\mathbf{U}$. As such, we can reconstruct (i.e., re-estimate) all the spatial covariance

matrices under the constraint of a common eigenspace by using the formula $\tilde{\mathbf{C}}_i = \mathbf{U}diag(\mathbf{U}^\top \mathbf{C}_i \mathbf{U})\mathbf{U}^\top$. Here, the $diag(\cdot)$ operator, which discards the non-diagonal elements of a matrix and obtaining a strictly diagonal matrix, is applied upon an almost diagonal matrix and hence achieves a good re-estimation of the original covariance matrix.

This estimation approach forces all the spatial covariance matrices to admit a common mixing matrix and, hence, acts as a denoising procedure that abides to well-established neuroscientific theories. In addition, the estimated covariance matrices are guaranteed to hold the SPD property which allows the employment of Riemannian geometry. A more detailed description about the advantages and the mathematical properties of this covariance estimation can be found in [16].

4   Results

4.1   Preliminary Study: Spectrotemporal Analysis and Sensor Selection

Taking into account the high subject variability encountered in EEG data, a preliminary analysis for each subject was performed that aimed to identify the exact brain areas (i.e. sensors), timing (i.e. trial segments) and spectral components (i.e. frequency ranges) that the phenomenon of imagined speech takes place, with the scope of decoding the underlying phenomenon in the best possible way. In this context, a wavelet filter bank approach that disentangles the input signal into multiple frequency components without losing the signal's temporal characteristics is employed. It is noted that wavelets are characterized by time locality, allowing an efficient capture of transient behavior in a signal, which is of essence in the case of imagined speech decoding. Working on the training set for each subject independently, we applied the continuous wavelet transform (FBCWT, based on morse wavelet function and Matlab filter bank implementation) within the [1-100]Hz frequency range and derive the associated scalogram for each trial separately. Following the aforementioned procedure, all single-trial scalograms were averaged, regardless their label, to derive a spectrotemporal profile of activation for every sensor. Finally, using the baseline period, the mean and std of each scale was estimated and used to derive a threshold value (mean+3std) that in turn was employed to reveal the significant event-related spectral perturbations. The process is completed with detection of the sensors, segments and frequencies of interest based on the thresholding process.

Fig.2 illustrates the averaged FBCWT patterns for an indicative set of sensors for an exemplar subject (i.e. subject S1), after the thresholding process is completed. It is important to note here, that for clarity purposes only a selected number of sensors is presented. The visual inspection of the figure provides answers regarding the three research questions posed in this subsection. Starting from identification of the brain areas that the imagined speech phenomenon takes place, it is evident that the most informative sensors are located over the Broca's area (e.g., FT7, FT9 and T7), a trend that aligns well with what is reported in relevant bibliography regarding the brain areas activated during the
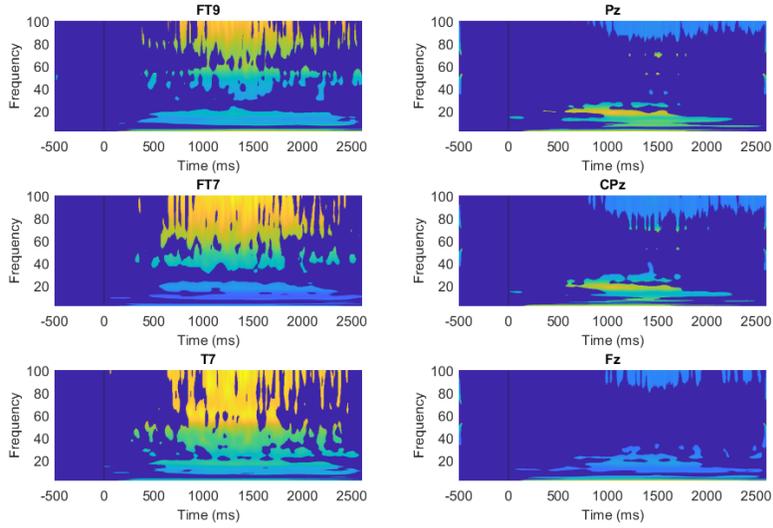
Fig. 2. Spectrotemporal analysis for the sensors characterized by the highest (left panel) and lowest (right panel) activation levels. The stimulus onset is indicated by the black vertical identified at t=0s and corresponds to cross disappearance as depicted in Fig. 1.

task of imagined speech [25]. On the contrary, the activation levels on sensors located over areas that are not associated with the mental speech task, like the middle area (e.g., sensors Pz, CPz and Fz), is significantly lower. Moving to the temporal domain, it is obvious that a reaction period of approximately 500ms is required before the mental imagery process is initiated by the participant, which is typical, while varying among individuals, when cue-based triggers mark the initiation of a task. Consequently, this process, upon appropriate modifications, can be employed as an onset detection procedure, which is of paramount importance in self-paced and online BCI paradigms. In the spectral domain, and specifically for the sensors characterized by high activity (such as FT7, FT9 and T7), three frequency ranges of interest can be identified: (i) Low ([5-20]Hz), (ii) Medium ([40-55]Hz) and, (iii) High ([$\geq$70]Hz), with the High frequency range being empirically identified, based on the validation set, as the one with the highest discriminative power. Finally, we should note that while the trends observed for the subject S1 are similar for the other subjects, the exact optimal sensors, segments and frequencies, as expected differ among them, showcasing the necessity and importance of this preliminary study.

4.2   Classification Results

Fig.3 presents the overall accuracy of the proposed decoder (Fig.3A, Fig.3B) and also the accuracy scores obtained for each subject independently (Fig.3C). In particular, the employed decoder is based on the estimation of covariance matrix (as described in section 3.2) from EEG signals in the frequency range above 70Hz while employing a Riemannian k-NN classifier where distance is calculated according to Eq. 2. By exploiting the validation set for each subject independently, $k = 3$ was identified as the most suitable value in terms of accuracy. We note that the provided test set is employed only for the purposes of obtaining and reporting the classification performance in this section.
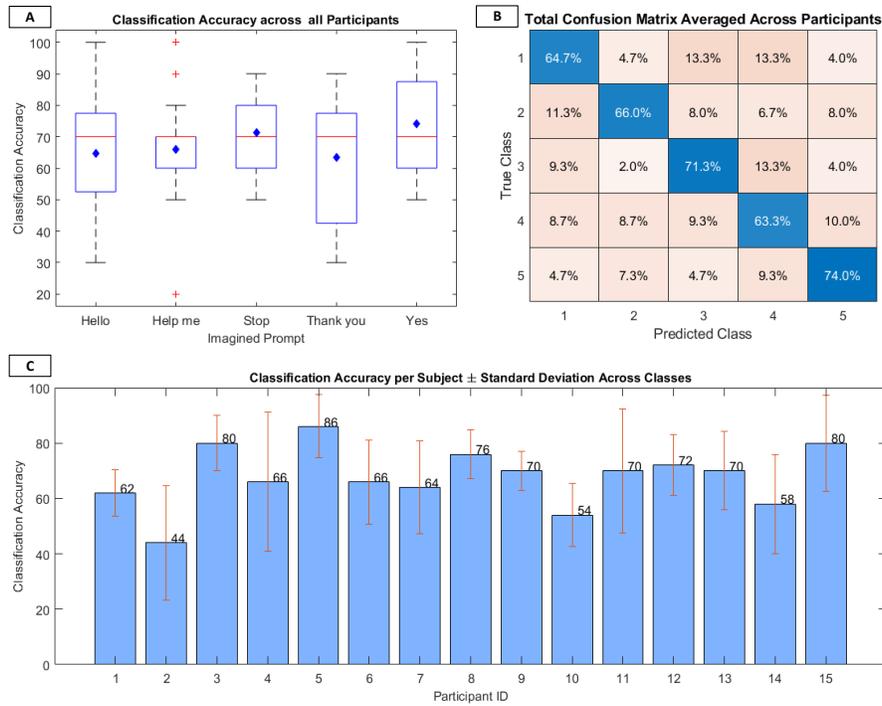
Fig. 3. The global and subject-wise performance of the proposed decoder. (A) The overall classification accuracy compartmentalized for each imagined prompt, (B) The total confusion matrix, and (C) The average classification accuracy per subject.

It is evident that despite the high subject variability, the majority of the subjects perform well when the imagined prompt is monosyllabic (see Fig.3A). This trend may imply that a different approach focusing on syllables rather than words may be required to better decipher the phenomenon of imagined

speech. In the same direction, disentangling the two prompts starting with the same syllable "He" (i.e. "Hello" and "Help me") seems highly challenging, given the high false positive values. Returning to the subject variability issue, while the accuracy for the majority of the subjects revolve around 70%, there are subjects with accuracy lower or barely exceeding 50% (i.e. S2, S10, S14), while there are also cases characterized by near-optimal performance (e.g. S3, S5). Considering the nature of the task (i.e. mental task) that in some cases may not be completely straightforward, it is not unlikely that some participants may require a familiarization period prior to the engagement with such tasks, as in the case of the motor imagery paradigm [9].

Despite the aforementioned, the proposed decoding scheme provides classification scores that significantly exceed the random level for this five class problem that comes at 20%. Finally, it must be noted that the achieved performance surpasses all but one the competitive approaches regarding the selected dataset [14]. Additionally, the employed AJD-based covariance estimator surpasses the classical covariance estimator, under the same classification setting, by 3.1% while exhibiting the same trends in class-specific classification results.

5 Discussion and Conclusion

In this paper we proposed a Riemannian geometry-based approach that relies on a delicate and neuroscientificaly valid estimation of the covariance matrix combined with tools of Riemannian geometry (as a result of endowing the SPD manifold with the AIRM). The obtained results demonstrate the effectivess on the employed scheme while showing the potential of Riemmanian geometry in the demanding task of EEG-based imagined speech decoding. As presented in the Results section our approach achieves State-of-the-Art results that are well-above the 20% random chance of the employed competition dataset and our presented approach is only surpassed by one competitor. It is important to note that while competition outcomes have been published [14], the corresponding methodologies deployed for the purposes of competition are not available. Hence the comparison presented in our work is confined to the classification performance and cannot be extended to more qualitative characteristics since they remain unknown.

In addition to the proposed decoding scheme our work presents a spectro-temporal preliminary analysis that investigates the frequency scales, time segments and sensors for the imagined speech phenomenon. Our findings suggest that the imagined speech takes place 500ms after the cue onset while the consistency among the wavelet-based activations paves the way for self-paced BCIs that fall under the imagined speech paradigm. Moreover, our preliminary analysis uncovered the involvement of three distinct frequency ranges in the speech imagery task. Although only the higher one (70Hz $\leq$) is exploited in the proposed decoding scheme, it becomes evident that all three could play a crucial role and therefore their effect in the decoding process should be further investigated. Moreover, our findings leave room for improvement by hinting that the combina-

tion and fusion of multiple frequency ranges may lead to more robust imagined speech decoders [9]. Ultimately, our preliminary analysis led to an unsupervised sensor selection procedure by means of keeping the sensors that exhibit the most powerful activations during the imagined speech task. Although the identified as the most 'informative' sensors are well-aligned with the existing literature, they may lack in terms of discriminability. Therefore, more suitable sensors selection approaches could be employed (e.g., [15, 10]) in an effort to achieve superior classification performance.

In more broad terms, imagined speech as a BCI paradigm is still in its infancy and several aspects should be explored thoroughly. As our study indicates some imagined prompts (i.e., monosyllabic words) are characterized by higher decoding robustness whereas close-to-echoing words seem to be conflated by the employed decoding scheme. Although particular claims cannot be made considering the extent of our study, valuable insights with respect to the most informative brain areas, physiology of the anticipated cortical activations and the most suitable prompts (e.g., words, phonemes or syllables) could be derived upon further exploration. The aforementioned will constitute the basis for future work towards conceptualizing novel decoding frameworks that will take advantage of the neural processes underpinning the imagined speech paradigm.

References

1. Abdulkader, S.N., Atia, A., Mostafa, M.S.M.: Brain computer interfacing: Applications and challenges. Egyptian Informatics Journal **16**(2), 213–230 (2015)
2. Bakhshali, M.A., Khademi, M., Ebrahimi-Moghadam, A., Moghimi, S.: Eeg signal classification of imagined speech based on riemannian distance of correntropy spectral density. Biomedical Signal Processing and Control **59**, 101899 (2020)
3. Cardoso, J.F., Souloumiac, A.: Jacobi angles for simultaneous diagonalization. SIAM journal on matrix analysis and applications **17**(1), 161–164 (1996)
4. Chengaiyan, S., Retnapandian, A.S., Anandan, K.: Identification of vowels in consonant–vowel–consonant words from speech imagery based eeg signals. Cognitive Neurodynamics **14**(1), 1–19 (2020)
5. Chu, Y., Zhao, X., Zou, Y., Xu, W., Song, G., Han, J., Zhao, Y.: Decoding multiclass motor imagery eeg from the same upper limb by combining riemannian geometry features and partial least squares regression. Journal of neural engineering **17**(4), 046029 (2020)
6. Congedo, M., Barachant, A., Bhatia, R.: Riemannian geometry for eeg-based brain-computer interfaces; a primer and a review. Brain-Computer Interfaces **4**(3), 155–174 (2017)
7. Cooney, C., Korik, A., Folli, R., Coyle, D.: Evaluation of hyperparameter optimization in machine and deep learning methods for decoding imagined speech eeg. Sensors **20**(16), 4629 (2020)
8. García-Salinas, J.S., Villaseñor-Pineda, L., Reyes-García, C.A., Torres-García, A.A.: Transfer learning in imagined speech eeg-based bcis. Biomedical Signal Processing and Control **50**, 151–157 (2019)
9. Georgiadis, K., Laskaris, N., Nikolopoulos, S., Kompatsiaris, I.: Connectivity steered graph fourier transform for motor imagery bci decoding. Journal of neural engineering **16**(5), 056021 (2019)

10. Georgiadis, K., Adamos, D.A., Nikolopoulos, S., Laskaris, N., Kompatsiaris, I.: A graph-theoretic sensor-selection scheme for covariance-based motor imagery (mi) decoding. In: 2020 28th European Signal Processing Conference (EUSIPCO). pp. 1234–1238. IEEE (2021)
11. Georgiadis, K., Kalaganis, F.P., Oikonomou, V.P., Nikolopoulos, S., Laskaris, N.A., Kompatsiaris, I.: Rneumark: A riemannian eeg analysis framework for neuromarketing. Brain Informatics **9**(1), 22 (2022)
12. Herff, C., de Pesters, A., Heger, D., Brunner, P., Schalk, G., Schultz, T.: Towards continuous speech recognition for bci. Brain-Computer Interface Research: A State-of-the-Art Summary 5 pp. 21–29 (2017)
13. Hyvärinen, A.: Survey on independent component analysis (1999)
14. Jeong, J.H., Cho, J.H., Lee, Y.E., Lee, S.H., Shin, G.H., Kweon, Y.S., Millán, J.d.R., Müller, K.R., Lee, S.W.: 2020 international brain–computer interface competition: A review. Frontiers in Human Neuroscience **16**, 898300 (2022)
15. Kalaganis, F.P., Laskaris, N.A., Chatzilari, E., Nikolopoulos, S., Kompatsiaris, I.: A riemannian geometry approach to reduced and discriminative covariance estimation in brain computer interfaces. IEEE Transactions on Biomedical Engineering **67**(1), 245–255 (2019)
16. Kalaganis, F.P., Laskaris, N.A., Oikonomou, V.P., Nikopolopoulos, S., Kompatsiaris, I.: Revisiting riemannian geometry-based eeg decoding through approximate joint diagonalization. Journal of Neural Engineering **19**(6), 066030 (2022)
17. Lee, S.H., Lee, M., Jeong, J.H., Lee, S.W.: Towards an eeg-based intuitive bci communication system using imagined speech and visual imagery. In: 2019 IEEE international conference on systems, man and cybernetics (SMC). pp. 4409–4414. IEEE (2019)
18. Lee, Y.E., Lee, S.H.: Eeg-transformer: Self-attention from transformer architecture for decoding eeg of imagined speech. In: 2022 10th International Winter Conference on Brain-Computer Interface (BCI). pp. 1–4. IEEE (2022)
19. Li, F., Xia, Y., Wang, F., Zhang, D., Li, X., He, F.: Transfer learning algorithm of p300-eeg signal based on xdawn spatial filter and riemannian geometry classifier. Applied Sciences **10**(5), 1804 (2020)
20. Lopez-Bernal, D., Balderas, D., Ponce, P., Molina, A.: A state-of-the-art review of eeg-based imagined speech decoding. Frontiers in Human Neuroscience **16** (2022)
21. Müller, K.R., Vigario, R., Meinecke, F., Ziehe, A.: Blind source separation techniques for decomposing event-related brain signals. International Journal of Bifurcation and Chaos **14**(02), 773–791 (2004)
22. Nguyen, C.H., Karavas, G.K., Artemiadis, P.: Inferring imagined speech using eeg signals: a new approach using riemannian manifold features. Journal of neural engineering **15**(1), 016002 (2017)
23. Pennec, X., Fillard, P., Ayache, N.: A riemannian framework for tensor computing. International Journal of computer vision **66**(1), 41–66 (2006)
24. Proix, T., Delgado Saa, J., Christen, A., Martin, S., Pasley, B.N., Knight, R.T., Tian, X., Poeppel, D., Doyle, W.K., Devinsky, O., et al.: Imagined speech can be decoded from low-and cross-frequency intracranial eeg features. Nature communications **13**(1), 48 (2022)
25. Si, X., Li, S., Xiang, S., Yu, J., Ming, D.: Imagined speech increases the hemodynamic response and functional connectivity of the dorsal motor cortex. Journal of Neural Engineering **18**(5), 056048 (2021)
26. Ziehe, A.: Blind source separation based on joint diagonalization of matrices with applications in biomedical signal processing. Ph.D. thesis, Universität Potsdam (2005)