

EEG-Based Decoding of Rhythm Imagery for Brain-Computer Musical Interfaces

Massimiliano Zanoni

Dept. of Electronics and Telecommunications
Politecnico di Torino, Turin, Italy
massimiliano.zanoni@polito.it

Andrea Eugenio Losi

Conservatorio F.A. Bonporti
Trento, Italy
andreaeugenio.losi@mail.polimi.it

Cristina Rottondi

Dept. of Electronics and Telecommunications
Politecnico di Torino, Turin, Italy
cristina.rottondi@polito.it

Abstract—Brain-Computer Musical Interfaces (BCMIs) represent an emerging class of interactive systems in which neural activity complements conventional audio and gesture-based control. Within the broader vision of the Internet of Sounds (IoS), BCMIs have the potential to enable silent musical interaction, accessible performance technologies, and multimodal communication paradigms in which musical intentions are inferred directly from brain activity. While previous research has extensively investigated motor imagery, auditory perception, and generic music imagination, the decoding of internally generated rhythmic structures remains largely unexplored despite rhythm constituting one of the fundamental components of musical performance.

This paper investigates whether distinct imagined rhythmic patterns can be discriminated from electroencephalographic (EEG) recordings acquired from trained percussionists. A controlled experimental protocol was designed in which participants imagined three rhythmic patterns while synchronizing to a metronome, providing repeatable temporal alignment across trials. Neural synchronization was analysed through phase locking value, while rhythm classification was performed using the lightweight EEGNet architecture.

Experimental results demonstrate measurable neural synchronization associated with rhythm imagery and show that subject-specific classifiers achieve accuracies substantially above chance level. Although the study represents a pilot investigation involving a limited number of participants, the results suggest that imagined rhythmic patterns contain discriminative neural signatures that may support future BCMIs.

Index Terms—Brain-Computer Musical Interface (BCMI); Electroencephalography (EEG); Neural Decoding; Rhythm Imagery; EEGNet; Phase Synchronization

I. INTRODUCTION

Recent advances in networking, artificial intelligence, immersive multimedia, and wearable sensing are rapidly transforming digital music technologies into interconnected ecosystems in which performers, audiences, intelligent agents, and distributed devices interact seamlessly in real time. Within this vision, often referred to as the *Internet of Sounds* (IoS) [1], musical interaction extends beyond conventional audio transmission to integrate heterogeneous modalities—including gesture, physiological sensing, visual information, haptic feedback, and semantic information—enabling richer and more expressive

forms of music creation, performance, and collaboration. As these environments become increasingly multimodal, novel interaction paradigms capable of capturing users' intentions more naturally are expected to play a central role.

Among these paradigms, Brain-Computer Musical Interfaces (BCMIs) represent a promising direction by exploiting neural activity as an additional interaction modality for musical systems [2], [3]. Unlike traditional Digital Musical Instruments (DMIs), which primarily rely on keyboards, MIDI controllers, touch interfaces, or body movements, BCMIs aim to infer musical intentions directly from brain activity, opening new opportunities for silent interaction, adaptive musical interfaces, accessible performance technologies, and future multimodal IoS applications. Besides improving accessibility for users with motor impairments, such interfaces also offer able-bodied musicians new possibilities for externalizing creative processes that are inherently internal, such as rhythm imagination and musical planning.

Electroencephalography (EEG) provides a non-invasive means of investigating these internally generated musical processes. However, despite extensive research on neural responses to music and motor processes, existing EEG-based approaches exhibit two important limitations. On one hand, a large body of work focuses on externally driven paradigms, including music listening, auditory attention decoding, and visually guided musical tasks such as score reading. These studies [4], [5] have successfully characterized neural responses associated with auditory perception and continuous music processing, but remain inherently constrained by their reliance on external stimuli, thus providing only limited insight into internally generated musical representations.

On the other hand, studies investigating endogenous neural activity have predominantly relied on classical Brain-Computer Interface (BCI) paradigms such as motor imagery [6], [7]. While these approaches have demonstrated reliable decoding of internally generated neural activity, they are generally designed to recognize discrete mental commands (e.g., left- versus right-hand imagery) rather than temporally evolving cognitive processes. Consequently, they are not well suited to representing rhythmic structures, whose musical meaning emerges from the temporal organization of successive events rather than from isolated mental states.

More recently, several studies have demonstrated that EEG

The authors acknowledge the support of the MUSMET project funded by the EIC Pathfinder Open scheme of the European Commission (grant agreement n. 101184379). This manuscript reflects only the authors' views and opinions and the European Union or the European Innovation Council cannot be considered responsible for them.

contains information related to higher-level musical cognition. Neural decoding techniques have successfully reconstructed continuous musical features from EEG recordings, while music imagery experiments have revealed reproducible neural signatures associated with internally generated melodies and imagined musical content [8]–[10]. Large-scale neuroimaging investigations have further shown that music perception and imagination recruit distributed cortical networks involving auditory, motor, frontal, and parietal regions [11]. Likewise, neurophysiological studies on rhythm perception have demonstrated the close interaction between auditory and motor systems during beat tracking and temporal prediction [12], [13], whereas rhythm imagery has been shown to induce measurable neural synchronization even in the absence of overt movement [14], [15]. Nevertheless, the practical problem of decoding distinct imagined rhythmic patterns for interactive musical applications remains largely unexplored.

This gap motivates the need for approaches specifically targeting internally generated and temporally structured musical representations. In this work, rhythm imagination is regarded not only as a cognitive phenomenon but also as a potential interaction modality for future BCMI. Rather than investigating generic music imagery, we focus on internally generated rhythmic patterns produced by trained percussionists, exploiting the close relationship between rhythm, motor planning, auditory imagery, and sensorimotor coordination.

To this end, EEG recordings are analyzed through complementary signal processing and deep learning techniques. Independent Component Analysis (ICA) is first employed to isolate neural activity associated with rhythm imagination while reducing physiological and environmental artifacts. The resulting independent components are subsequently exploited in two complementary analyses. First, Empirical Mode Decomposition (EMD) is employed to investigate the temporal correspondence between brain activity and imagined rhythmic structures, providing an interpretable characterization of rhythm-related neural synchronization. Second, rhythm classification is performed using EEGNet [16], a compact convolutional neural network specifically designed for EEG decoding.

Unlike previous studies primarily aimed at understanding the neural correlates of rhythm perception, the objective of this work is to evaluate whether imagined rhythmic patterns contain sufficiently discriminative neural signatures to support practical BCMI. By combining interpretable signal analysis with deep learning classification, the proposed framework provides both insight into the neural dynamics of rhythm imagery and an initial assessment of its feasibility as an interaction modality for future IoS applications.

The main contributions of this work are summarized as follows:

- the design of a controlled EEG protocol for rhythm imagery involving trained percussionists;
- an interpretable analysis of neural synchronization during rhythm imagination through ICA and EMD;
- the evaluation of EEGNet for subject-specific classification of imagined rhythmic patterns;

- a discussion of the opportunities and current limitations of EEG-based rhythm decoding for future BCMI and IoS applications.

The remainder of the paper is organized as follows. Section II reviews the state of the art in BCMI, EEG-based music imagery, and rhythm decoding. Section III discusses the neurophysiological foundations of rhythm imagery, while Section IV-A presents the experimental protocol and EEG acquisition procedure. The proposed signal processing and classification framework is described in Section V, and its performance is evaluated in Section VI through both neural synchronization analysis and rhythm classification experiments. Finally, Section VII concludes the paper and outlines future research directions.

II. RELATED WORK

A. Brain–Computer Musical Interfaces

BCMI extend conventional BCIs by exploiting neural activity as an interaction modality for musical applications. Unlike traditional DMIs, which primarily rely on physical gestures or electromyographic (EMG) signals, BCMI establish a direct communication pathway between users and musical systems through EEG activity. Early BCMI mainly employed motor imagery, event-related potentials, and steady-state visually evoked potentials to control musical parameters or generate adaptive soundscapes. More recently, the field has expanded towards affective computing, neurofeedback, and adaptive music generation, enabling musical systems capable of responding to users’ cognitive and emotional states [2], [3].

Within the broader vision of the IoS, BCMI are increasingly regarded as complementary sensing modalities rather than standalone interaction techniques. By combining EEG with conventional audio, gesture, motion capture, and physiological sensing, future multimodal musical systems may achieve more expressive, accessible, and context-aware interaction paradigms.

B. EEG-based Music Perception and Imagery

The relationship between music and brain activity has received considerable attention in recent years. EEG has been successfully employed to investigate music perception, auditory attention, musical imagery, and neural tracking of continuous musical stimuli. Neural encoding models have demonstrated that cortical activity follows temporal characteristics of musical signals, enabling reconstruction of musical features directly from EEG recordings [8]. Likewise, several studies have shown that internally generated musical representations, including imagined melodies and musical phrases, produce reproducible neural signatures that can be decoded using machine learning approaches [9], [10].

Large-scale neuroimaging studies have further revealed that music perception engages distributed cortical networks involving auditory, motor, frontal, and parietal regions rather than isolated functional areas [11]. These findings support the hypothesis that internally generated musical representations

constitute measurable neural processes suitable for future BCMI.

C. Neural Correlates of Rhythm Perception and Imagery

Among the different musical dimensions, rhythm plays a central role in temporal organization, synchronization, and sensorimotor coordination. Neurophysiological investigations have demonstrated that beat perception and temporal prediction involve tightly coupled auditory and motor cortical networks, even in the absence of overt movement [12], [17], [18]. Rhythm imagery similarly recruits distributed neural mechanisms responsible for maintaining internal temporal representations while synchronizing imagined events with an external temporal reference.

Previous studies have investigated neural synchronization during rhythm imagination using EEG and magnetoencephalography, reporting significant phase synchronization and oscillatory activity associated with internally generated rhythmic structures [14] [15]. Nevertheless, existing work has primarily focused on understanding the neural mechanisms underlying rhythm cognition, whereas considerably less attention has been devoted to decoding imagined rhythmic patterns as an interaction modality for musical interfaces.

D. Deep Learning for EEG Decoding

Traditional EEG decoding approaches have relied on hand-crafted temporal, spectral, or spatial features combined with conventional machine learning classifiers. More recently, deep neural networks have demonstrated remarkable performance by jointly learning discriminative feature representations directly from raw or minimally processed EEG signals.

Among these architectures, EEGNet has become one of the most widely adopted models owing to its compact architecture, low computational complexity, and ability to generalize across multiple EEG decoding paradigms [16]. By combining temporal convolutions with depthwise and separable spatial filtering, EEGNet efficiently captures both temporal dynamics and spatial dependencies while requiring substantially fewer trainable parameters than conventional convolutional neural networks.

Motivated by these characteristics, the present work adopts EEGNet to investigate whether internally imagined rhythmic patterns can be discriminated from EEG recordings acquired from trained percussionists. Unlike previous studies focusing primarily on neuroscientific interpretation, our objective is to evaluate rhythm imagery as a practical interaction modality for future BCMI.

III. RHYTHM IMAGERY AND NEURAL FOUNDATIONS

Rhythm constitutes one of the most fundamental components of musical cognition, supporting temporal organization, synchronization, prediction, and coordinated motor behavior during musical performance. Unlike isolated auditory events, rhythmic structures unfold over time and therefore require the brain to continuously generate internal temporal expectations. Consequently, rhythm imagination cannot be regarded as a

purely auditory phenomenon but rather emerges from the interaction between auditory and motor cortical systems.

Recent neurophysiological studies have shown that imagined rhythmic sequences recruit many of the same neural circuits involved in rhythmic performance. In particular, internally generated rhythms are sustained through top-down interactions from motor planning areas towards auditory regions, allowing temporal predictions to be maintained even in the absence of external acoustic stimulation [9], [15]. This predictive mechanism distinguishes rhythm imagery from passive listening, where sensory information primarily drives cortical activation.

A central component of this process is the close coupling between auditory perception and motor planning. Sensorimotor synchronization, observed for example during finger tapping or percussion performance, reflects the bidirectional interaction between these systems, enabling accurate temporal coordination and beat prediction [14]. Rather than merely responding to auditory events, motor regions actively participate in generating internal temporal representations that guide both perception and action.

Several cortical and subcortical structures contribute to these mechanisms. The Supplementary Motor Area (SMA) and pre-Supplementary Motor Area (pre-SMA) play a major role in internally maintaining beat representations and temporal prediction, while the basal ganglia contribute to relative timing processes underlying rhythmic entrainment. In contrast, the cerebellum primarily supports absolute timing and interval estimation [13]. Together, these structures generate oscillatory neural activity synchronized with internally represented rhythmic events.

This synchronization phenomenon, commonly referred to as neural entrainment, represents one of the principal neurophysiological signatures of rhythm processing. Previous studies have shown that phase synchronization persists during rhythm imagination even without external auditory stimulation, suggesting that internally generated rhythms produce measurable temporal organization of cortical oscillations [12], [14], [15]. Such synchronization provides the physiological basis for investigating rhythm imagery using EEG.

Motivated by these observations, the present work investigates whether the neural signatures generated during rhythm imagination contain sufficiently discriminative information to support practical BCMI. Rather than focusing exclusively on neuroscientific interpretation, the following sections exploit these synchronization mechanisms to develop an EEG-based rhythm decoding framework for future IoS applications.

IV. METHOD

This section presents the experimental methodology adopted to investigate rhythm imagery through EEG recordings. After describing the experimental protocol and the data acquisition procedure, the organization of the recorded data and the two analysis pipelines are introduced. The signal processing algorithms composing these pipelines are described in detail in Section V.



Fig. 1: Rhythmic patterns adopted during the experimental protocol.

A. Experimental Protocol

Selection of participants. The measurement campaign involved four experienced percussionists. Participants with extensive rhythmic training were intentionally selected because percussion practice strengthens auditory–motor coupling and supports the formation of stable internal temporal representations, both of which are fundamental for rhythm synchronization and imagery tasks [12], [13], [19].

Choice of rhythmic patterns.

Three simple metric rhythms were adopted to facilitate accurate mental reproduction while limiting cognitive load. Rhythms characterized by regular perceptual accents and integer temporal ratios naturally align with beat-based timing mechanisms [12]. Furthermore, high pulse clarity improves temporal discrimination and neural synchronization during rhythm processing [11], [20]. The three patterns were intentionally selected to share the same meter, tempo, duration, and a comparable level of rhythmic complexity, while differing in the temporal organization of rhythmic events and rests. Their relative similarity provides a controlled experimental setting in which differences associated with markedly different metrical structures or levels of complexity are limited, allowing us to investigate whether subtle differences in imagined rhythmic organization are reflected in the EEG signals. The rhythmic patterns used throughout the experiments are illustrated in Fig. 1.

Tempo selection. The experiment was performed at a tempo of 60 BPM, corresponding to a beat period of 1 s and an eighth-note subdivision every 500 ms. This temporal scale is consistent with previous studies on beat perception and facilitates stable synchronization during rhythm imagery [13], [21]. Since each rhythmic pattern spans one measure in a 4/4 meter, each imagery trial lasts four seconds and naturally defines one EEG epoch.

Experimental setup. Before EEG acquisition, participants rehearsed each rhythmic pattern between 20 and 100 times on a rubber practice pad in order to reinforce motor familiarity and stabilize the corresponding internal representation [22]. EEG recordings were then acquired in a quiet environment while participants remained seated with their eyes closed and mentally reproduced the rhythms without performing overt movements. Participants were instructed to imagine both the auditory feedback and the kinesthetic sensation associated with percussion performance.

A metronome operating at 60 BPM provided a common temporal reference throughout the experiment. Trigger events synchronized with the metronome were simultaneously recorded together with the EEG signals, allowing the contin-

TABLE I: Number of valid epochs retained after the cleaning procedure.

Participant	R1	R2	R3
S1	128	76	73
S2	90	75	104
S3	63	69	65
S4	31	86	44

uous recordings to be segmented into consecutive 4-s epochs corresponding to complete imagined rhythmic patterns.

For each participant and each rhythm, three recording sessions were acquired. Participants continued the imagery task until reporting mental fatigue or inability to maintain sufficient concentration. Since rhythm conditions were acquired in separate recording sessions, possible temporal-order effects cannot be completely excluded. Consequently, the present work should be regarded as a proof-of-concept investigation, while future studies will employ randomized or block-randomized presentation of rhythmic conditions together with cross-session validation.

Technical setup. EEG signals were acquired using the SD LTM 32/64 EXPRESS system (Micromed), equipped with 64 channels, of which 60 dedicated to cerebral activity were considered in the subsequent analyses. Signals were sampled at 512 Hz, while metronome triggers were simultaneously recorded to enable accurate temporal alignment between the imagined rhythmic cycles and the EEG recordings.

B. Data Processing Pipeline

Before feature extraction, all recordings underwent a preliminary cleaning procedure to improve EEG signal quality and retain only epochs corresponding to stable rhythm imagery. Initial epochs associated with the synchronization phase with the metronome were discarded, since they preceded the actual imagery task and could reflect adaptation processes rather than stable internal rhythm generation. In addition, epochs affected by evident movement artifacts or excessive noise were removed following visual inspection by an expert operator.

For each participant and each rhythmic condition, the remaining valid epochs collected across the three recording sessions were concatenated into a single cleaned sequence. The final number of valid epochs retained after the cleaning procedure is reported in Table I.

The cleaned EEG recordings were subsequently processed through two complementary analysis pipelines, corresponding to the two objectives of the present study.

The first pipeline aims at investigating the neurophysiological correlates of rhythm imagery. For each participant, epochs belonging to the same rhythmic condition were concatenated and subsequently analyzed through ICA, EMD, and Phase Locking Value (PLV) analysis in order to quantify neural synchronization with the imagined rhythmic structures.

The second pipeline addresses the rhythm decoding problem. In this case, the cleaned epochs belonging to the three

rhythmic conditions were combined into a single subject-specific dataset, allowing the classifier to learn discriminative neural representations across all rhythm classes. After concatenation, epochs data were analyzed through ICA. The resulting dataset was randomly partitioned into training (80%), validation (15%), and test (5%) subsets.

The adopted partition was intentionally selected to maximize the amount of training data available given the limited size of the pilot dataset while preserving independent validation and testing sets for model optimization and final performance evaluation. Hyperparameter tuning was performed independently for each participant, since the objective of the present study is to assess the feasibility of subject-specific rhythm decoding rather than to develop a subject-independent classifier.

Because the present work constitutes a proof-of-concept investigation involving a limited number of recordings, cross-session validation was not considered. Consequently, the reported results should be interpreted as an initial assessment of the discriminative information contained in rhythm imagery rather than as an estimate of population-level generalization. Future work will investigate larger participant cohorts together with randomized rhythm presentation, cross-session validation, and subject-independent decoding strategies.

The detailed signal processing algorithms composing the two pipelines are described in the following section.

V. SIGNAL PROCESSING FRAMEWORK

Starting from the cleaned EEG epochs obtained after the data cleaning stage described in Section IV-B, the proposed framework consists of two complementary processing pipelines sharing the same initial preprocessing stages. Both pipelines include (i) EEG preprocessing and (ii) Independent Component Analysis (ICA).

The first pipeline aims at characterizing the neurophysiological correlates of rhythm imagery. Following EEG preprocessing and ICA, the retained neural components are further processed through (iii) Empirical Mode Decomposition (EMD) and (iv) Phase Locking Value (PLV) analysis to quantify neural synchronization with the imagined rhythmic structures.

The second pipeline addresses rhythm classification. Following EEG preprocessing and ICA, the resulting EEG signals are directly provided as input to (iii) EEGNet to evaluate whether imagined rhythmic patterns contain sufficiently discriminative neural information to support automatic rhythm classification.

A. EEG Preprocessing

Raw EEG recordings were first visually inspected by an expert operator to identify corrupted channels and recording anomalies. Subsequently, the continuous recordings were segmented into consecutive 4s epochs according to the trigger events generated by the metronome, so that each epoch corresponded to one complete imagined rhythmic pattern.

To improve signal quality while preserving the temporal dynamics associated with rhythm imagery, each epoch was

subjected to a standard EEG preprocessing pipeline. A band-pass filter was applied to retain the frequency components relevant for cortical activity while suppressing slow drifts and high-frequency noise. In addition, a notch filter was employed to attenuate power-line interference.

The resulting multichannel EEG epochs constitute the input to the subsequent blind source separation stage.

B. Independent Component Analysis

EEG recordings represent linear mixtures of multiple neural and non-neural sources. Besides cortical activity, EEG channels inevitably contain artifacts generated by eye movements, muscle contractions, cardiac activity, and environmental interference. Since these artifacts may significantly affect both synchronization analysis and rhythm classification, an additional source separation stage was introduced before feature extraction.

ICA was adopted to decompose the multichannel EEG recordings into statistically independent latent sources. Under the standard linear mixing model,

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad (1)$$

where $\mathbf{x}(t)$ denotes the observed EEG channels, $\mathbf{s}(t)$ represents the unknown independent neural sources, and $\mathbf{A}$ is the mixing matrix.

ICA estimates an unmixing matrix

$$\mathbf{W} \approx \mathbf{A}^{-1}, \quad (2)$$

allowing the independent components to be recovered as

$$\hat{\mathbf{s}}(t) = \mathbf{W}\mathbf{x}(t). \quad (3)$$

The extracted independent components were subsequently classified using ICLabel [23], an automated component-labeling framework that assigns probabilistic estimates to several component categories, including Brain, Muscle, Eye, Heart, Line Noise, Channel Noise, and Other.

Only components whose probability of belonging to the Brain class exceeded 90% were retained for further analysis, whereas the remaining components were discarded as likely artifacts or non-neural sources. The cleaned EEG signals were finally reconstructed from the retained components and subsequently employed in both the neural synchronization and rhythm classification pipelines.

Compared with directly processing the original EEG channels, ICA provides a cleaner representation of cortical activity while reducing the influence of volume conduction and common artifacts.

C. Empirical Mode Decomposition

To investigate the temporal organization of rhythm-related neural activity, the retained ICA components were further analyzed using EMD.

TABLE II: Training Parameters and Architecture Configuration by Subject

Subject	Number of training epochs	Learning rate	Number of temporal filters	Number of spatial filters
S1	400	0.001911	4	12
S2	400	0.000635	9	9
S3	400	0.001911	4	12
S4	400	0.001911	4	12

Unlike conventional Fourier-based techniques, EMD does not assume signal stationarity or linearity, making it particularly suitable for EEG analysis. The method iteratively decomposes a signal into a finite set of Intrinsic Mode Functions (IMFs),

$$x(t) = \sum_{k=1}^K IMF_k(t) + r(t), \quad (4)$$

where $r(t)$ denotes the residual trend.

Each IMF captures oscillatory activity occurring at a characteristic temporal scale, allowing rhythm-related dynamics to be analyzed independently across different frequency bands. The resulting decomposition provides an interpretable representation of neural oscillations while preserving the intrinsic temporal variability of EEG signals.

The extracted IMFs were subsequently employed to investigate the correspondence between internally imagined rhythmic structures and the temporal organization of cortical oscillations.

D. Neural Synchronization Analysis

The first objective of the proposed framework is to investigate whether rhythm imagination produces measurable neural synchronization across repeated trials.

Synchronization was quantified through PLV, which measures the temporal consistency of the phase relationship between two oscillatory signals independently of their amplitudes. For two instantaneous phases $\phi_1(t)$ and $\phi_2(t)$, the PLV is defined as

$$PLV = \left| \frac{1}{N} \sum_{n=1}^N e^{j(\phi_1(n) - \phi_2(n))} \right|, \quad (5)$$

where N denotes the number of temporal samples.

PLV values range from 0 to 1, where values close to one indicate highly stable phase synchronization, whereas values close to zero correspond to weak or absent synchronization.

Although PLV is known to be influenced by volume conduction, its widespread use in rhythm perception and music cognition studies enables direct comparison with previous literature. Future work will investigate complementary connectivity measures, including the Phase Lag Index (PLI), which is generally considered more robust against zero-lag synchronization effects.

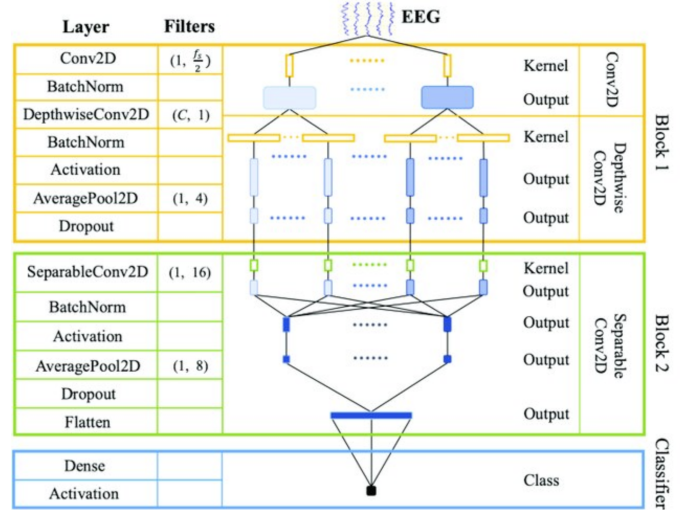


Fig. 2: EEGNet architecture summary for processing EEG signals.

E. Rhythm Classification using EEGNet

Beyond neurophysiological interpretation, the second objective of this work is to determine whether imagined rhythmic patterns can be automatically discriminated from EEG recordings.

To this end, rhythm classification was performed using EEGNet [16], a compact convolutional neural architecture specifically designed for EEG decoding. As reported in Fig. 2, EEGNet combines temporal convolutions with depthwise and separable spatial convolutions, enabling simultaneous extraction of temporal dynamics and spatial dependencies while maintaining a relatively small number of trainable parameters.

The cleaned EEG epochs obtained after ICA constitute the network input. Each epoch corresponds to one complete imagined rhythmic pattern and is associated with one of the three rhythm classes introduced in Section IV-A. Subject-specific models were trained independently for each participant in order to evaluate the discriminative information contained in individual rhythm imagery recordings while minimizing inter-subject variability.

EEGNet was selected as a suitable compromise between classification performance, computational complexity, and interpretability. Its compact architecture relies on temporal convolutions to capture temporal EEG dynamics, depthwise convolutions to learn spatial filters across electrodes, and separable convolutions to efficiently combine the extracted features. This design substantially reduces the number of trainable parameters compared with conventional CNN architectures, while preserving the ability to learn both temporal and spatial representations directly from EEG signals. Moreover, the explicit separation between temporal and spatial filtering provides a more interpretable representation of the learned EEG features than generic convolutional architectures. These characteristics make EEGNet particularly suitable for

the present proof-of-concept study, where the available dataset is limited, and represent a favorable property for future BCMI applications requiring lightweight EEG decoding. EEGNet training parameters used in this study are reported in Table II.

VI. RESULTS AND DISCUSSION

The proposed framework was evaluated from two complementary perspectives. First, we investigated whether internally imagined rhythmic patterns induce measurable neural synchronization by analyzing the phase relationship between EEG oscillations and rhythm-specific temporal references. Subsequently, we assessed whether these neural signatures contain sufficient discriminative information to enable automatic rhythm classification using EEGNet. Unless otherwise stated, all reported results correspond to subject-specific analyses performed independently for each participant.

A. Neural Synchronization Analysis

The first objective of the experimental evaluation is to determine whether rhythm imagination produces consistent neural synchronization reflecting the temporal organization of the imagined rhythmic patterns. Following the methodology described in Section V-D, EEG signals were first decomposed through EMD, and the three lowest-frequency IMFs were retained, as they capture the slow oscillatory dynamics most closely associated with rhythm processing.

For each imagined rhythm, the PLV was computed between each IMF and the corresponding rhythm-specific reference signal, represented by a rhythm-tracking frequency-modulated sinusoid. For every epoch, only the maximum PLV among the three IMFs was retained, under the assumption that the most synchronized component best reflects rhythm-related neural entrainment.

To quantify synchronization consistency, the proportion of epochs whose PLV exceeded a set of predefined thresholds was computed for each participant and subsequently averaged across subjects. Standard deviations were also calculated to quantify inter-subject variability. The resulting distributions are reported in Fig. 3(a).

As a reference condition, the same analysis was repeated using a swept-sine control signal whose instantaneous frequency linearly increases from the beat frequency to its double. While preserving the same frequency range, this baseline removes the temporal organization characteristic of the rhythmic patterns, allowing rhythm-specific synchronization to be distinguished from generic oscillatory coupling. The corresponding results are shown in Fig. 3(b).

At low PLV thresholds ($PLV \geq 0.1$ and $PLV \geq 0.3$), a large proportion of epochs exceeds the selected thresholds for all three rhythms. Since high exceedance rates are expected when relatively low thresholds are adopted, these values alone should not be interpreted as evidence of strong phase synchronization. More informative is the comparison with the swept-sine baseline: across the considered thresholds, the rhythm-specific reference signals generally yield higher proportions

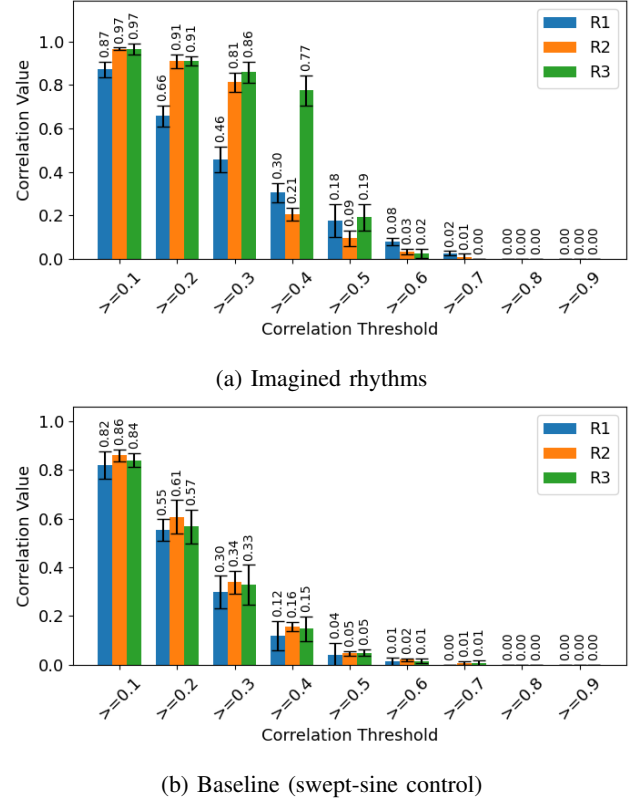


Fig. 3: Comparison of neural synchronization across imagined rhythmic patterns. (a) Percentage of epochs exceeding different PLV thresholds. (b) Corresponding swept-sine baseline. Bars represent mean $\pm$ standard deviation across participants.

of epochs exceeding the PLV thresholds than the control condition. This difference suggests that the observed phase locking is related to the temporal organization of the imagined rhythmic patterns rather than solely to generic oscillatory coupling.

As the threshold increases ($PLV \geq 0.5$), the percentage of synchronized epochs progressively decreases for all rhythms, reflecting the intrinsic variability associated with internally generated cognitive processes. Nevertheless, rhythm R3 consistently maintains the highest synchronization levels. A possible explanation is that its longer silent interval enhances internal temporal prediction by increasing the salience of rhythmic segmentation, thereby facilitating more stable neural entrainment.

At the highest thresholds ($PLV \geq 0.7$ and $PLV \geq 0.9$), synchronized epochs become comparatively rare for all rhythms. Rather than indicating the absence of rhythm-related neural activity, this behaviour is consistent with previous observations showing that internally generated rhythmic representations exhibit lower phase consistency than externally driven sensory responses.

Overall, these findings support the hypothesis that rhythm imagination induces reproducible neural synchronization following the temporal organization of the imagined rhythmic

TABLE III: EEGNet classification performance (5-fold average). Accuracy is reported for the overall three-class problem, whereas Recall, Precision, and F1-score are reported separately for rhythm classes R1, R2, and R3, in the order [R1, R2, R3]

Subject	Acc.	Recall	Prec.	F1
S1	0.867	[1 .65 .85]	[1 .86 .75]	[1 .69 .78]
S2	0.757	[.72 .75 .80]	[.78 .56 1]	[.75 .64 .89]
S3	0.727	[1 .75 .50]	[.60 .75 1]	[.75 .75 .67]
S4	0.667	[0 1 .50]	[0 .63 1]	[0 .77 .67]

patterns, providing a physiological basis for the subsequent rhythm classification analysis.

B. Rhythm Classification

The second objective of the experimental evaluation is to assess whether the neural signatures generated during rhythm imagination are sufficiently discriminative to support automatic rhythm decoding.

Classification performance was evaluated using Accuracy, Precision, Recall and F1-score. In the context of the present three-class rhythm decoding problem, Accuracy measures the overall proportion of EEG epochs assigned to the correct imagined rhythm. Recall is computed separately for each rhythm and indicates the proportion of epochs belonging to a given rhythm that are correctly identified by the classifier. Precision quantifies, for each rhythm, the proportion of epochs predicted as that rhythm that actually belong to the corresponding class, while the F1-score summarizes the balance between Precision and Recall. These class-wise metrics are particularly relevant in the present setting, as they reveal whether the classifier performs consistently across the three imagined rhythms or whether its overall Accuracy is driven primarily by successful decoding of only a subset of the rhythmic patterns. Since the study focuses on subject-specific rhythm decoding, an independent EEGNet model was trained for each participant.

As mentioned in section IV-B dataset was randomly partitioned into training (80%), validation (15%), and test (5%) subsets, using 5-fold cross-validation procedure. Results obtained for four subjects are reported in Table III and are the average of the 5 folds. Accuracy is reported for the overall three-class classification problem, whereas Recall, Precision, and F1-score are reported separately for rhythm classes R1, R2, and R3.

Table III summarizes the subject-specific classification performance achieved by EEGNet. Classification accuracies range from 66.7% to 86.7%, consistently exceeding the chance level associated with three-class classification. Subject S1 achieves the highest overall accuracy, whereas Subjects S3 and S4 exhibit lower but still above-chance performance. Subject S4 exhibits a particularly rhythm-dependent classification profile. While the classifier correctly identifies all R2 epochs (Recall = 1.00), no R1 epoch is correctly recognized (Recall = 0 and F1 = 0), whereas R3 shows intermediate performance (Recall = 0.50 and F1 = 0.67). Therefore, the overall accuracy of 66.7% for S4 should not be interpreted as uniform decoding performance across the three imagined

rhythms, but rather as the result of markedly different classification performance across rhythm classes.

Precision and Recall remain high for several rhythm classes, while the corresponding F1-scores indicate that EEGNet is able to recover discriminative information despite the intrinsic variability of EEG recordings. The class-wise values also reveal that performance is not uniform across subjects and rhythms. For example, some participants show strong performance on specific rhythmic patterns but weaker performance on others, suggesting that rhythm imagery may produce subject-dependent neural signatures.

Taken together, the synchronization analysis and classification results provide preliminary indications that internally imagined rhythmic patterns may be associated with measurable neural signatures containing discriminative information. While the observed phase-locking patterns and above-chance subject-specific classification performance are consistent with this hypothesis, the limited number of participants and recordings does not allow us to establish the robustness or generalizability of these findings. Therefore, the present results should be interpreted as an initial proof of concept motivating further investigation of rhythm imagery as a potential interaction modality for future BCMI. Moreover, since the rhythmic conditions were acquired in separate recording sessions, session-specific effects cannot be fully disentangled from rhythm-related information; therefore, the classification results should be regarded as preliminary rather than as conclusive evidence of rhythm-specific neural decoding.

VII. CONCLUSION

This paper investigated the feasibility of decoding internally imagined rhythmic patterns from EEG recordings as a potential interaction modality for Brain-Computer Musical Interfaces (BCMI). A processing framework combining Independent Component Analysis, Phase Locking Value analysis, Empirical Mode Decomposition, and EEGNet classification was proposed to characterize rhythm-related neural activity and evaluate its discriminative potential.

Experimental results showed measurable neural synchronization during rhythm imagery, while subject-specific EEGNet models achieved classification accuracies above chance level. These preliminary findings suggest that EEG recordings acquired during rhythm imagery may contain discriminative information relevant to EEG-based musical interaction.

Although the present work represents a proof-of-concept study involving a limited number of participants, it establishes a foundation for future BCMI within the Internet of Sounds. Future work will focus on larger cohorts, randomized experimental protocols, cross-session and cross-subject validation, and multimodal interaction frameworks integrating EEG with complementary sensing modalities such as audio, gesture, and motion capture.

REFERENCES

- [1] L. Turchet, M. Lagrange, C. Rottondi, G. Fazekas, N. Peters, J. Østergaard, F. Font, T. Bäckström, and C. Fischione, "The internet of sounds:

- Convergent trends, insights, and future directions,” *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11 264–11 292, 2023.
- [2] E. R. Miranda and J. Castet, *Guide to brain-computer music interfacing*. Springer, 2014.
 - [3] I. Daly, D. Williams, A. Kirke, J. Weaver, A. Malik, F. Hwang, E. Miranda, and S. J. Nasuto, “Affective brain-computer music interfacing,” *Journal of neural engineering*, vol. 13, no. 4, p. 046022, 2016.
 - [4] S. Hoeffle, A. Engel, R. Basilio, V. Alluri, P. Toiviainen, M. Cagy, and J. Moll, “Identifying musical pieces from fmri data using encoding and decoding models,” *Scientific Reports*, vol. 8, no. 1, p. 2266, 2018.
 - [5] L. Bellier, A. Llorens, D. Marciano, A. Gunduz, G. Schalk, P. Brunner, and R. T. Knight, “Music can be reconstructed from human auditory cortex activity using nonlinear decoding models,” *PLoS biology*, vol. 21, no. 8, p. e3002176, 2023.
 - [6] J. Zhang, K. Li, B. Yang, and X. Han, “Local and global convolutional transformer-based motor imagery eeg classification,” *Frontiers in Neuroscience*, vol. 17, p. 1219988, 2023.
 - [7] M. A. Abbasi, H. F. Abbasi, M. Z. Aziz, W. Haider, Z. Fan, and X. Yu, “A novel precisely designed compact convolutional eeg classifier for motor imagery classification,” *Signal, Image and Video Processing*, vol. 18, no. 4, pp. 3243–3254, 2024.
 - [8] G. M. D. Liberto, G. Marion, and S. A. Shamma, “Accurate decoding of imagined and heard melodies,” *Frontiers in Neuroscience*, vol. 15, p. 673401, 2021.
 - [9] G. Marion, G. M. D. Liberto, and S. A. Shamma, “The music of silence: part i: responses to musical imagery encode melodic expectations and acoustics,” *Journal of Neuroscience*, vol. 41, no. 35, pp. 7435–7448, 2021.
 - [10] R. S. Schaefer, R. J. Vlek, and P. Desain, “Music perception and imagery in eeg: Alpha band effects of task and stimulus,” *International Journal of Psychophysiology*, vol. 82, no. 3, pp. 254–259, 2011.
 - [11] V. Alluri, P. Toiviainen, I. P. Jääskeläinen, E. Glerean, M. Sams, and E. Brattico, “Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm,” *Neuroimage*, vol. 59, no. 4, pp. 3677–3689, 2012.
 - [12] J. A. Grahn and M. Brett, “Rhythm and beat perception in motor areas of the brain,” *Journal of Cognitive Neuroscience*, vol. 19, no. 5, pp. 893–906, 2007.
 - [13] H. Merchant, J. Grahn, L. Trainor, M. Rohrmeier, and W. T. Fitch, “Finding the beat: a neural perspective across humans and non-human primates,” *Philosophical Transactions of the Royal Society B*, vol. 370, no. 1664, p. 20140093, 2015.
 - [14] H. Okawa, K. Suefusa, and T. Tanaka, “Neural entrainment to auditory imagery of rhythms,” *Frontiers in Human Neuroscience*, vol. 11, p. 493, 2017.
 - [15] T.-H. Cheng, Z. S. C. Creel, and J. R. Iversen, “How do you feel the rhythm: Dynamic motor-auditory interactions are involved in the imagination of hierarchical timing,” *Journal of Neuroscience*, vol. 42, no. 3, pp. 500–512, 2022. [Online]. Available: <https://www.jneurosci.org/content/42/3/500>
 - [16] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, “Eegnet: A compact convolutional neural network for eeg-based brain-computer interfaces,” *Journal of Neural Engineering*, vol. 15, no. 5, p. 056013, 2018.
 - [17] T. Fujioka, L. J. Trainor, E. W. Large, and B. Ross, “Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations,” pp. 1791–1802, 2012.
 - [18] I. Tal, E. W. Large, E. Rabinovitch, Y. Wei, C. E. Schroeder, D. Poeppel, and E. Zion Golumbic, “Neural entrainment to the beat: The “missing-pulse” phenomenon,” *The Journal of Neuroscience*, vol. 37, no. 26, pp. 6331–6341, 2017.
 - [19] A. Guillot and C. Collet, “Duration of mentally simulated movement: a review,” *Journal of Motor Behavior*, vol. 37, no. 1, pp. 10–20, 2005.
 - [20] S. Hébert and L. L. Cuddy, “Detection of metric structure in auditory figural patterns,” *Perception Psychophysics*, vol. 64, no. 6, pp. 909–918, 2002.
 - [21] H. H. Kornhuber and L. Deecke, “Brain potential changes in voluntary and passive movements in humans: readiness potential and reafferent potentials,” *Pflügers Archiv-European Journal of Physiology*, vol. 468, pp. 1115–1124, 2016.
 - [22] M. Jeannerod, “Mental imagery in the motor context,” *Neuropsychologia*, vol. 33, no. 11, pp. 1419–1432, 1995.
 - [23] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, “Iclabel: An automated electroencephalographic independent component classifier, dataset, and website,” *NeuroImage*, vol. 198, pp. 181–197, 2019.