

Decoding a Cognitive Performance State From Behavioral Data in the Presence of Auditory Stimuli

Saman Khazaei¹, *Graduate Student Member, IEEE*, and Rose T. Faghih², *Senior Member, IEEE*

Abstract—Cognitive performance state is an unobserved state that refers to the overall performance of cognitive functions. Deriving an informative observation vector as well as the adaptive model and decoder would be essential in decoding the hidden performance. We decode the performance from behavioral observation data using the Bayesian state-space approach. Forming an observation from the paired binary response with the associated continuous reaction time may lead to an overestimation of the performance, especially when an incorrect response is accompanied by a fast reaction time. We apply the marked point process (MPP) framework such that the performance decoder takes the correct/incorrect responses and the reaction time associated with correct responses as an observation. We compare the MPP-based performance with two other decoders in which the pairs of binary and continuous signals are taken as the observation; one decoder considers an autoregressive (AR) model for the performance state, and the other one employs an autoregressive-autoregressive conditional heteroskedasticity (AR-ARCH) model which incorporates the time-varying and adaptive innovation term within the model. To evaluate decoders, we use the simulated data and the n-back experimental data in the presence of multiple music sessions. The Bayesian state-space approach is a promising way to decode the performance state. With respect to individual perspective, the estimated MPP-based and ARCH-based performance states outperform the AR-based estimation. Based on the aggregated data analysis, the ARCH-based performance decoder outperforms the other decoders. Performance decoders can be employed in educational settings and smart workplaces to monitor one's performance and contribute to developing a feedback controller in closed-loop architecture to improve cognitive performance.

Index Terms—Bayesian inference, music, cognitive performance, state-space estimation.

I. INTRODUCTION

THE cognitive brain states are often unobserved internal states that cannot be directly measured, and they should be decoded from the encoded information. To do so, a brain state decoder can be employed to provide an estimate of the state given the available observation. The cognitive performance decoder is a type of brain state decoder that quantifies the overall performance of one's cognitive functions. Similar to the majority of cognitive variables, cognitive performance state is dynamic, and clear-cut scores are not able to reflect its dynamic [1], [2], [3]. There is a need for continuous quantification of cognitive performance. In this research, we are particularly interested in decoding the performance state during the n-back task in the presence of music. The n-back task mainly engages the working memory as a cognitive function of interest. Working memory is one of the basic cognitive functions with temporary storage that can process the stored information [4], [5].

Recent frameworks in designing brain state decoders include the Bayesian state-space methods as well as machine learning-based (ML-based) approaches [2], [4], [6], [7], [8], [9], [10], [11], [12]. Most of the ML-based approaches depend on input labels, training data, and classification of segments of data [6], [7], [8]. Therefore, ML-based approaches may encounter challenges in the absence of labels and training data [13]. The Bayesian state-space approach is a statistical framework developed on Bayesian principles, and it is able to estimate a continuous state trajectory given the available sequential observations and independent of training data. A typical Bayesian state-space approach considers a linear state model with time-invariant process noise (i.e., innovation term) variance, and hidden state can be decoded from the specified observation. The autoregressive (AR) model is an example of such linear and time-invariant models that is typically used in cognitive neuroscience to model the performance state [9].

In the context of Bayesian state filtering, a wide range of observations can be used to decode an unobserved internal brain state within the state-space models. These observations include but are not limited to invasive and non-invasive

Received 25 March 2024; revised 25 October 2024; accepted 2 November 2024. Date of publication 11 November 2024; date of current version 10 December 2024. This work was supported in part by NSF CAREER: MINDWATCH: Multimodal Intelligent Noninvasive brain state Decoder for Wearable Adaptive Closed-loop architectures under Grant 1942585/2226123 and in part by New York University (NYU) start-up funds. (Rose T. Faghih served as the senior author.) (Corresponding author: Rose T. Faghih.)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board (IRB) at the University of Houston, TX, USA.

The authors are with the Department of Biomedical Engineering, New York University, New York, NY 10010 USA (e-mail: rfaghih@nyu.edu).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TNSRE.2024.3495704>, provided by the authors.

Digital Object Identifier 10.1109/TNSRE.2024.3495704

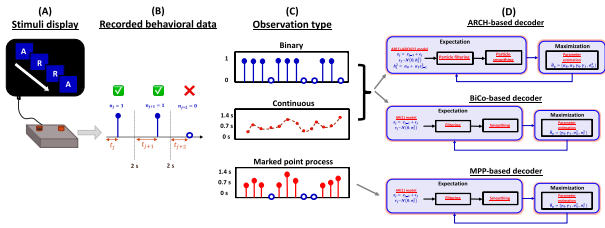


Fig. 1. An overview of decoding the cognitive performance state from behavioral data during the n-back task. (A) A series of stimuli is presented. (B) The behavioral data, including the sequence of correct/incorrect responses (binary) and reaction time (continuous), are collected. (C) The observation vector is formed based on various combinations of correct/incorrect responses and reaction time. (D) The performance state decoders take different observations and decode the performance state according to their specific designs: the ARCH-based decoder considers a time-varying process noise variance that follows the ARCH structure, and the decoder takes the pairs of correct/incorrect responses (binary) and reaction time (continuous); the BiCo-based decoder considers a time-invariant noise variance and takes the pairs of correct/incorrect responses (binary) and reaction time (continuous); the MPP-based decoder considers a time-invariant noise variance and takes the marked point process observation.

measurements such as neural recordings from the brain, peripheral physiological signals, and behavioral data [1], [3], [9], [10], [14], [15]. While the neural recordings and physiological signals are the informative metrics of one's brain activity, collecting them requires advanced setups. On the other hand, behavioral measurements are readily available and can be collected in everyday life settings, and we can assign different sets of behavioral observations to decode the performance state. The subject's binary sequence of correct and incorrect responses is one of the common behavioral observations that can be used in this paradigm [16], [17], [18]. Another informative behavioral signal that can be employed as an observation is the continuous reaction time (i.e., response time) [18], [19], [20]. In the field of cognitive neuroscience, various forms of observation vectors have been developed to decode the performance. Previously, the pairs of binary responses and continuous reaction time observation (i.e., BiCo observation) associated with the presented cognitive task have been used. Particularly, in [9], the performance state was decoded using the BiCo observation. One of the possible concerns that can be raised when using BiCo type observation is that the decoder can overestimate the performance when a pair of extremely fast reaction time and an incorrect response occurs. To address this concern, the marked point process can be applied to form the behavioral observation. The MPP framework is widely used in neuroscience to relate the ensemble neural spiking activity (point process) to any relevant covariates (mark) [21], [22], [23], [24]. For instance, in [15], the hidden arousal state is decoded from the sequence of arousal-related neural impulses and their corresponding amplitudes. Here, we consider an AR(1) performance state and assign a new observation vector consisting of correct/incorrect responses and the reaction times associated with the sequence of correct responses (successful trials), solely. Then, we decode the performance from the sequence of responses as well as the reaction time associated with successful trials.

The cognitive performance state can be drastically affected by internal and external variables such as environmental changes and emotional status [2], [4], [25]. Thus, an adaptive and time-varying performance state in the presence of non-linearity should be considered to resemble real-world settings. Although the AR model presents a pretty decent estimation outcome, it mostly assumes a time-invariant process noise variance within a linear model. The autoregressive Conditional heteroskedasticity (ARCH) framework has been widely applied for analyzing volatility in financial time series [26]. The term volatility in financial returns refers to the degree and rate of variation over a period of time [27], [28]. Particularly, the ARCH framework expresses the conditional variance of a time series as a function of past residuals, which enables the model to capture the time-varying nature of the volatility. The ARCH framework can be implemented within the state-space model to ensure the time-variability of the model. The Bayesian state-space methods are often employed within an expectation-maximization (EM) algorithm to estimate the hidden state and recover the unknown model parameter [15], [29], [30], [31], [32]. Including the ARCH framework within the state-space model induces the non-linearity, which prevents us from having a closed-form solution at the expectation step (E-step) [33]. A particle filter can be seen as a particular case of a Bayesian filter that is developed to cope with non-linear and non-Gaussian systems [34]. To handle the non-linear model structure and approximate an expected value of a function, an appropriate particle filter (PF) can be used [4]. Following the originally developed framework in [4], we model the performance state using the AR-ARCH state-space representation, and we employ the particle filtering to decode the performance state from the binary and continuous observation in the presence of non-linear and time-varying process noise variance.

Here, following the approach presented in [2] and [9], we first decode the performance state from one binary and one continuous (i.e., BiCo) observation using the AR state model. Thereafter, we develop a new performance state model to decode the performance from the MPP-type observation. Then, to induce the time-varying process noise variance, we follow the proposed method in [4], and we model the performance state using the AR-ARCH state-space representation. Then, we decode the ARCH-based performance using the sequence of responses and reaction time [4]. Hence, two types of decoders use a pair of binary and continuous observations, and one of them relies on the MPP observation. Fig. 1 depicts an overview of the presented cognitive performance decoders in this research. We demonstrate our findings on simulated and experimental datasets, followed by a discussion. Finally, we summarize our findings and draw a conclusion accordingly.

II. METHODS

A. Data

We investigate two datasets collected at the University of Houston [25], [35]. The experiments were designed to study the feasibility of employing safe intervention (i.e., music).

Particularly, using music, high and low arousing environments were simulated to modulate the cognitive brain states. The informed consent from all participants/subjects was obtained, and the behavioral and physiological signals were recorded during a working memory task called the n-back task while different types of background music were used. To preserve the personalized nature of the intervention, the applied music was provided by the individuals. The n-back task involves working memory usage, and subjects are presented with a series of stimuli such that they need to recall the n^{th} previous one and determine whether the current stimulus is the same as the n^{th} previous one or not. We use the recorded behavioral signals, which include the sequence of correct and incorrect responses as well as the response time of individuals. The individuals who performed the n-back task in dataset 1 are noted as “participants”, while individuals who performed the n-back task in dataset 2 are stated as “subjects”. The list of participants IDs and subject IDs are available in the supplementary information.

1) **Dataset 1:** The experimental procedures in the experiment were approved by the institutional review board (IRB) at the University of Houston, TX, USA (STUDY00002013) [35]. The total number of 6 participants (participants 1 to 6) performed an equal number of 1-back and 3-back task blocks within two music sessions. The calming background music was used throughout the first session, while the exciting background music was played during the second session. A total number of 16 task blocks were presented at each session, and the 1-back and 3-back blocks were equally and randomly distributed. Each task block was initiated with 5 seconds of instruction followed by 22 trials in which a letter was displayed for 0.5 seconds, and a plus sign was provided for 1.5 seconds (2 seconds trial window). Hence, the participants had two seconds to respond. At the end of each block, the participant had a 10 seconds resting time. Also, a 20 seconds resting segment was contrived in the middle of each session (at the end of the 8th task block). A detailed description of the experiment is available in [35].

2) **Dataset 2:** The original study includes 2 sets of n-back experiments in the presence of safe interventions such as music, coffee, and perfume [25]. The complete dataset is publicly available through the PhysioNet database [36]. The experimental procedures in the original study were approved by the IRB at the University of Houston, TX, USA (STUDY 00002490). In this study, we only consider the first set of the experiment that incorporates the music intervention. A total number of 17 subjects participated in the experiment while seven subjects were excluded from the original study due to measurement errors and artifacts [25]. Hence, 10 subjects (subjects A1 to A10) are studied. An equivalent number of 1-back and 3-back task blocks were randomly distributed within 4 sessions: The first session was presented with no background music; the second session included the calming music; the exciting music was played within the third session, and a newly generated relaxing music was provided in the last session (fourth session). The generated relaxing music was produced based on the subject’s preference and using

deep learning neural networks. Each session included 16 task blocks, each of which had 5 seconds of instruction followed by 22 trials. Each trial consisted of 0.5 seconds of stimulus presentation, along with a 1.5 seconds plus sign. Thus, subjects could deliver their response in 2 seconds trial window. At the end of each task block, 10 seconds resting section was provided, and 20 seconds resting section in the middle of each session was implemented. A subject performed a total number of 1408 trials (i.e., 4 sessions $\times$ 16 task blocks $\times$ 22 trials). A comprehensive description of the experiment is provided in [25].

B. State-Space Framework

The state-space framework is a common approach in control theory to present the unobserved state of interest [15]. In this research, the state of interest is the hidden cognitive performance. Once the state-space model for cognitive performance is specified, we wish to estimate the unobserved performance state. To do so, the available observation vector is employed, and the hidden state can be decoded through different methods. In this research, we decode the performance state by considering three different frameworks:

- AR(1) state model (time-invariant process noise variance) given one binary and one continuous observation (BiCo-based decoder) [9].
- AR(1) state model (time-invariant process noise variance) based on MPP observation (MPP-based decoder).
- AR(1)-ARCH(1) state model (time-variant process noise variance) given one binary and one continuous observation (ARCH-based decoder) [4].

Fig. 1 presents an overview of the presented cognitive performance decoders in this research.

C. Performance State Estimation in Presence of a Time-Invariant Process Noise Variance – One Binary and One Continuous Observation

First, the unobserved performance state is modeled based on the presence of time-invariant process noise variance [2], [4], [9]. Hence, the following AR is employed to present the hidden performance state (z_j):

$$z_j = z_{j-1} + \epsilon_j, \quad (1)$$

where $\epsilon_j \sim \mathcal{N}(0, \sigma_\epsilon^2)$.

Once the performance state model is specified, the observation vector may be identified to estimate the hidden performance state. Similar to [4], [9], and [18], the sequence of responses (n_j) and reaction times (t_j) are considered as the available observations.

The binary responses is assumed to follow the Bernoulli distribution. The $p_j = P(n_j = 1)$ can be related to performance state employing the logit transformation [4], [9], [18]:

$$p_j = \frac{1}{1 + e^{-(\mu + z_j)}}, \quad (2)$$

where the constant μ may be found by considering $z_j \approx 0$. Hence, $\mu = \log(\frac{p_0}{1-p_0})$ where p_0 stands for the baseline

probability [9]. The baseline probability can be derived from the average probability of an correct response occurrence during the entire experiment.

Similar to [9], the log of reaction time can be linearly related to the state as

$$r_j = \log(t_j) = \gamma_0 + \gamma_1 z_j + v_j, \quad (3)$$

where $v_j \sim \mathcal{N}(0, \sigma_v^2)$ stands for the observation noise, and γ_0 , γ_1 , and σ_v^2 are the unknown parameters to be determined.

Thus, based on the observation vector $Y^J = \{(n_1, r_1), (n_2, r_2), \dots, (n_J, r_J)\}$, the performance state (z_j) and unknown model parameters $\theta_p = \{\sigma_\epsilon^2, \gamma_0, \gamma_1, \sigma_v^2\}$ can be decoded by employing the EM algorithm. More information on the decoder equations can be found in the supplementary information and in [4] and [9].

D. Performance State Estimation in Presence of a Time-Invariant Process Noise Variance – A Marked Point Process Observation

1) *State-Space Model in Presence of a Time-Invariant Process Noise Variance – A Marked Point Process Observation:* Here, while we follow the mentioned AR model to present the performance state, we use different observation vector to estimate the hidden performance state (z_j) . Hence, the hidden performance state (z_j) is presented using the following AR model:

$$z_j = z_{j-1} + \epsilon_j \quad (4)$$

Inspired by [15], the observation vector can take a marked point process (MPP) form such that the log of reaction time presents the marked for the observed point process where the correct response is presented (i.e., $n_j = 1$). Let us indicate the correct response indices by $\tilde{J} = \{j | n_j = 1\}$. Hence, the marked observation can be represented as [15]:

$$u_{j \in \tilde{J}} = \gamma_0 + \gamma_1 z_j + v_j, \quad (5)$$

where the continuous observation value ($u_j = \log(t_j)$) is considered only when the correct response exists (i.e., $j \in \tilde{J}$). Please recall that we assume the Bernoulli distribution for the binary responses ($p_j = P(n_j = 1)$). The joint density function for the observed correct response is

$$p(n_j \cap u_j | z_j) = \begin{cases} 1 - p_j & \text{if } n_j = 0 \\ p_j \frac{1}{\sqrt{2\pi\sigma_v^2}} e^{-\frac{(u_j - \gamma_0 - \gamma_1 z_j)^2}{2\sigma_v^2}} & \text{if } n_j = 1 \end{cases}. \quad (6)$$

Similar to [15], we can apply an expectation-maximization approach to estimate the hidden state z_j and unknown parameters $\theta_p = \{\sigma_\epsilon^2, \gamma_0, \gamma_1, \sigma_v^2\}$, simultaneously.

2) *State-Space Decoder in Presence of a Time-Invariant Process Noise Variance – A Marked Point Process Observation:* The MPP-based decoders have been used to decode the arousal from the MPP-type observation [15]. Similarly, we apply an EM framework to determine the unknown model parameters $\theta_p = \{\sigma_\epsilon^2, \gamma_0, \gamma_1, \sigma_v^2\}$, and estimate the continuous performance state z_j at the same time.

E-Step:

The E-step equations are presented as:

Predict:

$$z_{j|j-1} = z_{j-1|j-1}, \quad (7)$$

$$\sigma_{j|j-1}^2 = \sigma_{j-1|j-1}^2 + \sigma_\epsilon^2, \quad (8)$$

Update:

if $n_j = 0$

$$z_{j|j} = z_{j|j-1} + \sigma_{j|j-1}^2 (n_j - p_{j|j}), \quad (9)$$

$$\sigma_{j|j}^2 = \left[\frac{1}{\sigma_{j|j-1}^2} + p_{j|j}(1 - p_{j|j}) \right]^{-1}, \quad (10)$$

if $n_j = 1$

$$C_j = \frac{\sigma_{j|j-1}^2}{\gamma_1^2 \sigma_{j|j-1}^2 + \sigma_v^2}, \quad (11)$$

$$z_{j|j} = z_{j|j-1} + C_j \left[\sigma_v^2 (n_j - p_{j|j}) + \gamma_1 (u_j - \gamma_0 - \gamma_1 z_{j|j-1}) \right], \quad (12)$$

$$\sigma_{j|j}^2 = \left[\frac{1}{\sigma_{j|j-1}^2} + p_{j|j}(1 - p_{j|j}) + \frac{\gamma_1^2}{\sigma_v^2} \right]^{-1}. \quad (13)$$

Employing numerical solvers such as Newton-Raphson, we may solve for $z_{j|j}$.

By reversing the direction, the smoother can be implemented, and the sets of smoothed states and variance can be found:

$$B_j = \frac{\sigma_{j|j}^2}{\sigma_{j+1|j}^2}, \quad (14)$$

$$z_{j|J} = z_{j|j} + B_j (z_{j+1|J} - z_{j+1|j}), \quad (15)$$

$$\sigma_{j|J}^2 = \sigma_{j|j}^2 + B_j^2 (\sigma_{j+1|J}^2 - \sigma_{j+1|j}^2). \quad (16)$$

M-Step:

The expected values of z_j^2 , and $z_j z_{j-1}$ can be written as

$$\mathbb{E}[z_j^2] = z_{j|J}^2 + \sigma_{j|J}^2, \quad (17)$$

$$\mathbb{E}[z_{j+1} z_j] = z_{j+1|J} z_{j|J} + B_j \sigma_{j+1|J}^2. \quad (18)$$

The expected value of log-likelihood function Q can take the following form:

$$\begin{aligned} Q = & \sum_{j=1}^J \mathbb{E}[n_j (\mu + z_j) - \log(1 + e^{\mu + z_j})] \\ & + \frac{-\|\tilde{J}\|}{2} \log(2\pi\sigma_v^2) - \sum_{j \in \tilde{J}} \frac{\mathbb{E}[(u_j - \gamma_0 - \gamma_1 z_j)^2]}{2\sigma_v^2} \\ & + \frac{-J}{2} \log(2\pi\sigma_\epsilon^2) - \sum_{j=1}^J \frac{\mathbb{E}[(z_j - z_{j-1})^2]}{2\sigma_\epsilon^2}, \end{aligned} \quad (19)$$

where $\|\tilde{J}\|$ denotes the length of $\tilde{J}$, and the unknown parameters are estimated such that they maximize Q . The iteration will persist between the E-step and the M-step until meeting the convergence criteria.

E. Performance State Estimation in Presence of a Time-Variant Process Noise Variance – One Binary and One Continuous Observation

1) *State-Space Model in Presence of a Time-Varying Process Noise Variance*: To induce the non-linearity, we implement the ARCH noise [4]. Hence, the performance state model follows the AR-ARCH structure in which the process noise variance is time-variant [4]. The AR(1)-ARCH(1) model for the hidden performance state (z_j) can be shown as

$$z_j = z_{j-1} + \epsilon_j, \quad (20)$$

where $\epsilon_j \sim \mathcal{N}(0, h_j^2)$ is a process noise.

In the ARCH(1) model the noise term can be presented as

$$\epsilon_j = h_j w_j, \quad (21)$$

where $h_j^2 = \alpha_0 + \alpha_1 \epsilon_{j-1}^2$, and it stands for the time-varying conditional variance. The terms α_0 and α_1 are the ARCH model parameters in which $\alpha_0 > 0$ and $0 \leq \alpha_1 < 1$ are the constraints [26].

Considering the binary response (n_j) and the log of reaction time ($r_j = \log(t_j)$) as the available observations, we wish to estimate the unobserved performance in the presence of a time-varying process noise variance. Following [2], [4], and [9], the performance state can be related to the probability of having correct response using $\log\left(\frac{p_j}{1-p_j}\right) = \mu + z_j \Rightarrow$

$p_j = \frac{1}{1+e^{-(\mu+z_j)}}$; the reaction time is expressed in terms of performance state via $r_j = \log t_j = \gamma_0 + \gamma_1 z_j + v_j$.

2) *State-Space Decoder in Presence of a Time-Varying Process Noise Variance*: In this paradigm, the observed vector is $Y^J = \{(n_1, r_1), (n_2, r_2), \dots, (n_J, r_J)\}$, and we aim to estimate the performance state z_j and the unknown model parameters $\theta_p = \{\gamma_0, \gamma_1, \sigma_v^2, \alpha_0, \alpha_1\}$. We apply the EM framework to estimate the hidden state and recover the model parameters at the same time. However, in this case, deriving the expected value of the log-likelihood function would be a concern to be addressed. The log-likelihood function takes the non-linear form such that finding a closed form solution becomes challenging [33]. Therefore, an effective approximation is required. Inspired by [4] and [33], we use a particle filter for the AR-ARCH model to decode the hidden performance.

E-Step:

The developed particle filter in [4] enables us to estimate the performance state z_j in the presence of a time-varying process noise variance. Following the proposed approach in [4], the algorithm implementation is described below.

- Step 1: Generate K number of particles with an arbitrary initial mean state value $\bar{z}_0$ and initial state variance σ_0^2 .
- Step 2: Proceed from $j-1$ to j and draw particles $\hat{z}_j(k)$ based on the conditional state mean ($\bar{z}_j$) and conditional variance (σ_j^2) derived from the Bayes' rule:

$$\begin{aligned} \bar{z}_j(k) = & \frac{\sigma_{j-1}^2(k) + h_j^2(k)}{\gamma_1^2(\sigma_{j-1}^2(k) + h_j^2(k)) + \sigma_v^2} \\ & \times \left[\sigma_v^2(n_j - \bar{p}_j(k)) \right. \\ & \left. + \gamma_1(r_j - \gamma_0 - \gamma_1 \hat{z}_{j-1}(k)) \right] + \hat{z}_{j-1}(k), \quad (22) \end{aligned}$$

$$\begin{aligned} \sigma_j^2(k) = & \left[\frac{1}{\sigma_{j-1}^2(k) + h_j^2(k)} \right. \\ & \left. + \bar{p}_j(k)(1 - \bar{p}_j(k)) + \frac{\gamma_1^2}{\sigma_v^2} \right]^{-1}, \quad (23) \end{aligned}$$

where $h_j^2(k) = \alpha_0 + \alpha_1(\hat{z}_{j-1}(k) - \hat{z}_{j-2}(k))^2$.

By inserting $\bar{p}_j(k) = [1 + e^{-(\mu + \bar{z}_j(k))}]^{-1}$ in (22), the term $\bar{z}_j(k)$ would appear on both sides of (22), and $\bar{z}_j(k)$ can be derived using numerical approaches. Given the $\bar{z}_j(k)$ and $\sigma_j^2(k)$, the particles can be generated from $\hat{z}_j(k) \sim \mathcal{N}(\bar{z}_j(k), \sigma_j^2(k))$.

- Step 3: Assign the importance weight $w_j(k)$ to each sample $\hat{z}_j(k)$. The weight density function can be specified based on the problem of interest. Similar to [4] and [33], the optimal importance weight w_j can be derived from

$$w_j^{(k)} = \frac{\mathcal{N}(Y^j; \hat{z}_j(k), \sigma_v^2) \mathcal{N}(\hat{z}_j(k); \hat{z}_{j-1}(k), h_j^2(k))}{\mathcal{N}(\hat{z}_j(k); \bar{z}_j(k), \sigma_j^2(k))}. \quad (24)$$

- Step 4: Normalize the weights $w_j^{(k)} = \frac{w_j^{(k)}}{\sum_{j=1}^K w_j^{(k)}}$ [4], [37], and perform the resampling to avoid the particle degeneracy. As described in [38], the effective sample size can be approximated: $N_{ess} = \frac{1}{\sum_{k=1:K} w_j^2(k)}$. Considering $K/2$ as the threshold, if $N_{ess} < K/2$, the systematic resampling can be performed [39], [40].
- Step 5: At $j = J$, we reverse the direction and to improve the estimation by deriving a set of smoothed state $\tilde{z}_j(k)$ and equally smoothed weights $\tilde{w}_j(k) = 1/K$ [4]. We consider $p(\tilde{z}_j(k-1)|\hat{z}_j(k))$ as the distribution of interest, and smooth the state solely based on the dynamics of the system to avoid the potential overfitting [41].

M-Step:

During the M-step, model parameters can be determined such that they maximize the expected value of our log-likelihood function. We approximate the expected value of the log-likelihood by using the particles and their weights [4], [33], [42]. Hence, $\mathbb{E}[Q(z_j, \theta)] \approx \frac{1}{K} \sum_{k=1}^K Q(\tilde{z}_j(k), \theta)$, where the log-likelihood function (Q) is presented as

$$\begin{aligned} Q = & \sum_{j=1}^J \left[n_j(\beta + z_j) - \log(1 + e^{\beta + z_j}) \right] \\ & + \frac{-J}{2} \log(2\pi\sigma_v^2) - \sum_{j=1}^J \frac{(r_j - \gamma_0 - \gamma_1 z_j)^2}{2\sigma_v^2} \\ & + \frac{-J}{2} \log(2\pi) - \frac{1}{2} \sum_{j=1}^J \left[\frac{(z_j - z_{j-1})^2}{\alpha_0 + \alpha_1(z_{j-1} - z_{j-2})^2} \right. \\ & \left. + \log(\alpha_0 + \alpha_1(z_{j-1} - z_{j-2})^2) \right]. \quad (25) \end{aligned}$$

We recover a set of parameters $\theta_p = \{\gamma_0, \gamma_1, \sigma_v^2, \alpha_0, \alpha_1\}$ that maximizes the $\mathbb{E}[Q]$ [4]. A detailed description of the parameter estimation in M-step is provided in the supplementary information. Following the proposed overfitting control technique in [13], we update the unknown parameters progressively and consider the early-stopping to bypass the potential overfitting. Let us indicate the model parameters associated with the reaction time by $\theta_r = \{\gamma_0, \gamma_1, \sigma_v^2\}$; we update those parameters at the m^{th} iteration as:

$$\theta_{r, \text{updt}}^{(m+1)} = \theta_r^{(m+1)} + \lambda \left[\theta_{r, \text{pred}}^{(m+1)} - \theta_r^{(m)} \right], \quad (26)$$

where $0 < \lambda \leq 1$ is a hyper-parameter to be set based on the dataset, and $\theta_{r, \text{pred}}^{(m+1)}$ refers to the initially recovered parameters (available in the supplementary information). Also, we set the early-stopping criteria such that if $\sigma_{v, \text{updt}}^{2, (m+1)} \leq \sigma_{v, \text{threshold}}^2$, the iteration stops; the $\sigma_{v, \text{threshold}}^2$ is a hyper-parameter that can be fine-tuned accordingly [13]. The algorithm iterates between E-step and M-step until the convergence.

Once the hidden performance state (z_j) is decoded, we can derive a generalized and person-specific index of performance called high performance index (HPI). Particularly, given the differences in human brain structure and variation in cognitive abilities from one individual to another [43], an individual's baseline can be considered, and HPI can be derived from $p(z_j > \text{threshold})$, where the threshold is indicative of the cognitive performance baseline. The threshold can be set to the median of state [15]. This type of index is between zero and one, and it is inspired by the introduced term called observer certainty level in [1]. It can reflect the probability that an event (e.g., correct response) occurs more than just by chance in a behavioral learning experiment [15]. The HPI can be used to reduce the inter-subject variability and evaluate the performance estimates among the participant pool. Also, the probability of observing a correct response (p_j) can be derived by plugging the estimated performance state in $p_j = \frac{1}{1 + e^{-(\mu + z_j)}}$. Another performance metric that can be found from the decoder's output would be the reconstructed input observation (e.g., binary correct responses or continuous reaction time) by plugging the estimated state and parameters in the observation model equations such that the reconstructed signal can be compared with the observed signal to evaluate the goodness of fit.

F. Data Simulation

To further evaluate the decoders, a set of data can be simulated based on the previously developed framework in [4]. Particularly, we simulate the hidden state using the AR process with time-varying process noise variance, a continuous observation signal using a linear function in (3), and the probability of observing a binary event using the logit transformation [15], and the sequence of events can be simulated where an event will occur at index j if $p_j \geq$ baseline probability.

III. RESULTS

In the context of cognitive brain states, the absence of continuous ground truth leads to challenges in the

TABLE I
THE COEFFICIENT OF DETERMINATION (R^2) VALUES FOR THE ESTIMATED STATE AND PROBABILITY IN THE SIMULATED STUDY

Variable	R^2		
	ARCH-based	MPP-based	BiCo-based
z_j	0.8599	0.8486	0.7316
p_j	0.8464	0.8257	0.8698

decoder's evaluation process [13]. One may address the issue by implicitly evaluating the available cues. As an instance of such implicit evaluations, different types of n-back tasks and music interventions exist to be considered. Also, according to the observations equations in (3), the reconstructed observation can be derived by inserting the estimated state and parameters in $\gamma_0 + \gamma_1 z_j$, and the goodness of fit with respect to the actual observation may help us in evaluating the decoder's outcome. Furthermore, a simulation study can provide valuable information regarding the decoder's performance [13]. We evaluate the estimated performance using simulated data as well as two different experimental datasets.

A. Simulated Data

To evaluate the ARCH-based, MPP-based, and BiCo-based performance decoders, we use a set of simulated data. Mainly, we simulate the reaction time, a sequence of binary responses, and the performance state. To simulate the data, we set the parameters according to the findings from experimental data in [4]. The sub-figures of Fig. 2 present the state estimation results for ARCH-based, MPP-based, and BiCo-based performance decoders. The left sub-figure is related to the ARCH-based performance decoder, the middle sub-figure demonstrates the MPP-based decoder's outcome, and the right sub-figure depicts the BiCo-based decoder result. Sub-panels, from top to bottom, present (1) the applied observation and reconstructed one, (2) the simulated sequence of incorrect response, (3) the simulated state and estimated one, (4) the simulated probability and estimated one, and (5) the HPI, respectively. The reconstructed continuous observation in the first sub-panel is shown using the black signal, and the actual one is presented in red. The simulated signals (ground truth) are shown in red in the third and fourth sub-panels, and the estimated variables are presented in blue. Table I presents the coefficient of determination (R^2) values for the estimated state (z_j) and probability (p_j) in the simulated study. The higher R^2 value stands for a better alignment.

B. Experimental Data

1) *Individual Data*: Two unique datasets are presented to evaluate each filter based on the experimental data. Considering ARCH-based decoder, the performance state estimation results for one participant/subject from each dataset are illustrated in Fig. 3, where sub-figure (A) depicts the results associated with dataset 1 [4], and sub-figure (B) is linked to dataset 2. Similarly, the MPP-based and BiCo-based decoders' findings on both datasets are presented in sub-figures (A) and (B) of Fig. 4 and Fig. 5, respectively. In

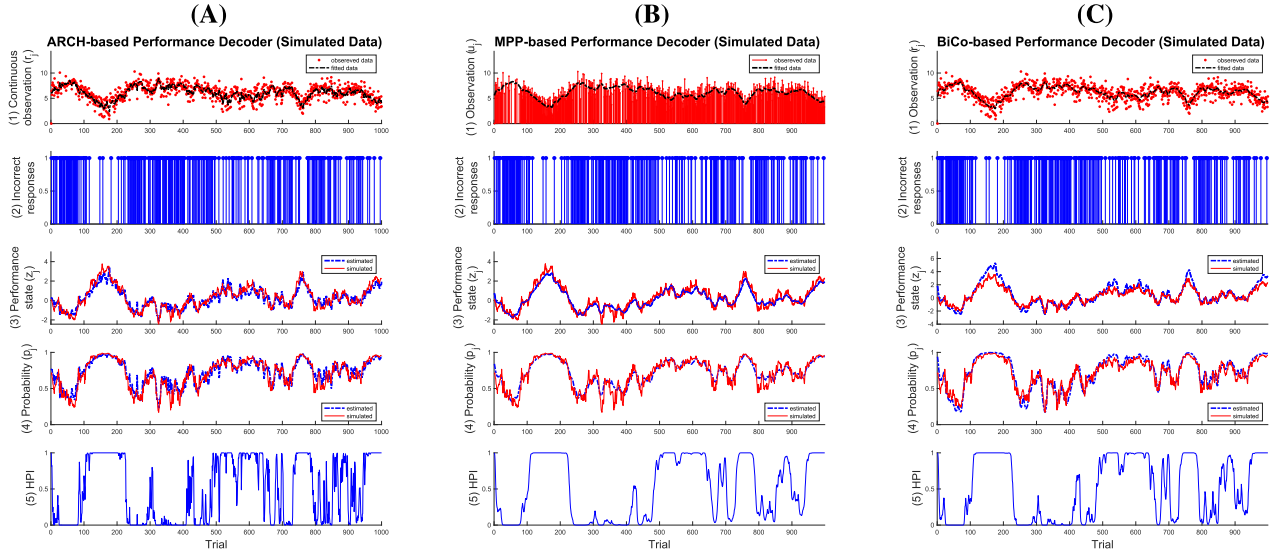


Fig. 2. State estimation with simulated data using ARCH-based, MPP-based, and BiCo-based decoders. (A) The left sub-figure presents the ARCH-based decoder findings on the simulated data. (B) The middle sub-figure shows the MPP-based decoder findings on the simulated data. (C) The right sub-figure shows the BiCo-based decoder findings on the simulated data. Within each sub-figure, the sub-panels respectively depict: (1) the simulated observation based on reaction time (red) and the reconstructed one (black); (2) the sequence of simulated incorrect response (blue vertical lines); (3) the simulated (red) and estimated (blue) state; (4) the probability of observing a correct response p_j and its estimate; (5) the HPI.

TABLE II

THE COEFFICIENT OF DETERMINATION (R^2) VALUES FOR THE FITTED OBSERVATIONS WITH RESPECT TO EXPERIMENTAL DATA

R^2 - Dataset 1			
Participant ID	BiCo-based $\hat{r}_j$	MPP-based $\hat{u}_j$	ARCH-based $\hat{r}_j$
1	0.3720	0.4419	0.8314
2	0.1279	0.0312	0.7665
3	0.0913	0.0082	0.8532
4	0.1523	0.1198	0.7740
5	0.2700	0.2171	0.7743
6	0.1695	0.0420	0.8076
R^2 - Dataset 2			
Subject ID	BiCo-based $\hat{r}_j$	MPP-based $\hat{u}_j$	ARCH-based $\hat{r}_j$
A1	0.0320	0.0321	0.8447
A2	0.2116	0.2205	0.7898
A3	0.4859	0.5430	0.7847
A4	0.2050	0.2886	0.9141
A5	0.5071	0.6170	0.7833
A6	0.4166	0.6429	0.8889
A7	0.0261	0.0504	0.8073
A8	0.4199	0.5264	0.8512
A9	0.4414	0.5080	0.7986
A10	0.3116	0.2894	0.7541

each sub-figure of Fig. 3 to Fig. 5, the sub-panels present: (1) the applied observation (r_j or u_j) and its reconstructed signal (black curve); (2) the incorrect responses during the experiment; (3) the decoded performance state during the experiment; (4) the estimated probability of observing a correct response; (5) the calculated HPI. The corresponding green and red background colors in sub-figure (A) of Fig. 3 to Fig. 5 refer to calming and exciting sessions, respectively; the corresponding cyan, green, red, and yellow background colors in sub-figure (B) of Fig. 3 to Fig. 5, in turn, indicate sessions with no music, relaxing music, exciting music, and the relaxing generated music. For Fig. 3 to Fig. 5, the milder background

colors stand for the 1-back task blocks, and the more intense background colors indicate the 3-back task blocks. The outcome of decoders for other participants/subjects is in the supplementary information.

In order to evaluate the status of the estimated states, the coefficient of determination (R^2) value for the reconstructed continuous observation, and the receiver operating characteristic curve (ROC curve) for the classified correct responses are considered: In Table II, the R^2 values for each decoder are shown. Note that to calculate the R^2 for the reconstructed observation in the MPP-based decoder ($u_{j \in \tilde{J}}$), we only consider data points of $\tilde{J} = \{j | n_j = 1\}$. Additionally, we aim to classify the correct responses from the estimated probability p_j in each decoder. To do so, we classify successful trials from $p_j > p_{\text{threshold}}$, where the range of threshold values is set between 0.01 and 0.99. This wide range of thresholds enables us to contemplate people with various cognitive baselines. Fig. 6 displays the ROC curves associated with the classified correct responses for one participant/subject of each dataset. The area under the curve (AUC) values for participant 1 in dataset 1 are 0.87 for the ARCH-based decoder (red curve), 0.63 for the MPP-based decoder (black curve), and 0.70 for the BiCo-based decoder (blue curve) estimates. The AUC values for subject A10 estimates are 0.68 for the ARCH-based decoder (red curve), 0.70 for the MPP-based decoder (black curve), and 0.64 for the BiCo-based decoder (blue curve). The high AUC value implies a reliable classification performance. The ROC curves for all the studied participants/subjects can be found in the supplementary information.

One of the aims of incorporating the ARCH structure process noise is to capture the surrounding impacts (e.g., auditory stimuli) on the performance state. Hence, we represent a session-wise perspective of the findings on one

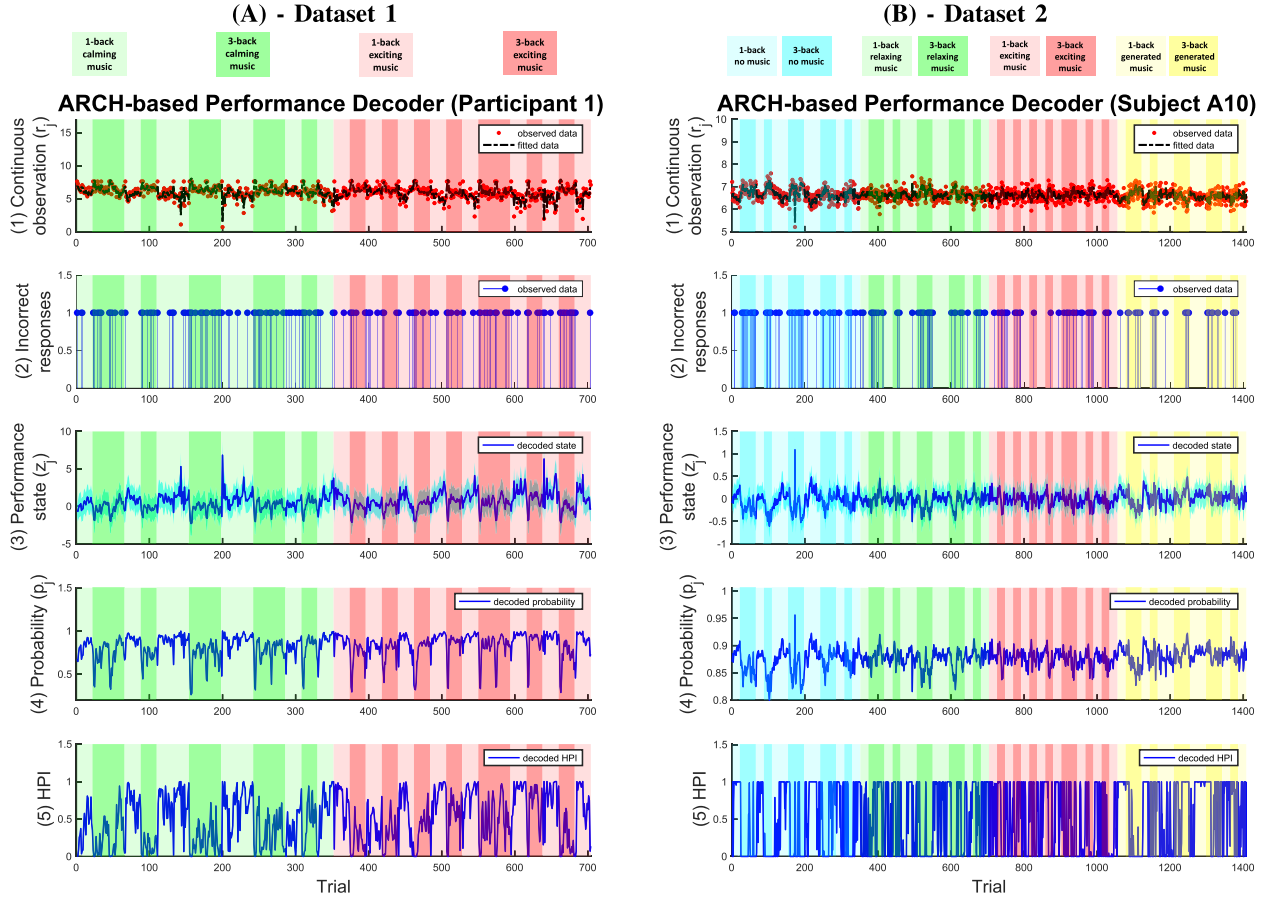


Fig. 3. The ARCH-based performance state estimation for the example individuals during the n-back experiments in datasets 1 and 2. The sub-figures from left to right present: (A) the ARCH-based performance estimates of dataset 1 [4]; (B) the ARCH-based performance estimates of dataset 2. Within each sub-figure, the sub-panels respectively present: (1) The continuous value of the applied observation (red) and its fit (black); (2) the sequence of incorrect responses (blue vertical line); (3) estimated performance z_j (blue) and its 95% confidence limits (blue); (4) the estimated probability p_j (blue); (5) the HPI (blue). In regard to background colors in sub-figure (A): the green background color indicates a calming music session, and the red background color represents an exciting music session. In regard to background colors in sub-figure (B): the cyan background color corresponds to the no music session; the green background color stands for the relaxing music session; the red background color indicates the exciting music session; and the yellow background color refers to the generated relaxing music session. The milder background colors stand for the 1-back task blocks, and the more intense background colors indicate the 3-back task blocks.

subject from dataset 2. Fig. 7 present the ARCH-based decoder variables in a session-wise manner. The sub-panels of the figure depict the performance state (z_j), process noise variance (h_j^2), process noise term (ϵ_j), l^2 -norm of process noise variance, and frequency spectrum of process noise variance, respectively. The bottom sub-panel of Fig. 7 displays the frequency spectrum of process noise variance during the no music (black curve), relaxing music (blue curve), exciting music (red curve), and generated relaxing music (yellow curve) sessions. The cyan, green, red, and yellow background colors in Fig. 7, in turn, refer to the no music, relaxing music, exciting music, and relaxing generated music sessions. In Fig. 7, the milder background colors indicate the 1-back task blocks, and the more intense background colors present the 3-back task blocks. Such session-wise perspective of variables for all the studied participants/subjects from both datasets are presented in the supplementary information.

2) Aggregated Data: To evaluate the performance of decoders in the context of aggregated data, we should be cautious and address the potential inter-subject variability. To do so, since the HPI is calculated based on the subject's

baseline, it may be considered as a normalized index of performance that can be used in an aggregated perspective, in which we analyze a vector of data that includes performance metrics from multiple subjects. More specifically, we evaluate the correlation between the mean HPI and the average of observed performance metrics (i.e., reaction time and correct response) among the participants/subjects of dataset 1 and 2 (Fig. 8 and Fig. 9).

Findings on dataset 1 present that the Pearson correlation coefficient of average correct response and HPI associated with the ARCH-based, MPP-based, and BiCo-based decoders are $r_n = 0.7060$, $r_n = 0.2620$, and $r_n = 0.4747$, respectively. Considering the correlation between the reaction time and decoded HPI via ARCH-based, MPP-based, and BiCo-based decoders, the correlation coefficients are $r_{rt} = -0.7821$, $r_{rt} = -0.2716$, and $r_{rt} = -0.4837$.

A similar analysis is performed on dataset 2, and the Pearson correlation coefficients between the average correct response and HPI derived from ARCH-based, MPP-based, and BiCo-based decoders are $r_n = 0.3842$, $r_n = 0.3308$, and $r_n = 0.2740$, respectively. Considering the correlation between the reaction time and decoded HPI via ARCH-based, MPP-

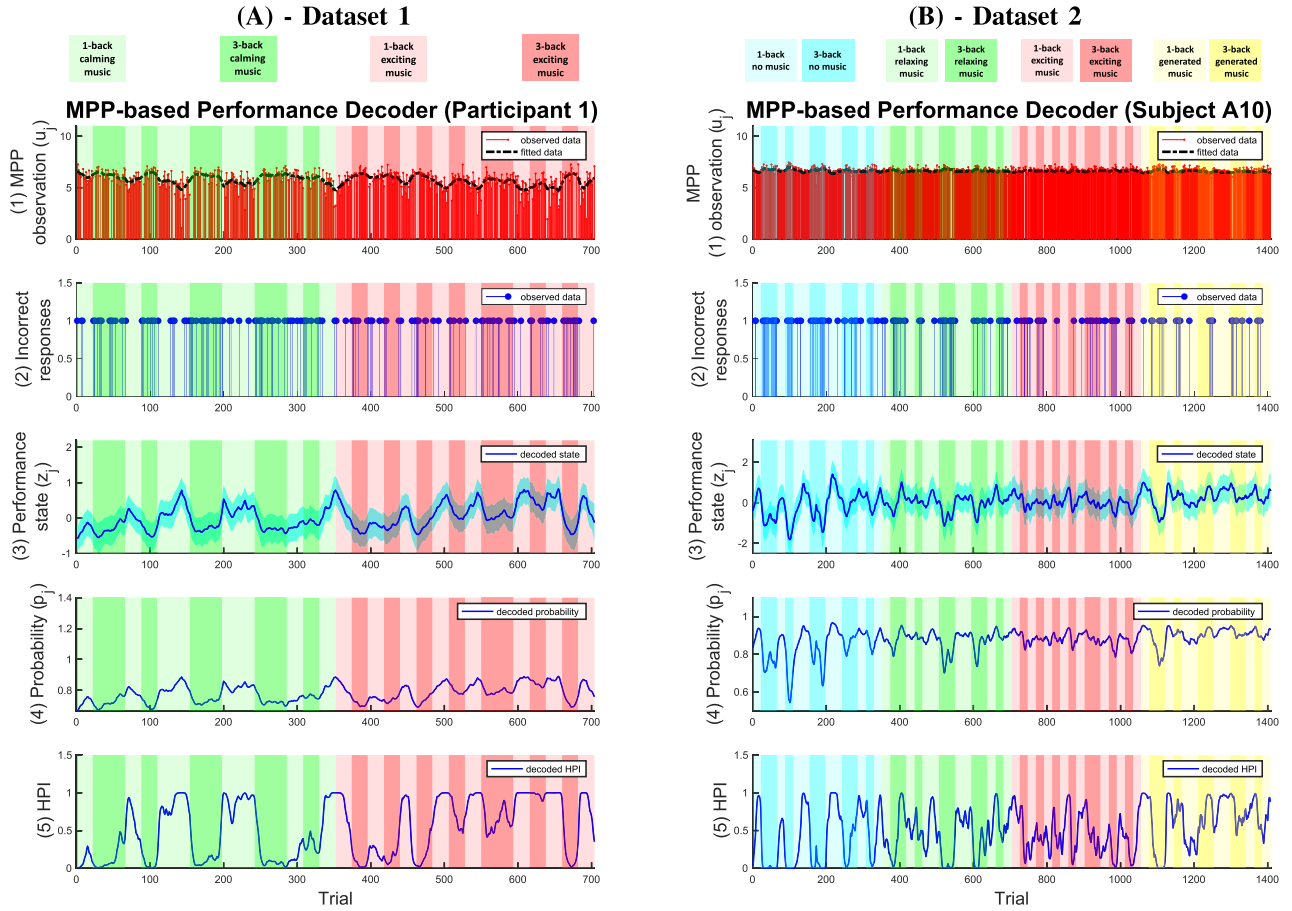


Fig. 4. The MPP-based performance state estimation for the example individuals during the n-back experiments in datasets 1 and 2. The sub-figures from left to right present: **(A)** the MPP-based performance estimates of dataset 1; **(B)** the MPP-based performance estimates of dataset 2. Within each sub-figure, the sub-panels respectively present: (1) The MPP value of the applied observation (red) and its fit (black); (2) the sequence of incorrect responses (blue vertical line); (3) estimated performance z_j (blue) and its 95% confidence limits (blue); (4) the estimated probability p_j (blue); (5) the HPI (blue). In regard to background colors in sub-figure **(A)**: the green background color indicates a calming music session, and the red background color represents an exciting music session. In regard to background colors in sub-figure **(B)**: the cyan background color corresponds to the no music session; the green background color stands for the relaxing music session; the red background color indicates the exciting music session; and the yellow background color refers to the generated relaxing music session. The milder background colors stand for the 1-back task blocks, and the more intense background colors indicate the 3-back task blocks.

based, and BiCo-based decoders, the correlation coefficients are $r_{rt} = -0.8257$, $r_{rt} = -0.6617$, and $r_{rt} = -0.6479$.

IV. DISCUSSION

The main aim of this research is to decode cognitive performance from behavioral data. The absence of ground truth in this paradigm can turn the decoder's evaluation process to a challenging step. One way to cope with this challenge is to interpret the findings based on the available observations and cues. We evaluate the ARCH-based, MPP-based, and BiCo-based decoders in the context of experimental setups as well as simulation study. The studied experiments include the n-back task as a cognitive task of interest, and incorporate different personalized music interventions.

The results of the simulated study for ARCH-based and MPP-based decoders present a decent agreement between the simulated variables and the estimated ones (Fig. 2). Specifically, the R^2 measurements for both ARCH-based and MPP-based decoded performance states are higher than 0.80. While we simulated a dataset based on the previous findings

in [4], simulating a performance state with a different dynamic is possible using the unique baseline probability (p_0) [15].

Considering Fig. 3, the first noticeable property is the noisy estimates of the ARCH-based decoder, while the other decoders in Fig. 4 and Fig. 5 present the smooth estimates. We implement a time-varying process noise variance to preserve the encoded environmental information within the noise. However, there would be a trade-off between having a noisy measure of a state and the smooth one. While the ARCH-based model can store the information within the process noise, this non-linear model can introduce additional noise. The other possible reason is that the reaction time dynamic, as one of the applied observations, is fast time-varying, and the particle filtering approach is highly prone to overfit to the dynamic of reaction time. Also, we can observe the range of decoded performance varies between each decoder, which arises from the decoder's nature. Hence, to compare the estimated performance derived from each decoder, we can rely on the HPI, which is the generalized form of performance estimates.

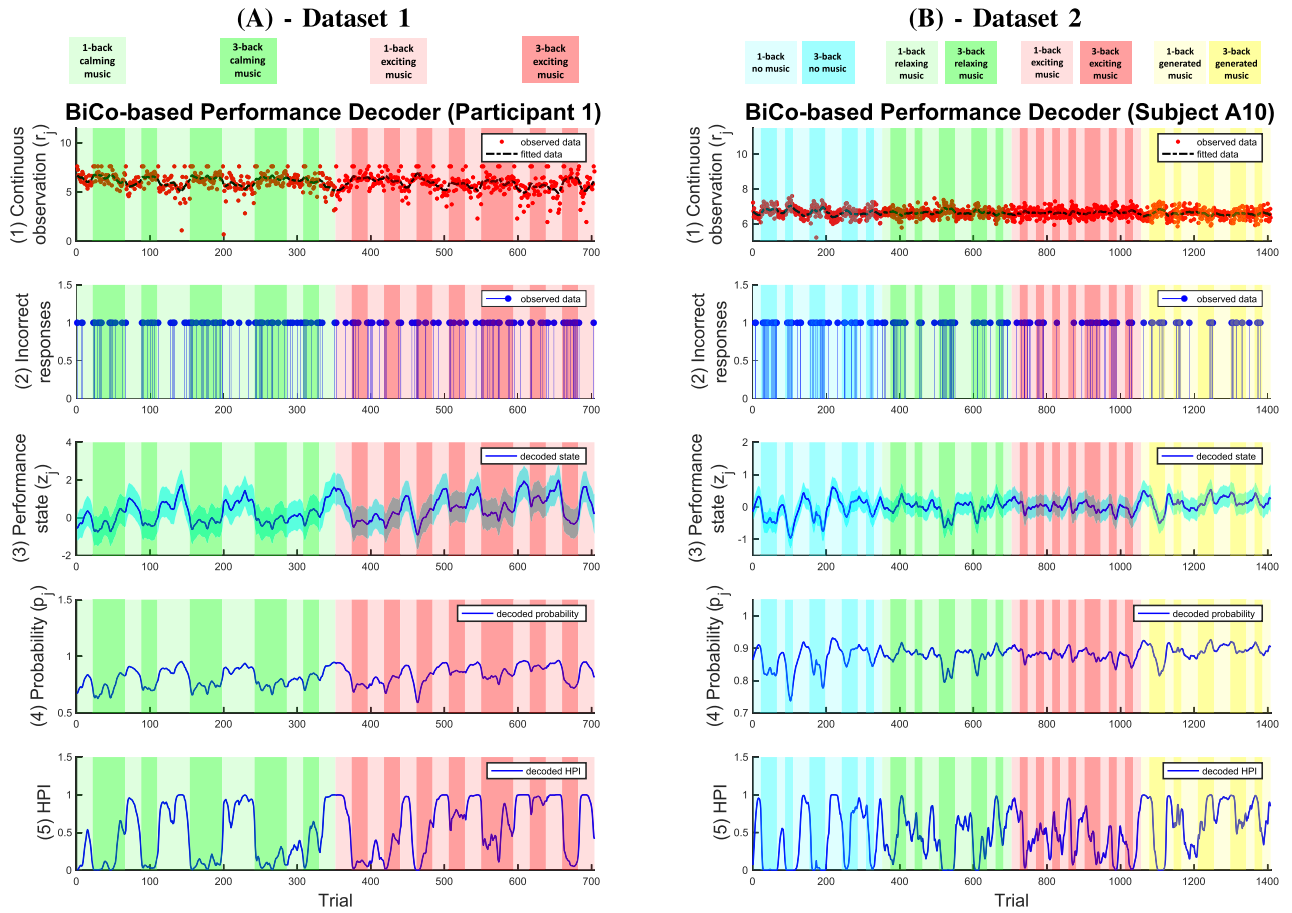


Fig. 5. The BiCo-based performance state estimation for the example individuals during the n-back experiments in datasets 1 and 2. The sub-figures from left to right present: (A) the BiCo-based performance estimates of dataset 1 [9]; (B) the BiCo-based performance estimates of dataset 2 [25]. Within each sub-figure, the sub-panels respectively present: (1) The continuous value of the applied observation (red) and its fit (black); (2) the sequence of incorrect responses (blue vertical line); (3) estimated performance z_j (blue) and its 95% confidence limits (shaded blue); (4) the estimated probability p_j (blue); (5) the HPI (blue). In regard to background colors in sub-figure (A): the green background color indicates a calming music session, and the red background color represents an exciting music session. In regard to background colors in sub-figure (B): the cyan background color corresponds to the no music session; the green background color stands for the relaxing music session; the red background color indicates the exciting music session; and the yellow background color refers to the generated relaxing music session. The milder background colors stand for the 1-back task blocks, and the more intense background colors indicate the 3-back task blocks.

The variance of the decoded performance using the ARCH-based decoder is high, while this is not the case in the other evaluated decoders. This implies that the ARCH coefficient α_1 maintains a high value, which increases the recovered time-varying process noise variance and leads to a better fit of reconstructed observation (the first subplot of Fig. 3). However, in the other decoders, the performance state model does not count for the time-varying property of the process noise variance. Table II presents the R^2 for the fitted $r_j = \gamma_0 + \gamma_1 z_j$ based on the estimated states where the R^2 values associated with the ARCH-based decoder are drastically higher than the other decoders. While such goodness of fit arises from the higher degree of freedom that the ARCH model induces, there should be a specific protocol to avoid overfitting [4]. To do so, we follow a proposed approach in [13], and we update the recovered model parameters progressively. Additionally, in the backward smoothing step, we set the weights based on the state transition and avoid relying on the observation data. Using the overfitting control technique, we find that while the R^2 value for the fitted

observation is lower than the reported one in [4], we still maintain a high goodness of fit (Table II).

According to the ROC curves in Fig. 6, the ARCH-based decoder outperforms in classifying the correct responses associated with this particular participant from dataset 1. However, the right sub-panel of Fig. 6 presents that the MPP-based decoder posits a higher AUC value for a studied subject from dataset 2. This can inspire us to develop an AR-ARCH model and decoder that only takes the MPP-type observation.

Despite the subtle difference between the estimated performance in an MPP-based decoder and a BiCo-based one, the difference can play a crucial role, and BiCo-based may encounter overestimation of performance. As the apparent instance of such cases in sub-figure (A) of Fig. 5, in block numbers 24 (trials 528-550) and 30 (trials 572-594), the baselines of continuous observation (r_j) are low while the population of incorrect responses is relatively dense; this can lead to overestimation of the performance state by overfitting to the continuous observation and dedicating less importance to the incorrectness of responses in BiCo-

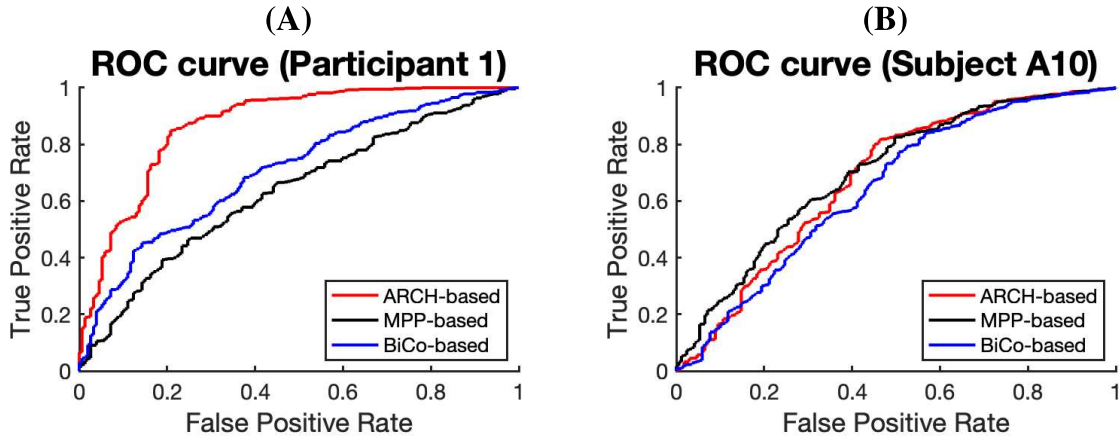


Fig. 6. The receiver operating characteristic (ROC) curve for the classified correct response by estimated probabilities within each decoder. The sub-figures from left to right present: (A) The ROC curve related to the classified correct response for a participant in dataset 1. (B) The ROC curve related to the classified correct response for a subject in dataset 2.

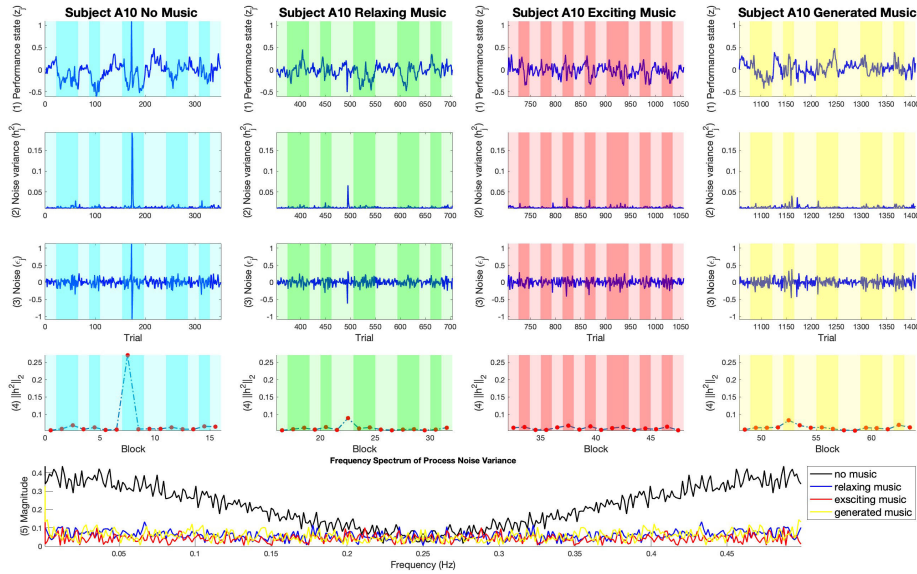


Fig. 7. Session-wise representation of ARCH-based decoder properties for one subject in dataset 2. The sub-panels from top to bottom present: (1) The estimated performance state (z_j) at each session; (2) the recovered process noise variance (h_j^2) at each session; (3) the recovered process noise (ϵ_j) at each session; (4) the l^2 -norm of process noise variance at each session, and (5) the frequency spectrum of process noise variance for all sessions. The cyan, green, red, and yellow background colors represent no music, relaxing music, exciting music, and generated relaxing music sessions, respectively. The black, blue, red, and yellow curves at the bottom sub-panel are associated with no music, relaxing music, exciting music, and generated relaxing music sessions, respectively. The milder background colors stand for the 1-back task blocks, and the more intense background colors indicate the 3-back task blocks.

based filter. Implementing the MPP-based observation can overcome this issue by ignoring the low baselines of r_j when incorrect responses occur and imposing more weights on the correct/incorrect responses. A similar outcome can be seen in Fig. 4 and Fig. 5, where at the beginning of the 53th block (trials 1166-1188), consecutive incorrect responses as well as fast reaction time are presented, and the MPP-based performance state has the local minimum at the start of the block (sub-figure (B) of Fig. 4), while the BiCo-based decoder presents a relatively low variation of state in the entire block (sub-figure (B) of Fig. 5).

Fig. 7 is displayed to evaluate the capability of the ARCH model to capture the environmental impacts. In these session-wise views, we aim to distinguish the applied auditory stimuli. Particularly, considering the pick value of $\|h_j^2\|_2$ and the

frequency spectrum of process noise variance, there is a clear separation between the no music session (black curve in the bottom subplot) and other sessions. By visual inspection, we can see that the noise and noise variance at each session have a particular structure that can be applied toward the change point detection. However, the main scope of this study is dedicated to the decoder design paradigm, and the offered ARCH-based model can be investigated deeply in this context.

Fig. 8 and Fig. 9 are presented to evaluate the overall performance of each decoder, given the aggregated data of all participants/subjects in each dataset. Specifically, we use the average HPI and evaluate the association of the HPI with the available performance metrics such as average correct response and reaction time. The high positive correlation between the average HPI and the average correct response signal can be

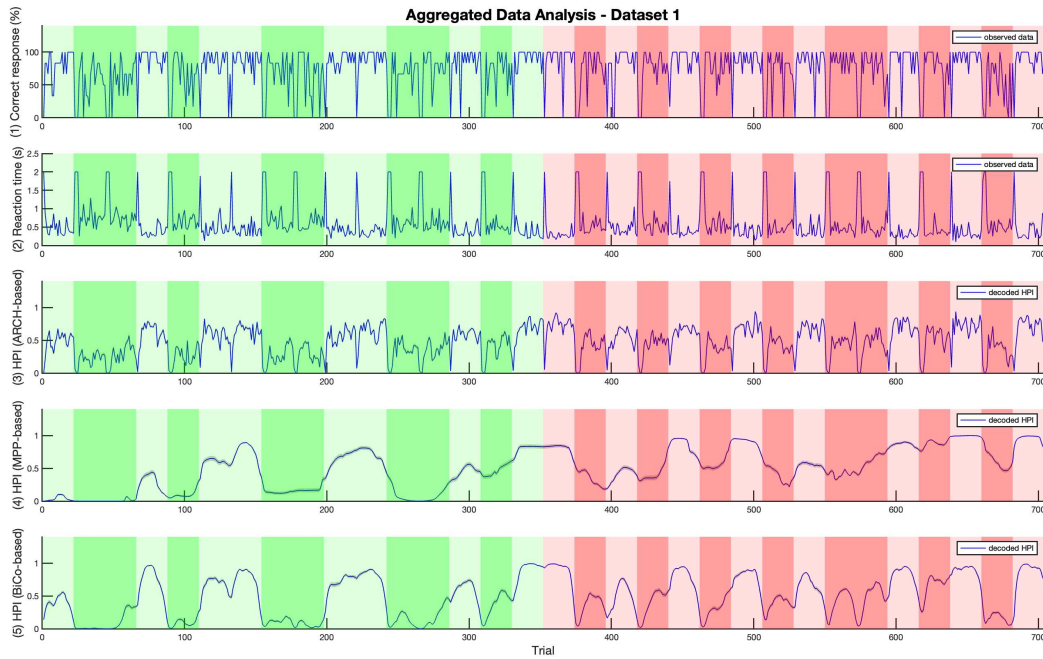


Fig. 8. Average performance indices within all the participants of experiment 1. The sub-panels, from top to bottom, presents (1) the percentage of average correct response among the participants, (2) the average of reaction times among the participants, (3) the average of decoded HPI from ARCH-based decoder among the participants, (4) the average of decoded HPI from MPP-based decoder among the participants, and (5) the average of HPI from BiCo-based decoder among the participants. The green background color indicates a calming music session, and the red background color represents an exciting music session. The milder background colors stand for the 1-back task blocks, and the more intense background colors indicate the 3-back task blocks.

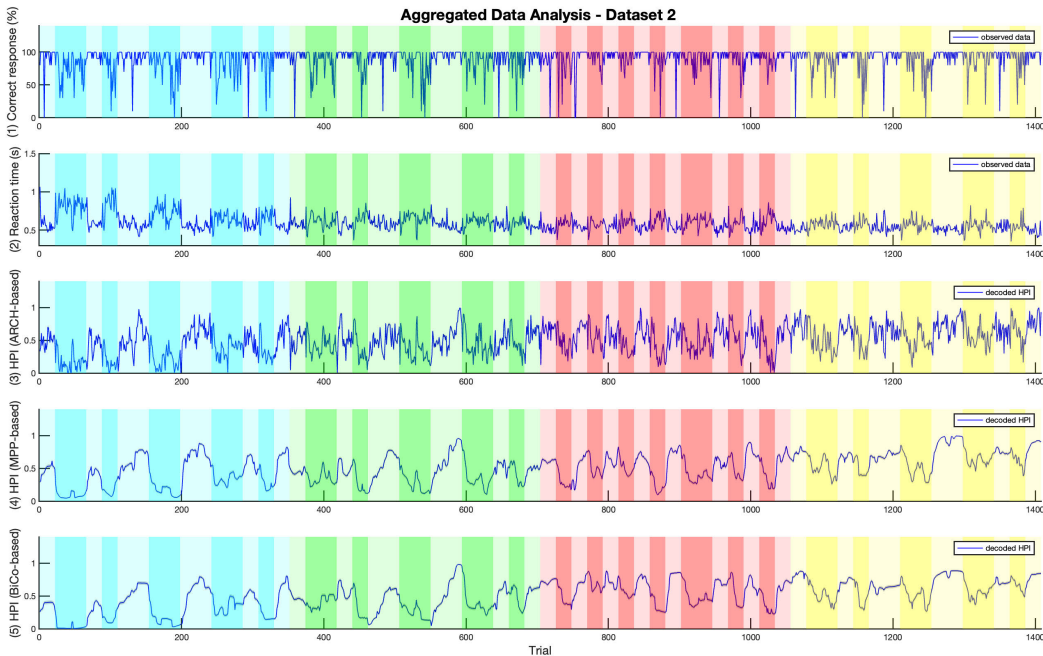


Fig. 9. Average performance indices within all the subjects of experiment 2. The sub-panels, from top to bottom, presents (1) the percentage of average correct response among the subjects, (2) the average of reaction times among the subjects, (3) the average of decoded HPI from ARCH-based decoder among the subjects, (4) the average of decoded HPI from MPP-based decoder among the subjects, and (5) the average of HPI from BiCo-based decoder among the subjects. The cyan background color corresponds to the no music session; the green background color stands for the relaxing music session; the red background color indicates the exciting music session; and the yellow background color refers to the generated relaxing music session. The milder background colors stand for the 1-back task blocks, and the more intense background colors indicate the 3-back task blocks.

interpreted as an indicator of reliable estimation, and the high negative association between the average HPI and the average reaction time may be deduced as an index of a reasonable fit. The derived Pearson correlation coefficients present a better

fit of derived ARCH-based HPI and applied performance metrics (i.e., average correct response and reaction time) in both datasets. Considering the aggregated data lens, one may interpret that the ARCH-based decoder outperforms the

MPP-based and BiCo-based decoder. However, it should be noted that an individual perspective should be considered in parallel. Specifically, one of the objectives of these hidden brain state decoders would be their implementation within the personalized automated closed-loop brain-machine interfaces that could potentially regulate the brain state according to the individual's attributes [25], [44].

V. CONCLUSION

The brain state decoder is an essential element of the closed-loop systems in understanding the impact of stimuli on the brain dynamic. Particularly, the cognitive performance state is an unobserved variable that can be affected by several internal and external factors. To closely resemble the dynamic of cognitive performance and decode the hidden state, a performance state model and decoder should be employed to quantify the performance from the available sets of observations.

In this research, we focus on the cognitive performance of individuals who were performing the n-back tasks in the presence of different types of music, and we study multiple types of decoders. We first study a presented performance decoder in [9], and present the findings accordingly. Inspired by the proposed framework in [15] for the hidden arousal state, we consider the MPP observation to decode the performance state. Particularly, the MPP-based decoder takes account of the reaction time when a correct response occurs. Thus, it only considers a reaction time associated with the correct response regardless of how fast an individual reacts within the incorrect trials. Then, to account for the environmental impacts on performance, we employ the developed ARCH-based decoder in [4]. Inducing the time-varying and non-linear process noise variance enables the model to be adaptable and encode the information within the process noise. However, the ARCH model introduces a higher degree of freedom, which may increase the likelihood of overfitting. If the tendency of overfitting is observed, the early-stopping approach can be one of the possible solutions to be considered.

According to the presented results on two studied datasets, we may conclude that the ARCH-based and MPP-based decoders outperform the BiCo-based one. It should be noted that considering the individual perspective and comparing the ARCH-based and MPP-based decoders together, we are not able to make such a solid judgment. However, the aggregated data view depicts a better performance of the ARCH-based decoder. Also, it should be noted that the MPP-based decoder shows a better performance compared to the BiCo-based decoder, and this can pave the way for a new avenue of decoder design procedures in which the ARCH noise structure can be considered in the modeling step, and the MPP-type observation can be assigned to the decoder (i.e., ARCH-MPP decoder). Hence, developing an ARCH-MPP decoder can be listed as a future direction of this research.

Our findings on both simulated and experimental data present the feasibility of performance decoder implementation within the closed-loop neural architecture. This architecture enables us to monitor and regulate the unobserved brain

states. Monitoring cognitive performance in a non-invasive manner can assist us in characterizing the interaction between affective brain states, identifying the factors that maximize the performance, and designing a safe neurofeedback mechanism to reach optimal performance [2], [25].

REFERENCES

- [1] A. C. Smith et al., "Dynamic analysis of learning in behavioral experiments," *J. Neurosci.*, vol. 24, no. 2, pp. 447–461, Jan. 2004.
- [2] S. Khazaei, M. R. Amin, and R. T. Faghih, "Decoding a neurofeedback-modulated cognitive arousal state to investigate performance regulation by the Yerkes-Dodson law," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 6551–6557.
- [3] X. Deng et al., "Estimating a dynamic state to relate neural spiking activity to behavioral signals during cognitive tasks," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 7808–7813.
- [4] S. Khazaei, M. R. Amin, and R. T. Faghih, "Decoding a neurofeedback-modulated performance state in presence of a time-varying process noise variance," in *Proc. 56th Asilomar Conf. Signals, Syst., Comput.*, Oct. 2022, pp. 990–996.
- [5] A. D. Baddeley and G. Hitch, "Working memory," in *Psychology of Learning and Motivation*, vol. 8. Amsterdam, The Netherlands: Elsevier, 1974, pp. 47–89.
- [6] M. Schütze et al., "Use of machine learning to predict cognitive performance based on brain metabolism in neurofibromatosis type 1," *PLoS ONE*, vol. 13, no. 9, Sep. 2018, Art. no. e0203520.
- [7] M. Stikic, R. R. Johnson, D. J. Levendowski, D. P. Popovic, R. E. Olmstead, and C. Berka, "EEG-derived estimators of present and future cognitive performance," *Frontiers Hum. Neurosci.*, vol. 5, p. 70, Aug. 2011.
- [8] M. R. Amin, D. S. Wickramasuriya, and R. T. Faghih, "A wearable exam stress dataset for predicting grades using physiological signals," in *Proc. IEEE Healthcare Innov. Point Care Technol. (HI-POCT)*, Mar. 2022, pp. 30–36.
- [9] M. R. Amin, M. Tahir, and R. T. Faghih, "A state-space investigation of impact of music on cognitive performance during a working memory experiment," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 757–762.
- [10] O. G. Sani, H. Abbaspourzad, Y. T. Wong, B. Pesaran, and M. M. Shanechi, "Modeling behaviorally relevant neural dynamics enabled by preferential subspace identification," *Nature Neurosci.*, vol. 24, no. 1, pp. 140–149, Jan. 2021.
- [11] A. Yousefi et al., "Predicting learning dynamics in multiple-choice decision-making tasks using a variational Bayes technique," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 3194–3197.
- [12] A. Yousefi et al., "Cognitive state prediction using an EM algorithm applied to gamma distributed data," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 7819–7824.
- [13] D. S. Wickramasuriya, S. Khazaei, R. Kiani, and R. T. Faghih, "A Bayesian filtering approach for tracking sympathetic arousal and cortisol-related energy from marked point process and continuous-valued observations," *IEEE Access*, vol. 11, pp. 137204–137247, 2023.
- [14] D. Hulsey, K. Zumwalt, L. Mazzucato, D. A. McCormick, and S. Jaramillo, "Decision-making dynamics are predicted by arousal and unstructured movements," *Cell Reports*, vol. 43, no. 2, 2024.
- [15] D. S. Wickramasuriya and R. T. Faghih, "A marked point process filtering approach for tracking sympathetic arousal from skin conductance," *IEEE Access*, vol. 8, pp. 68499–68513, 2020.
- [16] B. K. Hawes, T. T. Brunyé, C. R. Mahoney, J. M. Sullivan, and C. D. Aall, "Effects of four workplace lighting technologies on perception, cognition and affective state," *Int. J. Ind. Ergonom.*, vol. 42, no. 1, pp. 122–128, Jan. 2012.
- [17] C. Kirschbaum, O. T. Wolf, M. May, W. Wippich, and D. H. Hellhammer, "Stress- and treatment-induced elevations of cortisol levels associated with impaired declarative memory in healthy adults," *Life Sci.*, vol. 58, no. 17, pp. 1475–1483, Mar. 1996.
- [18] M. J. Prerau et al., "Characterizing learning by simultaneous analysis of continuous and binary measures of performance," *J. Neurophysiol.*, vol. 102, no. 5, pp. 3060–3072, Nov. 2009.

- [19] H. Peak and E. G. Boring, "The factor of speed in intelligence," *J. Experim. Psychol.*, vol. 9, no. 2, p. 71, 1926.
- [20] I. J. Deary, G. Der, and G. Ford, "Reaction times and intelligence differences: A population-based cohort study," *Intelligence*, vol. 29, no. 5, pp. 389–399, 2001.
- [21] D. Ba, S. Temereanca, and E. N. Brown, "Algorithms for the analysis of ensemble neural spiking activity using simultaneous-event multivariate point-process models," *Frontiers Comput. Neurosci.*, vol. 8, p. 6, Feb. 2014.
- [22] L. Tao, K. E. Weber, K. Arai, and U. T. Eden, "A common goodness-of-fit framework for neural population models using marked point process time-rescaling," *J. Comput. Neurosci.*, vol. 45, no. 2, pp. 147–162, Oct. 2018.
- [23] X. Deng, D. F. Liu, K. Kay, L. M. Frank, and U. T. Eden, "Clusterless decoding of position from multiunit activity using a marked point process filter," *Neural Comput.*, vol. 27, no. 7, pp. 1438–1460, Jul. 2015.
- [24] M. C. Burkhart, *A Discriminative Approach to Bayesian Filtering With Applications to Human Neural Decoding*. Providence, RI, USA: ProQuest Dissertations Publishing, 2019.
- [25] S. Parshi, "A functional-near infrared spectroscopy investigation of mental workload," Master's thesis, Univ. Houston, 2020.
- [26] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, Jul. 1982.
- [27] L. Bauwens, C. M. Hafner, and S. Laurent, *Handbook of Volatility Models and Their Applications*, vol. 3. Hoboken, NJ, USA: Wiley, 2012.
- [28] D. Ardia, *Financial Risk Management With Bayesian Estimation of GARCH Models Theory and Applications*. Cham, Switzerland: Springer, 2008.
- [29] A. C. Smith, M. R. Stefani, B. Moghaddam, and E. N. Brown, "Analysis and design of behavioral experiments to characterize population learning," *J. Neurophysiol.*, vol. 93, no. 3, pp. 1776–1792, Mar. 2005.
- [30] T. P. Coleman, M. Yanike, W. A. Suzuki, and E. N. Brown, "A mixed-filter algorithm for dynamically tracking learning from multiple behavioral and neurophysiological measures," in *The Dynamic Brain: An Exploration of Neuronal Variability and its Functional Significance*, M. Ding and D. L. Glanzman, Eds. New York, NY, USA: Oxford Univ. Press, 2011, pp. 1–16.
- [31] M. M. Shanechi, R. C. Hu, M. Powers, G. W. Wornell, E. N. Brown, and Z. M. Williams, "Neural population partitioning and a concurrent brain-machine interface for sequential motor function," *Nature Neurosci.*, vol. 15, no. 12, pp. 1715–1722, Dec. 2012.
- [32] M. M. Shanechi, R. C. Hu, and Z. M. Williams, "A cortical–spinal prosthesis for targeted limb movement in paralysed primate avatars," *Nature Commun.*, vol. 5, no. 1, pp. 1–9, Feb. 2014.
- [33] J.-C. Duan and A. Fulop, "A stable estimator of the information matrix under EM for dependent data," *Statist. Comput.*, vol. 21, no. 1, pp. 83–91, Jan. 2011.
- [34] V. Pesce, A. Colagrossi, and S. Silvestrini, *Modern Spacecraft Guidance, Navigation, and Control: From System Modeling to AI and Innovative Applications*. Amsterdam, The Netherlands: Elsevier, 2022.
- [35] S. Parshi, R. Amin, H. F. Azgomi, and R. T. Faghih, "Mental workload classification via hierarchical latent dictionary learning: A functional near infrared spectroscopy study," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, May 2019, pp. 1–4.
- [36] H. F. Azgomi et al., "Regulation of brain cognitive states through auditory, gustatory, and olfactory stimulation with wearable monitoring, (version 1.0.0)," *PhysioNet*, 2023, doi: [10.13026/jr6f-wc85](https://doi.org/10.13026/jr6f-wc85).
- [37] A. R. Poyatos-Bakker et al., "Load disaggregation through particle filtering of harmonic features," in *Proc. 20th Int. Conf. Harmonics Quality Power (ICHQP)*, 2022, pp. 1–6.
- [38] J. Elfring, E. Torta, and R. van de Molengraft, "Particle filters: A hands-on tutorial," *Sensors*, vol. 21, no. 2, p. 438, Jan. 2021.
- [39] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," *Handbook Nonlinear Filtering*, vol. 12, nos. 656–704, p. 3, 2009.
- [40] J.-L. Blanco. (2023). *Resampling Methods for Particle Filtering*. [Online]. Available: <https://www.mathworks.com/matlabcentral/fileexchange/24968-resampling-methods-for-particle-filtering>
- [41] A. Doucet, N. D. Freitas, and N. J. Gordon, *Sequential Monte Carlo Methods in Practice*, vol. 1. Cham, Switzerland: Springer, 2001.
- [42] J. F. G. D. Freitas, M. Niranjana, A. H. Gee, and A. Doucet, "Sequential Monte Carlo methods to train neural network models," *Neural Comput.*, vol. 12, no. 4, pp. 955–993, Apr. 2000.
- [43] J. Gu and R. Kanai, "What contributes to individual differences in brain structure?" *Frontiers Hum. Neurosci.*, vol. 8, p. 262, Apr. 2014.
- [44] M. M. Shanechi, "Brain–machine interfaces from motor to mood," *Nature Neurosci.*, vol. 22, no. 10, pp. 1554–1564, Oct. 2019.