Received: 8 March 2021 | Revised: 25 March 2021 | Accepted: 29 March 2021

DOI: 10.1002/hbm.25442

RESEARCH ARTICLE

WILEY

Bayesian estimation of maximum entropy model for individualized energy landscape analysis of brain state dynamics

Jiyoung Kang[1,2] | Seok-Oh Jeong[3] | Chongwon Pae[1,2] | Hae-Jeong Park[1,2,4]

[1]Center for Systems and Translational Brain Science, Institute of Human Complexity and Systems Science, Yonsei University, Seoul, South Korea

[2]Department of Nuclear Medicine, Psychiatry, Yonsei University College of Medicine, Seoul, South Korea

[3]Department of Statistics, Hankuk University of Foreign Studies, Yong-In, Seoul, South Korea

[4]Graduate School of Medical Science, Brain Korea 21 Project, Yonsei University College of Medicine, Seoul, South Korea

Correspondence
Hae-Jeong Park, Department of Cognitive Science, Center for Systems and Translational Brain Science, Yonsei University, 50 Yonsei-ro, Sinchon-dong, Seodaemun-gu 120-752, Seoul, South Korea.
Email: parkhj@yonsei.ac.kr

Funding information
National Research Foundation of Korea, Grant/Award Numbers: NRF-2017M3C7A1030750, NRF-2017H1D3A1A01053094

Abstract

The pairwise maximum entropy model (MEM) for resting state functional MRI (rsfMRI) has been used to generate energy landscape of brain states and to explore nonlinear brain state dynamics. Researches using MEM, however, has mostly been restricted to fixed-effect group-level analyses, using concatenated time series across individuals, due to the need for large samples in the parameter estimation of MEM. To mitigate the small sample problem in analyzing energy landscapes for individuals, we propose a Bayesian estimation of individual MEM using variational Bayes approximation (BMEM). We evaluated the performances of BMEM with respect to sample sizes and prior information using simulation. BMEM showed advantages over conventional maximum likelihood estimation in reliably estimating model parameters for individuals with small sample data, particularly utilizing the empirical priors derived from group data. We then analyzed individual rsfMRI of the Human Connectome Project to show the usefulness of MEM in differentiating individuals and in exploring neural correlates for human behavior. MEM and its energy landscape properties showed high subject specificity comparable to that of functional connectivity. Canonical correlation analysis identified canonical variables for MEM highly associated with cognitive scores. Inter-individual variations of cognitive scores were also reflected in energy landscape properties such as energies, occupation times, and basin sizes at local minima. We conclude that BMEM provides an efficient method to characterize dynamic properties of individuals using energy landscape analysis of individual brain states.

KEYWORDS
Bayesian estimation, energy landscape, maximum entropy model, nonlinear brain dynamics, resting state fMRI

Jiyoung Kang, Seok-Oh Jeong, and Hae-Jeong Park contributed equally to this study.

1 | INTRODUCTION

The human brain has been viewed as a dynamic complex system that possesses multiple stable states over a stable circuitry (Breakspear, 2017; J. Cabral, Kringelbach, & Deco, 2014; Deco & Jirsa, 2012; Deco, Tononi, Boly, & Kringelbach, 2015; Freyer et al., 2011; Freyer, Roberts, Ritter, & Breakspear, 2012; Kelso, 2012; M. I. Rabinovich & Varona, 2011; Tognoli & Kelso, 2014). In this view, the brain "at rest" is configured to have multistability (J. Cabral, Kringelbach, & Deco, 2017; Deco, Jirsa, & McIntosh, 2013; Deco, Senden, & Jirsa, 2012; Golos, Jirsa, & Dauce, 2015; Kang, Pae, & Park, 2017), which is associated with the formulation of diverse brain functions "at task" such as memory, decision-making, and sensory processing (Cabessa & Villa, 2014; Knierim & Zhang, 2012; Kringelbach, McIntosh, Ritter, Jirsa, & Deco, 2015; Mikhail I. Rabinovich, Huerta, Varona, & Afraimovich, 2008; Tognoli & Kelso, 2014). The multistable dynamic properties of the resting and task brain have been explored in terms of energy landscape of brain states, where energy of a state is defined by the negative log probability of the occurrence of the state (e.g., low energy for a frequent state). To estimate probability of all neural states, a pairwise maximum entropy model (MEM) has often been used. In MEM, the probability of a state is determined by interactions within the system and composes the energy landscape of the system's states. Energy landscape analysis of brain states enables identification of stable states (local minima, called "attractors") of the brain system and characterization of transition rates among those states, which may characterize individuals or brain diseases.

So far, energy landscape analysis has been applied to explore dynamics in the large-scale functional brain network, such as the default mode and prefrontal networks, in the resting state fMRI (rsfMRI Ezaki, Fonseca Dos Reis, Watanabe, Sakaki, & Masuda, 2020; Kang et al., 2017; Kang, Pae, & Park, 2019; Watanabe et al., 2013; Watanabe, Hirose, et al., 2014), sleep (Watanabe, Khan, et al., 2014), and vision-related areas for bistable perception (Watanabe, Masuda, Megumi, Kanai, & Rees, 2014). The abnormality of dynamic properties in autism spectrum disorder (Watanabe & Rees, 2017) as well as schizophrenia (Joana Cabral et al., 2013; Loh, Rolls, & Deco, 2007) has also been researched using the energy landscape analysis. In those studies, brain states are defined by distributed activity patterns across brain regions, and the brain belongs to one of the states at each measurement time point (Ezaki, Sakaki, Watanabe, & Masuda, 2018; Gu et al., 2018; Kang et al., 2017; Watanabe et al., 2013; Watanabe, Hirose, et al., 2014; Watanabe, Kan, et al., 2014; Watanabe, Masuda, et al., 2014).

Despite the unique advantage of energy landscape analysis for modeling nonlinear dynamics of the brain, the application of this approach to an individual has been hindered by the requirement of a large sample size for model parameter estimation. Since the acquisition of a large sample data from an individual is not practical in most clinical settings, the MEM analysis has mainly been restricted to the fixed-effect group-level studies by concatenating time series of individuals without considering individual differences. In the estimation of MEM model parameters with a network size N, the number of parameters equals to $M = N \times (N+1)/2$ (see Section 2) and the estimation procedure requires time series data of length approximately 10 times according to the rule of 10 (Harrell Jr., Lee, Califf, Pryor, & Rosati, 1984; Harrell Jr., Lee, & Mark, 1996). According to this rule, for example, a system with an $N = 15$ has $M = 120$ parameters needs approximately 1,200 (120×10) temporally independent time points, which is generally much longer than the conventional rsfMRI data for a subject (around 300 time points or $5 \sim 10$ min per person). According to Ezaki, Watanabe, Ohzeki, and Masuda (2017), the amount of data for a reliable estimation of the pairwise MEM is proportional to 2^N, and thus the reliable estimation of the system with $N = 15$ nodes requires $C \times 2^{15}$ sample points, where C is a constant. Due to the temporal dependency (redundancy) of rsfMRI samples, more data is needed to fit the model reliably. It is for this reason that fixed-effect analysis using group-concatenated data has been used so far. However, the fixed-effect group-level analysis limits the application of MEM and evaluation of state dynamics specific to each individual. To mitigate the small sample size issue in characterizing an individual's dynamic properties, we developed a Bayesian model estimation scheme for MEM using a variational Bayes approach (BMEM) to incorporate a priori information.

To evaluate the performance of BMEM, we conducted simulation studies by changing sample sizes and a priori information compared to conventional maximum likelihood estimation (MLE). We also conducted an experimental study using rsfMRI data from the Human Connectome Project (HCP) database (Van Essen et al., 2012) to assess how well our proposed method differentiates the model parameters of one individual from others (i.e., subject specificity) and to evaluate the usefulness of MEM in characterizing human cognition (available in the HCP database). For this, we used and compared sparse canonical correlation analysis (sCCA) results between cognitive scores and MEM parameters, between cognitive scores and energy landscape features with that of functional connectivity matrix (FC).

The current article is composed of three main parts. First, we described a mathematical formulation for the BMEM. Second, the BMEM scheme was tested using a simulated dataset with respect to sample sizes, empirical prior types, and estimation methods. Third, we applied BMEM to the resting-state of fMRI data of 468 HCP subjects and evaluated subject specificity and characterizability of human cognition using MEM parameters and local minima of energy landscapes compared with correlation-based functional connectivity. We expected that BMEM would provide a new framework for exploring individual brains from the perspective of dynamic systems and become a useful tool for characterizing or identifying healthy and abnormal brain dynamics.

2 | METHODS

2.1 | The pairwise maximum entropy model

The energy landscape analysis of brain states consists of steps for defining brain states, construction of a pairwise MEM model for state dynamics, optimization of MEM model parameters to fit probability distributions of empirical brain states using states generated by the model, and analysis of the energy landscape. In this paragraph, a brief

introduction of MEM model estimation is presented. Details of the model construction are provided in our previous study (Kang et al., 2017, 2019) (for the mathematical details, see a review [Yeh et al., 2010]).

Suppose the brain state space is represented by

$$\mathcal{S} = \left\{ \boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)^\top \in \left\{0,1\right\}^N \middle| \boldsymbol{\sigma} \text{ is a possible state} \right\}, \tag{1}$$

where the value of σ_i (i = 1, 2, ..., N) is either 0 (inactive) or 1 (active), indicating a local activity at a node (brain region) i, and N denotes the total number of nodes (or ROIs). Maximizing the entropy of the state space, $H(p) = -\sum_{\boldsymbol{\sigma}\in\mathcal{S}} p(\boldsymbol{\sigma}) \log p(\boldsymbol{\sigma})$, where $p(\boldsymbol{\sigma})$ is the probability mass for the state $\boldsymbol{\sigma}$, leads to a special form of the probability distribution of each state. This probability is given by a Boltzmann distribution as below

$$p(\boldsymbol{\sigma}|\boldsymbol{h},\boldsymbol{J}) = \frac{\exp\{-E(\boldsymbol{\sigma}|\boldsymbol{h},\boldsymbol{J})\}}{\sum_{\boldsymbol{\sigma}'\in\mathcal{S}} \exp\{-E(\boldsymbol{\sigma}'|\boldsymbol{h},\boldsymbol{J})\}}, \tag{2}$$

where

$$E(\boldsymbol{\sigma} \mid \boldsymbol{h},\boldsymbol{J}) = -\sum_{i=1}^{N} h_i\sigma_i - \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} J_{ij}\sigma_i\sigma_j, \tag{3}$$

is the energy, and h_i's and J_{ij}'s are model parameters. This model's total number of parameters is then $M = N + N(N-1)/2 = N(N+1)/2$. If h_i is large, the energy is smaller (the probability is larger) with $\sigma_i = 1$ than with $\sigma_i = 0$ and hence the i-th ROI tends to be activated. If J_{ij} is large, both the i-th ROI and j-th ROI tend to be activated at the same time. For convenience, we reparametrize the model $p(\boldsymbol{\sigma} \mid \boldsymbol{h}, \boldsymbol{J})$ as follows:

$$\tilde{\boldsymbol{\sigma}} = (\sigma_1, \sigma_2, \ldots, \sigma_N, \sigma_1\sigma_2, \sigma_1\sigma_3, \ldots, \sigma_{N-1}\sigma_N)^\top \in \mathbb{R}^M, \tag{4}$$

$$\boldsymbol{\theta} = (h_1, h_2, \ldots, h_N, J_{12}, J_{13}, \ldots, J_{N-1,N})^\top \in \mathbb{R}^M, \tag{5}$$

$$E(\boldsymbol{\sigma}|\boldsymbol{\theta}) = -\boldsymbol{\theta}^\top \tilde{\boldsymbol{\sigma}}, \tag{6}$$

$$p(\boldsymbol{\sigma}|\boldsymbol{h},\boldsymbol{J}) = \frac{\exp\{-E(\boldsymbol{\sigma}|\boldsymbol{\theta})\}}{\sum_{\boldsymbol{\sigma}'\in\mathcal{S}} \exp\{-E(\boldsymbol{\sigma}'|\boldsymbol{\theta})\}} = \frac{\exp(\boldsymbol{\theta}^\top\tilde{\boldsymbol{\sigma}})}{\sum_{\boldsymbol{\sigma}'\in\mathcal{S}} \exp(\boldsymbol{\theta}^\top\tilde{\boldsymbol{\sigma}}')}. \tag{7}$$

Note that Boltzmann distributions belong to the exponential family and that $\boldsymbol{\theta}$ is the vector of natural parameters.

2.2 | Model estimation: Maximum likelihood method

Let $\mathcal{D}$ be the set of observed brain states at time points t = 1, 2, ..., T:

$$\mathcal{D} = \{\boldsymbol{\sigma}(t), t = 1, 2, \ldots, T\}. \tag{8}$$

Then, the log-likelihood function of the data set $\mathcal{D}$ is given as

$$\ell(\boldsymbol{\theta}) = \log p(\mathcal{D}|\boldsymbol{\theta}) = \sum_{t=1}^{T} \boldsymbol{\theta}^\top \tilde{\boldsymbol{\sigma}}(t) - T\log\sum_{\boldsymbol{\sigma}\in\mathcal{S}} e^{\boldsymbol{\theta}^\top\tilde{\boldsymbol{\sigma}}} \tag{9}$$

The maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$ is defined by the maximizer of $\ell(\boldsymbol{\theta})$. One may conduct this maximization procedure using the gradient ascent algorithm:

$$\boldsymbol{\theta} \leftarrow \boldsymbol{\theta}_0 + \rho\nabla\ell(\boldsymbol{\theta}_0), \tag{10}$$

where $\boldsymbol{\theta}_0$ is the parameter vector before each updating step, $\nabla\ell(\boldsymbol{\theta}_0)$ is the gradient of $\ell(\boldsymbol{\theta})$ assessed at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, and $\rho > 0$ is a learning rate. Note that the gradient is simply given by

$$\nabla\ell(\boldsymbol{\theta}) = T\left(\langle\tilde{\boldsymbol{\sigma}}\rangle_{\text{empirical}} - \langle\tilde{\boldsymbol{\sigma}}\rangle_{\text{model}}\right), \tag{11}$$

where $\langle\tilde{\boldsymbol{\sigma}}\rangle_{\text{empirical}}$ is the average of observed $\tilde{\boldsymbol{\sigma}}(t)$'s and $\langle\tilde{\boldsymbol{\sigma}}\rangle_{\text{model}}$ is the model mean vector generated with model parameter $\boldsymbol{\theta}$.

2.3 | Bayesian formulation and variational Bayes method

Now we develop a Bayesian treatment by introducing a prior distribution for the parameters $\boldsymbol{\theta} \in \mathbb{R}^M$. Choosing the normal distribution for the prior, $\boldsymbol{\theta} \sim N\left(\boldsymbol{\eta}, \text{diag}(\boldsymbol{\alpha})^{-1}\right)$, with hyperparameters $\boldsymbol{\eta} \in \mathbb{R}^M$ and $\boldsymbol{\alpha} \in \mathbb{R}_+^M$, it is intractable to derive the posterior since the normal distribution is not a conjugate prior for the Boltzmann distribution. So we consider a variational approximation for the posterior (for details, see Bishop, 2006) by using the normal distribution $q(\boldsymbol{\theta}) = N\left(\boldsymbol{\mu}, \text{diag}(\boldsymbol{\beta})^{-1}\right)$ with $\boldsymbol{\mu} \in \mathbb{R}^M$ and $\boldsymbol{\beta} \in \mathbb{R}_+^M$. Precisely,

$$p(\boldsymbol{\theta}|\mathcal{D},\boldsymbol{\eta},\boldsymbol{\alpha}) \approx q(\boldsymbol{\theta}) = N\left(\boldsymbol{\mu}, \text{diag}(\boldsymbol{\beta})^{-1}\right) = \prod_{j=1}^{M} \frac{\beta_j^{1/2}}{\sqrt{2\pi}} \exp\left\{-\frac{\beta_j}{2}(\theta_j - \mu_j)^2\right\}, \tag{12}$$

which is, like the prior, a factorized form of univariate normal distributions again. It is well known that this variational approximate solution for the variational distribution, q (parameterized by $\boldsymbol{\mu}$ and $\boldsymbol{\beta}$) is obtained by optimizing the evidence lower bound (ELBO), often called free energy, $\mathcal{F}(q,\boldsymbol{\eta},\boldsymbol{\alpha})$:

$$\begin{aligned}\mathcal{F}(q,\boldsymbol{\eta},\boldsymbol{\alpha}) \approx &\sum_{t=1}^{T} \boldsymbol{\mu}^\top\tilde{\boldsymbol{\sigma}}(t) - T[\log\textstyle\sum_{\boldsymbol{\sigma}\in\mathcal{S}} e^{\boldsymbol{\eta}^\top\tilde{\boldsymbol{\sigma}}} + \langle\tilde{\boldsymbol{\sigma}}\rangle_{\boldsymbol{\eta}}^\top(\boldsymbol{\mu}-\boldsymbol{\eta}) \\ &+ \frac{1}{2}\left\{\text{tr}\left(\text{diag}(\boldsymbol{\beta})^{-1}C_{\boldsymbol{\eta}}\right) + (\boldsymbol{\mu}-\boldsymbol{\eta})^\top C_{\boldsymbol{\eta}}(\boldsymbol{\mu}-\boldsymbol{\eta})\right\} + \frac{1}{2}\textstyle\sum_{j=1}^{M}\log\alpha_j \\ &- \frac{1}{2}\textstyle\sum_{j=1}^{M}\alpha_j\left\{(\mu_j-\eta_j)^2 + \beta_j^{-1}\right\} - \frac{1}{2}\textstyle\sum_{j=1}^{M}\log\beta_j + \text{constant},\end{aligned} \tag{13}$$

with $C_{\boldsymbol{\eta}} = \text{Cov}_{\boldsymbol{\eta}}(\tilde{\boldsymbol{\sigma}})$. See Appendix A for technical details. Now, by maximizing $\mathcal{F}(q,\boldsymbol{\eta},\boldsymbol{\alpha})$ with respect to $\boldsymbol{\mu}$ and $\boldsymbol{\beta}$, we have the posterior mean $\boldsymbol{\mu}$ and precision $\boldsymbol{\beta}$ as follows:

$$\boldsymbol{\mu} = \boldsymbol{\eta} + T\mathbf{A}_{\eta,\alpha}^{-1}\left(\langle\tilde{\sigma}\rangle_{\text{empirical}} - \langle\tilde{\sigma}\rangle_{\eta}\right), \tag{14}$$

$$\boldsymbol{\beta} = \boldsymbol{\alpha} + T\boldsymbol{c}_{\eta}, \tag{15}$$

where $\mathbf{A}_{\eta,\alpha} = \text{diag}(\boldsymbol{\alpha}) + TC_{\eta}$ and $\boldsymbol{c}_{\eta}$ is the vector composed of the diagonal elements of C_{η}.

Thus, the prior distribution $N\left(\boldsymbol{\eta}, \text{diag}(\boldsymbol{\alpha})^{-1}\right)$ is updated to the posterior distribution $N\left(\boldsymbol{\mu}, \text{diag}(\boldsymbol{\beta})^{-1}\right)$. The observed data play a role in the Bayesian update algorithm so that the mean vector is updated from the prior along the direction of the gradient of log-likelihood. The precision is improved by the magnitude of the variance of observed data. This work does not consider any kind of iterative procedure such as expectation–maximization (EM-like) algorithms to obtain the hyperparameters based on the dataset. We simply propose an approximate algorithm to obtain the posterior *given* prior information on the model parameters.

For the prior, we considered three types of information: (a) zero-mean prior with $\boldsymbol{\eta}$ all components being equal to zero and with the precision α_j = 6.67 for all; (b) empirical prior obtained by group averaging; and (c) empirical prior from the group concatenated data. For (b), we first obtained variational Bayes model parameter estimates for each individual using the zero-mean prior in (a), and then averaged them to get the mean vector $\boldsymbol{\eta}$ for the empirical prior. For (c), we concatenated group data and estimated parameters for group concatenated data using the zero-mean priors in (a) and use it for the mean vector $\boldsymbol{\eta}$ of the empirical prior. The prior precision $\boldsymbol{\alpha}$ can be viewed as a tuning parameter controlling the amount of shrinkage to the prior mean for the resulting variational Bayes parameter estimates. In this work, for parsimony, we used a common positive number 6.67 for the prior precision so that the probability for a normal random variable to deviate from its mean by 2.58-*SD* ($2.58 \times \frac{1}{\sqrt{6.67}} \approx 1$) is about 1% (99% stays within the range of ±1). This conforms with the results observed in our empirical studies that the absolute values of the MEM parameter estimate rarely exceeded 1.

Figure 1 presents a procedure for BMEM analysis of rsfMRI data of an individual.

2.4 | Resting-state fMRI data of HCP

For the simulation and experimental analysis, we used rsfMRI time series at the 15 subcortical and limbic brain regions of 468 participants (192 males, 276 females, age: 29.2 ± 3.5 years) from the HCP database (Van Essen et al., 2012), which were used in our previous study (Kang et al., 2017). For the simulation, we estimated MEM of two differently sized systems, that is, a system with eight regions of interest

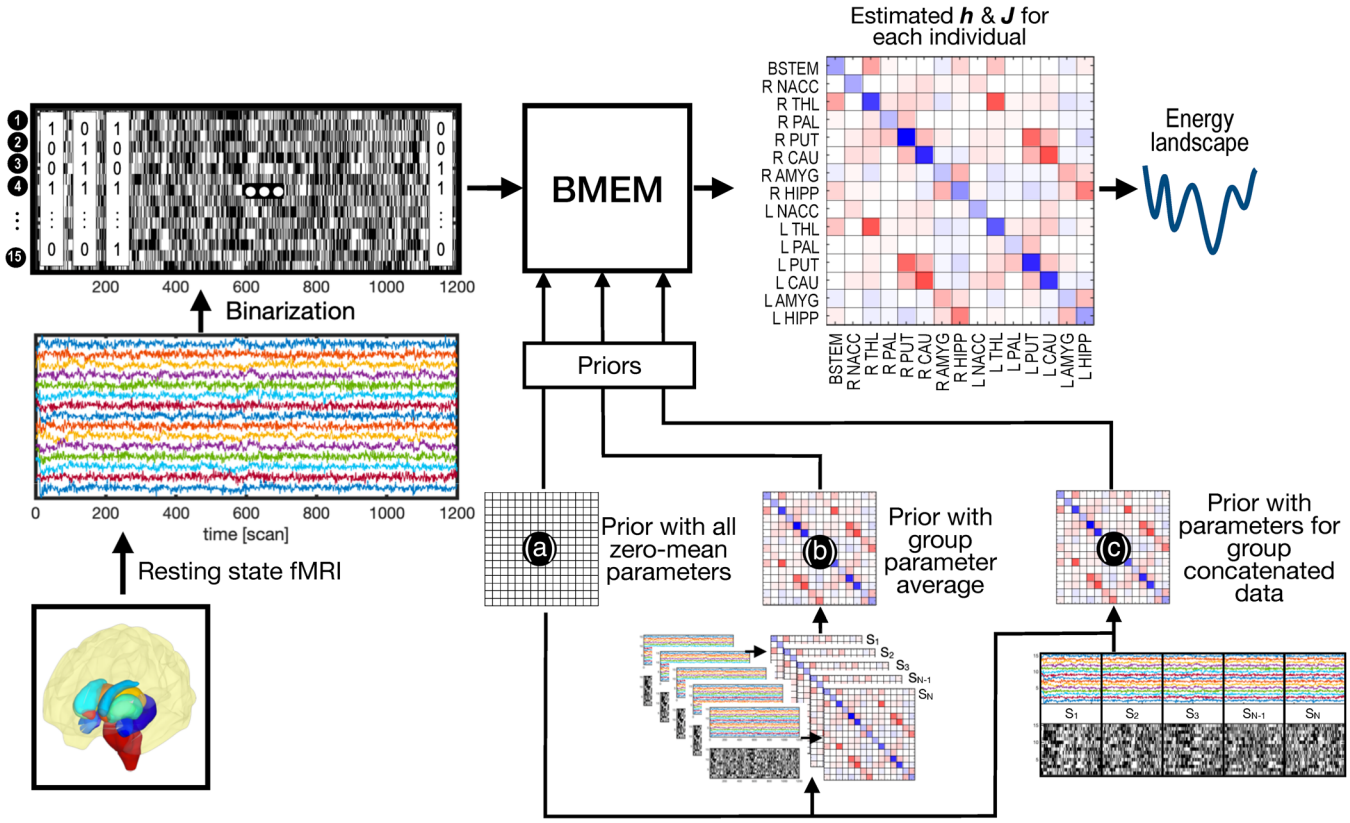

FIGURE 1 An illustration of the variational Bayesian MEM analysis (BMEM) procedure. The rsfMRI signals were binarized to represent local brain states used in BMEM to estimate *h* and *J* parameters. Model parameters were estimated using BMEM with three different types of priors—(a) the zero-mean prior, (b) empirical priors with a group average of individual MEM parameter estimates with the zero-mean prior, and (c) empirical priors with MEM parameter estimates for group concatenated data using the zero-mean prior

(ROI) and a system with 15 ROIs in the subcortex and limbic cortex. Eight regions were the hippocampus (a limbic cortex), amygdala, caudate, putamen, pallidum, thalamus, nucleus accumbens of the left hemispheres, and brainstem. For the 15-ROIs system, seven homologous regions in the right hemispheres were added to the 8-ROIs system. A time series of the first eigenvariate of principal component analysis was used for the time series of all the voxels within each region in the subsequent analysis. This eigenvariate is advantageous in representing the time course of an ROI over a simple averaging when an ROI is composed of voxels with heterogeneous activities (B. Park, Ko, Lee, & Park, 2013). In line with our previous study (Kang et al., 2017), we conducted detrending, despiking and regressing out motion artifacts (rigid motion and their derivatives), and selected zero as the threshold to binarize regional states after global regression (global signals in the whole-brain mask) to emphasize regional changes in signal levels in specific brain states.

All (binarized) 8 or 15 regional states compose a state vector. Due to a demand for a large sample size to estimate brain states (for all 2^8 or 2^{15} possible states), we concatenated all temporal samples from four sessions of 468 participants (the total number of state samples are 1,200 samples × 4 sessions × 468 participants) and estimated parameters of the pairwise MEM for the concatenated data using the methods described in the following sections. The estimated two parameter sets (8-ROIs and 15-ROIs) from the concatenated data were used as ground-truth parameters to make the simulation biologically plausible. The details of signal processing can be found in Kang et al. (2017).

2.5 | Evaluation of BMEM with simulations

Simulation studies were conducted to explore the performance of BMEM with regard to sample sizes and prior types for the two networks of 8- and 15-ROIs.

First, we performed two types of simulations for the network with 8-ROIs. To evaluate the performance of each estimation method (MLE and VB) with respect to sample sizes (50, 200, 500, and 1,200) and prior types (zero-mean and empirical priors), we generated 500 datasets of time series for each condition from the ground-truth model parameter vector.

Next, we implemented individual differences within a group to simulate group studies. MEM parameters of individual 15-ROIs systems were set by adding Gaussian noises to the ground-truth model parameter θ_{group} (as a group ground truth). Thus, a unique MEM parameter θ_i is given by

$$\theta_i = \theta_{group} + \varepsilon_i,\ \ \varepsilon_i \sim N\left(0, s^2\right), \tag{16}$$

where s represents *SD* of individuals (or jitters) and was set to $s = 0.1$. Considering that the interaction parameters are generally in the range between 0 and 0.5, one *SD* (0.1) of each parameter from the θ_{group} may be a reasonable choice, especially when multiple combinations of deviated parameters represent individual variations within a single group.

Sample signals for 30 individuals were generated based on 30 MEM parameter sets. For these data, the performance of all the estimation methods (MLE and VB) was evaluated with various sample sizes (50, 200, 500, and 1,200) and prior types (a zero-mean prior and two empirical priors). The hyperparameter $\boldsymbol{\eta}$ for two-types of empirical priors was obtained by the average of parameter estimates from all the subjects and by the parameter estimate from the concatenated sample dataset. In addition, we set the variances of the normal priors to 0.15 (precision = 6.67) by default; that is, $\boldsymbol{\theta} \sim N(\boldsymbol{\eta}, \text{diag}[0.15])$. In all simulations, to evaluate the accuracy of parameter estimation, we used root mean square error (RMSE) between estimated and true parameters:

$$RMSE = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(\theta_{m,\text{estimated}} - \theta_{m,\text{true}}\right)^2}, \tag{17}$$

where M is the total number of MEM parameters (h_i and J_{ij}).

2.6 | Evaluation with experimental data: Capability of detecting inter-individual variations

We evaluated the capability of BMEM in characterizing each individual's dynamic properties with MEM parameters and with energy landscape features in the rsfMRI of the HCP database by exploring an inter-individual similarity matrix for those features across individuals. The inter-individual similarity matrix was defined using a Pearson correlation coefficient for MEM parameters across subjects. The size of the inter-individual similarity matrix is the number of subjects by the number of subjects. Since we do not have an established criterion to evaluate the degree of the inter-individual variation of MEM parameters, we considered the inter-individual variation of functional connectivity as a reference to evaluate the level of inter-individual variation of the MEM model parameters. To evaluate the level of inter-individual variations of MEM parameters compared to that of the functional connectivity, we conducted a Pearson correlation analysis between the upper triangle of the inter-individual similarity matrix of MEM parameters and that of functional connectivity. If the Pearson correlation between the two inter-individual similarity matrices is high, the individuals with similar functional connectivity would have similar MEM parameters and vice versa. If so, the MEM parameters would show high inter-individual variation comparable to functional connectivity and reflect a biologically plausible aspect of each individual's resting-state brain.

The same evaluation was also applied to the energy landscape features to test the capability of representing state dynamics specific to each individual. For this, we constructed the inter-individual similarity matrix of energy landscapes and correlated it with that of functional connectivity. In this analysis, we also considered the inter-individual similarity matrix of functional connectivity as a reference to assess the degree of inter-individual similarity matrix of energy landscapes.

For this analysis, we constructed energy landscapes based on the pairwise MEM estimation, following the procedure described in previous papers (Kang et al., 2017; Watanabe, Hirose, et al., 2014). Briefly, to construct an energy landscape, we defined the distance between two states as the number of elements (bits) that differ between two state vectors. Based on a gradual state transition assumption, we examined the energy landscape by changing one element of the state vector for each step. The local minima (also called stable states) were defined to have lower energy (more frequent) relative to their neighbors. We calculated depths, basin sizes, and occupation times for local minima, which were used to evaluate the subject specificity of BMEM. The local minima, occupation times, and basin sizes were defined according to previous work by Kang et al. (2017). Local minima indicate states that have lower energy relative to their neighbors. A total of 26 local minima were defined by the union of (1) and (2). The basin region of a local minimum was defined as all the states with gradients toward the local minimum. The basin size of a local minimum was defined as the fraction of states that belong to the basin of the local minimum. The occupation time of each local minimum was defined as a sum of the probabilities of the states in its basin region.

Of note, although the energy landscape features (such as energies at the local minima, their occupation times, and basin sizes) are extracted from the MEM parameters, the energy landscape features directly reflect information about frequent brain states and their properties in the dynamics, which cannot be explored directly from the interaction parameters in FC or MEM. All the analysis was done with the first session (1,200 samples) rsfMRI data of the HCP database, with the zero-mean prior (precision = 6.67).

2.7 | Evaluation with experimental data: Capability of reflecting cognitive scores

In order to evaluate the applicability of the energy landscape analysis using BMEM in differentiating individuals and in exploring neural correlates for human behavior, we have associated MEM parameters and brain state energies with gender, age, and 15 cognitive scores; the 15 cognitions scores used are Mini Mental Status Exam Total Score (MMSE_Score), Pittsburgh Sleep Questionnaire Total Score (PSQI_Score), Penn Matrix Test: Number of Correct Responses (PMAT24_A_CR), Penn Matrix Test: Total Skipped Items (PMAT24_A_SI), Short Penn CPT Sensitivity (SCPT_SEN), Short Penn CPT Specificity (SCPT_SPEC), NIH Toolbox Flanker Inhibitory Control and Attention Test Age-Adjusted Scale Score (Flanker_AgeAdj), NIH Toolbox Picture Sequence Memory Test Age-Adjusted Scale Score (PicSeq_AgeAdj), NIH Toolbox Dimensional Change Card Sort Test Age-Adjusted Scale Score (CardSort_AgeAdj), NIH Toolbox List Sorting Working Memory Test Age-Adjusted Scale Score (ListSort_AgeAdj), NIH Toolbox Oral Reading Recognition Test Age-Adjusted Scale Score (ReadEng_AgeAdj), NIH Toolbox Picture Vocabulary Test Age-Adjusted Scale Score (PicVocab_AgeAdj), NIH Toolbox Pattern Comparison Processing Speed Test Age-Adjusted Scale Score (ProcSpeed_AgeAdj), Penn Line Orientation: Total Number Correct (VSPLOT_TC), Penn Word Memory: Total Number of Correct Responses (IWRD_TOT).

To associate MEM and energy landscape properties with cognitive performances, we conducted sCCA (Chu, Liao, Ng, & Zhang, 2013), which is multivariate data analysis for finding the correlation between behaviors (cognitions) and neurobiological data (BMEM parameters). sCCA is a sparse solution of the canonical correlation analysis (CCA), which is described below.

Given two sets of variables X and Y, which are composed of column vectors $X = [X_1, X_2, \cdots, X_p]$ and $Y = [Y_1, Y_2, \cdots, Y_q]$, let us consider two column vectors U_1' and V_1' that are driven from the weighted sum of the column vectors of X_i and Y_i as below.

$$U_1' = \boldsymbol{a}'^T X = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p, \tag{18}$$

$$V_1' = \boldsymbol{b}'^T Y = b_{11}Y_1 + b_{12}Y_2 + \cdots + b_{1q}Y_q. \tag{19}$$

The sizes of vectors U_1' and V_1' are S subjects $\times$ p variables and S subjects $\times$ q variables. CCA is to search the weights, called canonical coefficients, $\boldsymbol{a}'$ and $\boldsymbol{b}'$ that maximize the correlation between the two vectors U_1' and V_1'.

$$(\boldsymbol{a}, \boldsymbol{b}) = \underset{\boldsymbol{a}', \boldsymbol{b}'}{\operatorname{argmax}} \operatorname{corr}(U_1', V_1') \tag{20}$$

The first pair of canonical variables are $U_1 = \boldsymbol{a}^T X$ and $V_1 = \boldsymbol{b}^T Y$. The second pair of canonical variables maximize the same correlation under the constraint that they are uncorrelated from the first pair of canonical variables (U_1, V_1). This procedure continues up to minimum of p and q. sCCA finds the optimal solution $(\boldsymbol{a}, \boldsymbol{b})$ under the additional constraint of sparsity for $(\boldsymbol{a}', \boldsymbol{b}')$. For the details, see sCCA (Chu et al., 2013).

In the current study, cognitive scores were assigned to X, and neurobiological features were assigned to Y. Three different biological features were associated with cognition scores using sCCA: (a) MEM parameters (a vector of $\boldsymbol{h}$ and an upper triangle of $\boldsymbol{J}$); (b) a vector of energy landscape features; and (c) a vector of an upper triangle of functional connectivity of each individual. The energy landscape features are composed of the union of (a) energies at the local minima of the group concatenated data set, (b) energies at the local minima of the group average energy landscape derived from the averaged probability distribution of all individuals, and (c) the occupation times and basin sizes for the two lowest local minima specific to each individual. sCCA between functional connectivity and the cognitive scores was evaluated as a reference to sCCA for BMEM. All behavioral scores and neurobiological scores were normalized across subjects by z-transformation.

3 | RESULTS

RMSEs between the ground-truth parameters of the 8-ROIs system and estimated parameters by MLE, variational Bayes MEM (BMEM) approach with zero priors (VB_z) and variational Bayes approach with empirical priors (VB_{gc}) for the 500 generated datasets with four

different sample sizes (T = 50, 200, 500, and 1,200) are presented in Figure 2a. All methods tend to show better reliability as the sample size increases. For the datasets with short lengths (T = 50 and T = 200), VB with group empirical prior tends to give model parameter estimates having smaller RMSEs than the others. Similar results are found in Figure 2b, wherein VB with an empirical prior obtained from concatenated data (VBgc) shows the lowest RMSE in the 15-ROIs system across 30 different individuals. For datasets with varying sizes of sample (sample sizes T = 50, 200, 500, and 1,200), empirical priors improved the reliability of parameter estimation of VB for the individual data compared to MLE and VB with zero-mean priors. RMSEs for VB with empirical priors were lower than those of MLE or VB with zero priors. In particular, VB with empirical priors of a group-concatenated dataset showed better parameter estimation (lower RMSE) than VB with empirical priors of group averaged parameters. This tendency was exemplified in a subject using a system of 15 ROIs (Figure 3).

When we average estimated parameters of all subjects to estimate group-level model parameters, that is, θ_{group} in Equation (16), the averaged parameters are highly consistent with the group ground-truth model parameters (Figure 4). Both the 8-ROIs (Figure 4a) and 15-ROIs systems (Figure 4b) showed a high correlation with the ground-truth group model parameters regardless of estimation methods (MLE or VBgc). VBs for larger datasets were highly accurate in all analyses, but VBs for small datasets still worked reasonably well in the parameter estimation.

Since the best results were obtained using VB with priors obtained from concatenated datasets in the simulation, we applied it to our analysis of the rsfMRI data of HCP, used previously in the Kang et al. (2017). In this study, we used the zero-mean prior in the BMEM estimation for the group concatenated data with precision α_j = 6.67 for all j.

The relationship between the group union of all local minima (estimated) for all individuals (iLMs) and the distribution of local minima estimated using the group-concatenated data (gLM) is presented in Figure 5. There are correspondences between them, particularly in the frequently appearing local minima across subjects. This result suggests the capability of BMEM in analyzing the individual energy landscape of brain states, particularly in identifying local minima at each individual. Most local minima of the group concatenated data (gLMs) serve as references for clusters of iLMs. Some local minima frequently found in many individuals are not detected as local minima in the model of group concatenated data. Furthermore, various frequencies indicate variations of local minima across individuals. This result suggests that although the group model with the group concatenated data grossly reflects the group population's brain states, individual MEM is needed to explore the subject-specific variability.

We calculated inter-individual similarity matrices (cross-correlation matrices across subjects) of functional connectivity, MEM parameters, energy levels at all states, and energy levels at the group local minima. The inter-individual similarity matrix of functional connectivity was highly correlated with those of MEM, energy levels at all states, energy levels at the group local minima (correlation coefficients = .88 (p = .0000, degree of freedom [df] = 466), .86 (p = .0000), and .57 (p = .0000), respectively; Figure 6). Inter-individual variation (reflected in the inter-individual similarity matrix) of the energy levels

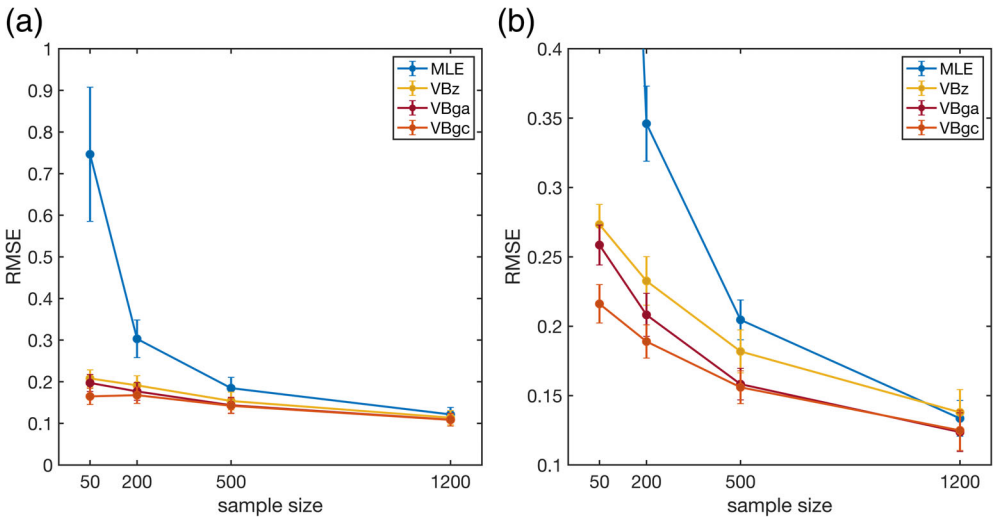

FIGURE 2 Performance evaluation of model parameter estimations using synthetic data. Root mean square errors (RMSE) for all the parameters estimated by maximum likelihood estimation (MLE), variational Bayes approach with the zero prior (VBz), variational Bayes approach with the prior from the group average (VBga), and variational Bayes approach with the prior derived from group concatenation (VBgc), compared to the ground truth. All state vectors that were used in the estimations were synthesized from the ground truth MEM parameters. (a) Mean and *SD* of RMSEs were evaluated for model estimation reliability within a subject (a model parameter set) using a total of 500 repetitions of sample generation with four sample sizes (T = 50, 200, 500, and 1,200) in the 8-ROIs systems. (b) Mean and *SD* of RMSEs for parameters in the 15-ROIs systems were evaluated at the four different sample sizes for 30 individuals deviated from (Gaussian noise) a group-truth parameter, as a group level analysis. The RMSE for MLE with 50 samples was very high beyond the current range between 0.1 and 0.4

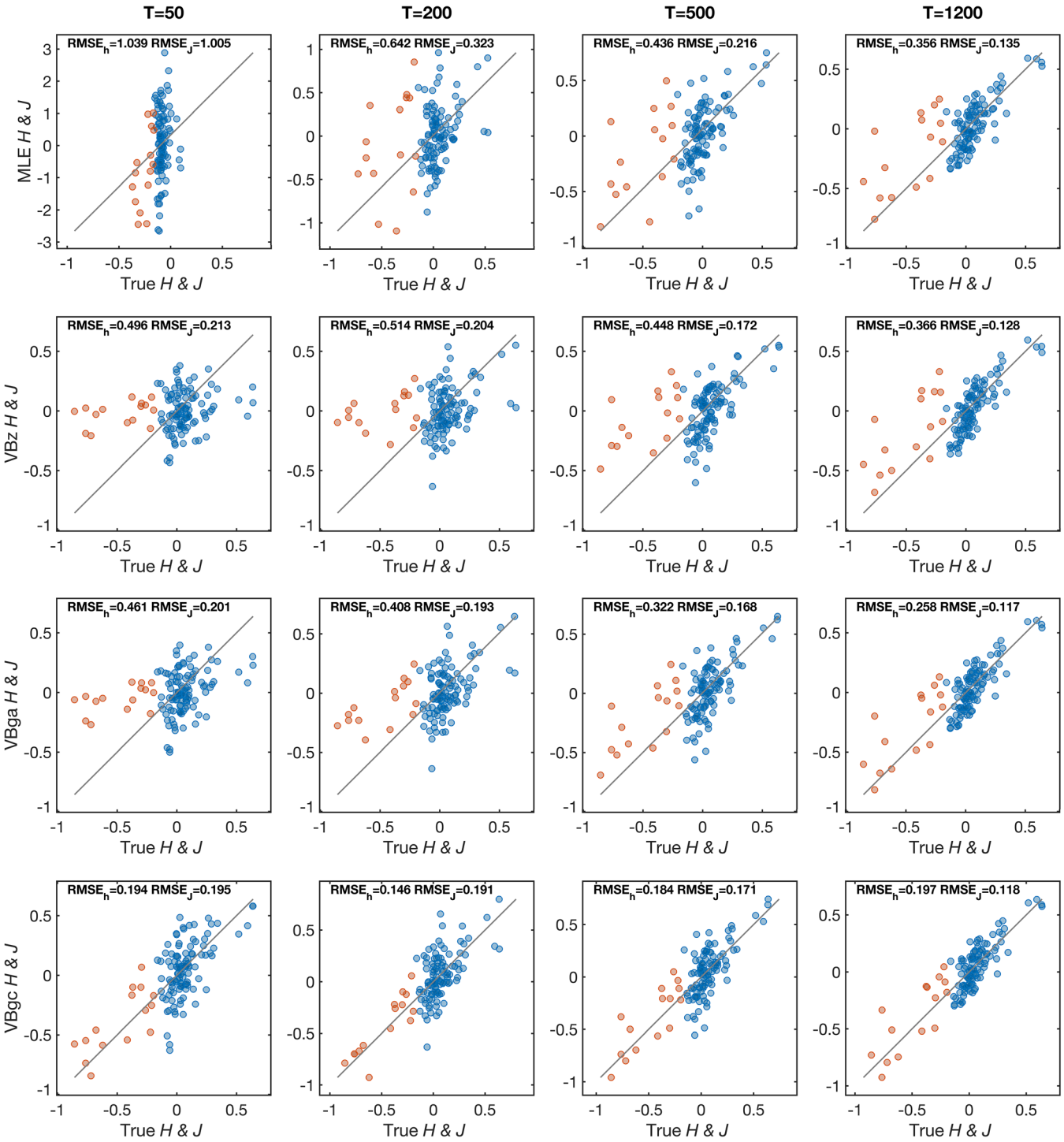

FIGURE 3 Evaluation of MEM parameter estimation from a simulated dataset of a single subject with a 15-ROIs system. The results for model parameter estimation with MLE and three different model priors for variational Bayes (VB) approach are displayed: zero prior expectations (VBz), priors from group average (VBga), and priors from group-concatenated data (VBgc). Each model estimation method is compared with the true parameter that generated state sequences. The red circles indicate *h* while blue circles indicate *J*. The *x*-axis shows the true parameters for a given model, while the *y*-axis indicates the estimated parameters. RMSE for *h* and *J* are displayed in each scatter plot (RMSEh and RMSEJ)

at the group local minima showed significantly reduced correlation with inter-individual variation of the functional connectivity compared to inter-individual variation of energy levels at all states (r = .67 vs. r = .86, p = .000) and MEM parameters (r = .67 vs. .88, p = .000). No significant difference exists between the inter-individual variation of MEM parameters and inter-individual variation of energy levels at all states in the correlation with the inter-individual variation of functional connectivity (p = .104). This result implies that MEM and energy

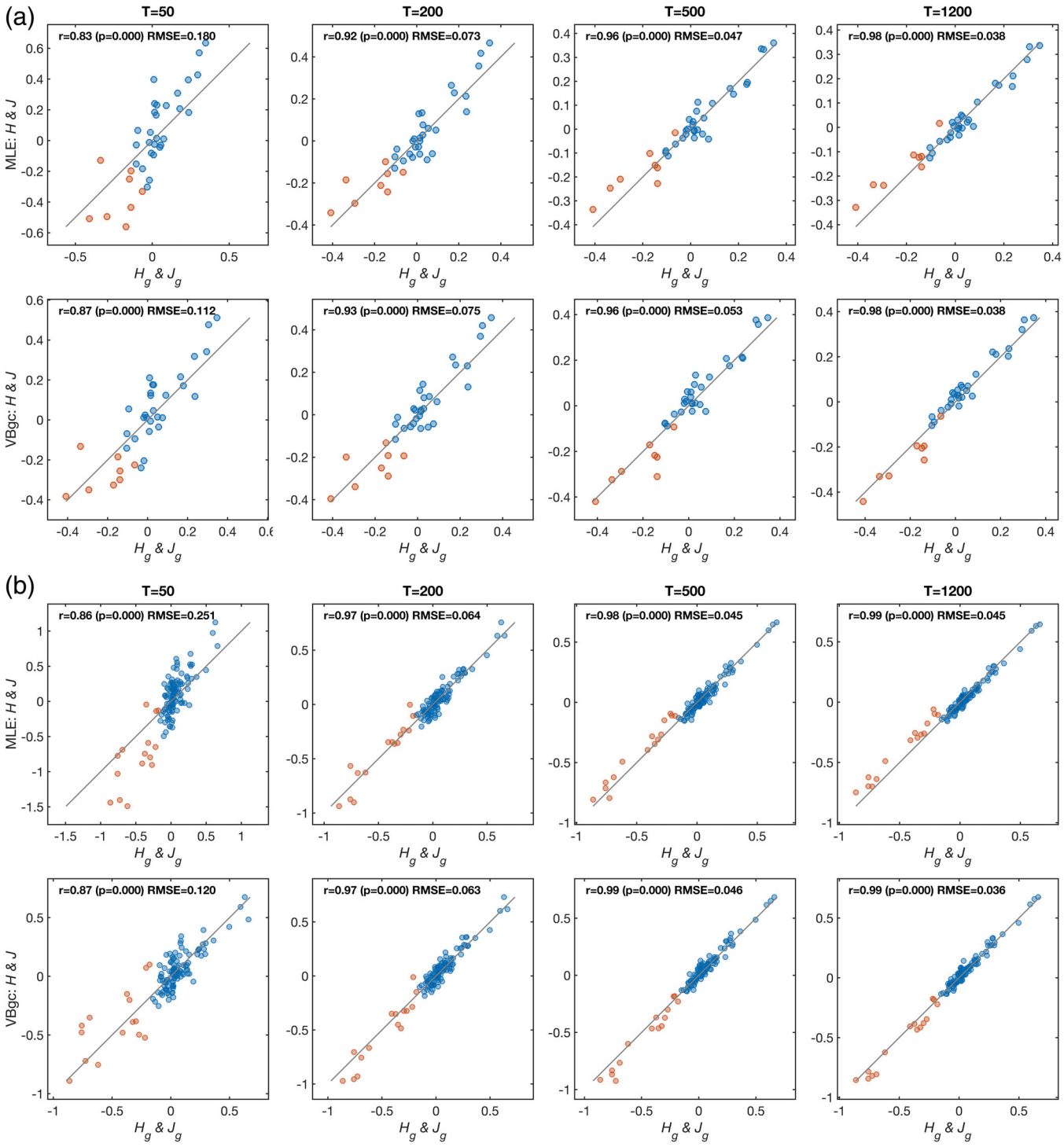

FIGURE 4 Comparison between the average of 30 subjects' estimated parameters versus the group ground-truth parameter for a network size of 8 (a) and of 15 (b). For MLE and VBgc, estimated model parameters for all 30 individuals were averaged for four different sample sizes (T = 50, 200, 500, and 1,200) and compared with those of the group ground-truth. The red colors indicate h while blue colors indicate J. r and p indicate correlation coefficient and its p-value. RMSE: root mean square error

landscape features (e.g., local minima) reflect inter-individual differences in a biologically meaningful way comparable to functional connectivity.

The applicability of BMEM and the energy landscape analysis in characterizing individuals or in identifying neural correlates for individual behaviors were evaluated using sCCA between cognitive scores and neurobiological features estimated from BMEM and energy landscape analysis. Figure 7 shows sCCA results between cognitive scores and MEM parameters. The first canonical variable for cognition scores composed of high fluid intelligence scores

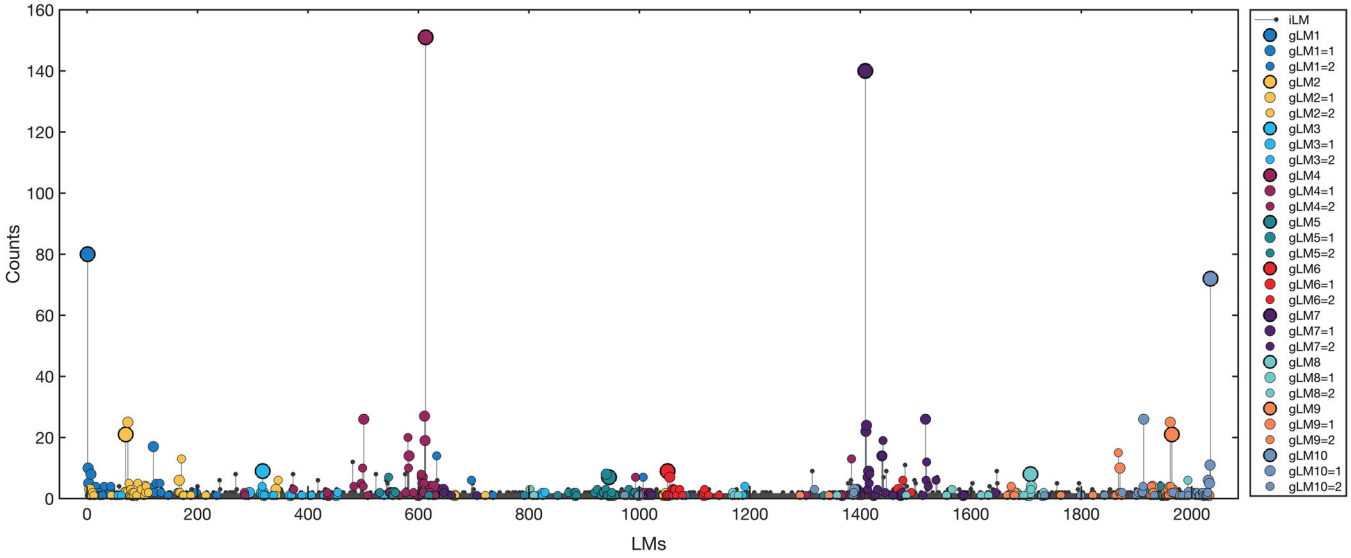

FIGURE 5 Local minima in the energy landscape of BMEM for the group-concatenated data (gLM) are presented with frequencies for all individual local minima found in all the subjects (iLM). According to each state's identification number, the local minima of individuals is presented in sorted order (*x*-axis). The frequencies are evaluated from the group union of all local minima for all individuals (iLMs). For an x-th gLM (i.e., gLMx), gLMx = 1 and gLMx = 2 indicate individual local minima with a Hamming distance = 1 and 2 from the gLMx. Most gLMs serve as references for clusters of iLMs. Some frequent states at the individual level (small black dots, iLM) are not seen as local minima in the group concatenated data (gLMs)

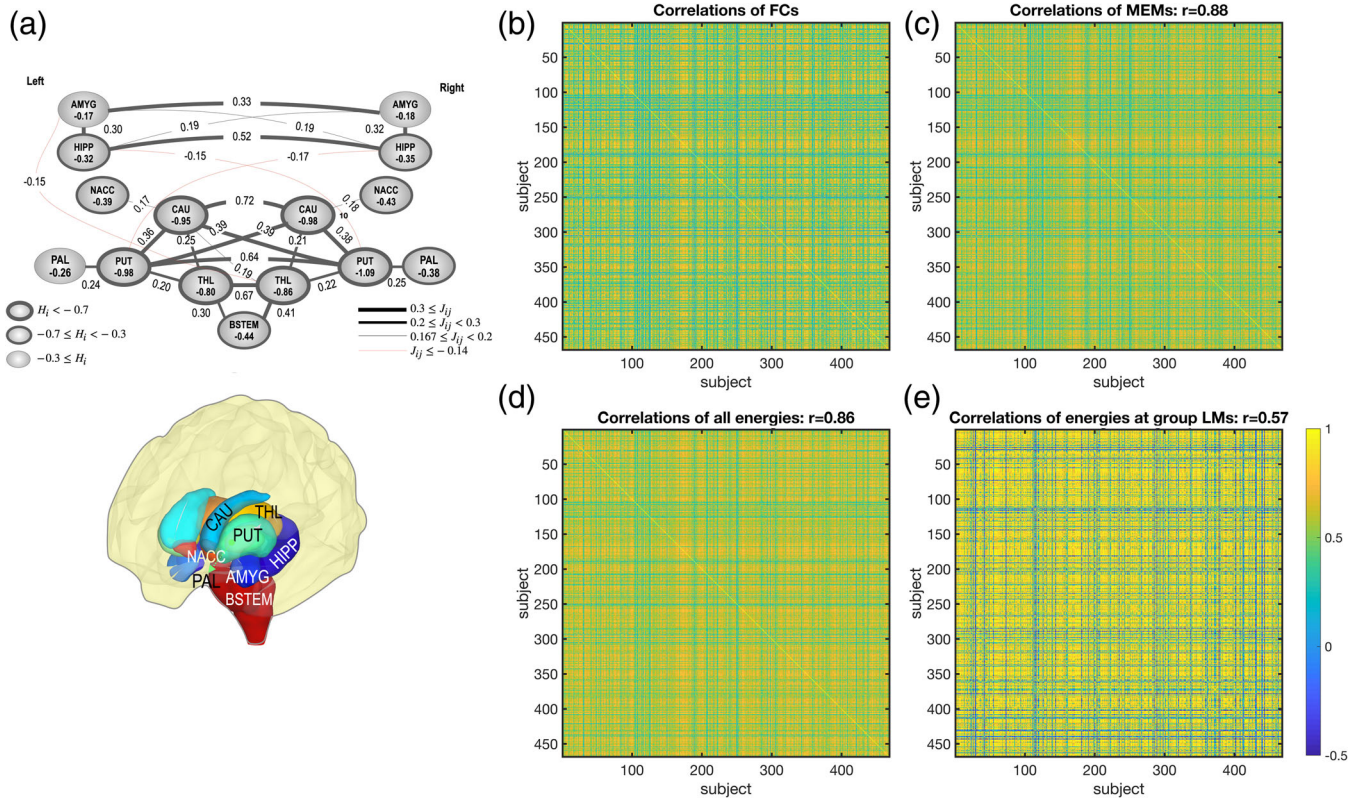

FIGURE 6 Group average interaction parameters (estimated from group-concatenated data) and regions of interests (ROI) (a). Inter-subject similarity matrices of functional connectivity (b), MEM parameters (c), energy landscapes (energies at all the states) (d), and energy levels at the group local minima (e). The adjacency matrices were evaluated by cross-correlation of the multivariate vectors (e.g., functional connectivity, MEM parameters, etc.) across 468 individuals. Part label (a) was modified from Kang et al. (2017). The 15 ROIs were the hippocampus (HIPP), amygdala (AMYG), caudate (CAU), putamen (PUT), pallidum (PAL), thalamus (THL), nucleus accumbens (NACC) of the left and right hemispheres, and brainstem (BSTEM)

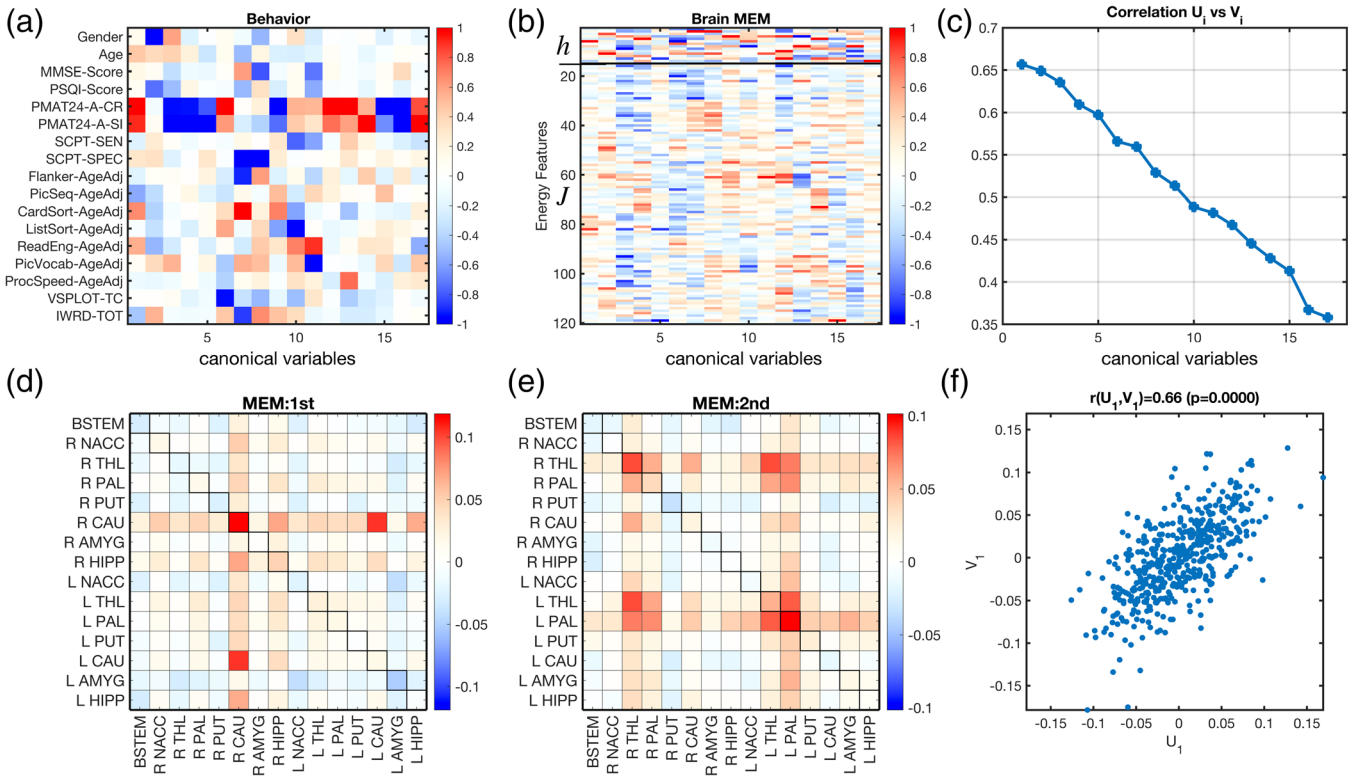

FIGURE 7 sCCA results between cognitive scores and maximum entropy model (MEM) parameters (*h* and *J*). (a) Canonical variables for cognitive scores (U, with gender and age), (b) canonical variables for MEM parameters (V, *h*: upper part, *J*: lower part), and (c) canonical correlations between every pairs of canonical variables for cognitive scores (each column in [a]) and MEM parameters (the column in [b], corresponding to the column in [a]) are displayed. (d) The first MEM canonical variable and (e) the second MEM canonical variable estimated are composed of diagonal elements of *h* and off-diagonal elements of *J*, restructured from the first and second columns in (b). (f) The individual samples of cognitive scores and MEM parameters projected onto the first canonical variable for cognition score (U1) and the first canonical variable for MEM features (V1) showed a correlation of $r = .66$. In (a) and (b), each row was scaled to its maximum value for a better display purpose. In (d and e), the *h* of the MEM (diagonal elements) was divided by 2 to balance the scales of *h* and off-diagonal elements *J*

(PMAT24_A_CR and PMAT24_A_SI) and high executive function/cognitive flexibility score (CardSort-AgeAdj) showed a strong correlation with MEM parameter scores, particularly for connectivity associated with the right caudate (first canonical variable for neurobiological features). The highest canonical correlation between cognitive scores and MEM parameters were $r = .66$, $p = .0000$, $df = 466$. The highest canonical correlation between cognitive scores and the functional connectivity were $r = .64$, $p = .0000$ for full data or $r = .62$, $p = .0000$ after excluding one outlier, as shown in Figure 8. Figure 9 shows sCCA results between cognitions and energy landscape properties of brain states. The highest canonical correlation between the first pair of canonical variables was $r = .46$ ($p = .0000$). Energy landscape features showed significantly reduced canonical correlation with cognitive scores compared to functional connectivity ($r = .46$ vs. $r = .64$, $p = .000$) and MEM parameters ($r = .46$ vs. $.66$, $p = .000$). No significant difference exists between MEM parameters and functional connectivity in the canonical correlation with cognitive scores ($p = .348$). In this evaluation, the first canonical variable for cognition scores (including high fluid intelligence scores) was positively associated with the basin sizes and negatively associated with the occupation times of the first and second lowest local minima of each individual.

The computation time for estimating MEM for a single data of HCP rsfMRI 1,200 scans are approximately 250.1 s for MLE and 7.3 s for BMEM using MATLAB code (Mathworks Inc., Natick, Massachusetts) on Intel(R) Xeon(R) 3.50GHz. BMEM is very fast as BMEM does not require iterations.

4 | DISCUSSION

In the current study, we propose a variational Bayesian scheme for estimating pairwise MEM parameters to analyze brain states' energy landscape using a small sample dataset obtained from a single individual in combination with empirical priors. The energy landscape of brain states, derived from MEM, has provided an important perspective to brain science in understanding state dynamics in terms of a complex brain system (Ezaki et al., 2018; Gu et al., 2018; Kang et al., 2017; Watanabe et al., 2013; Watanabe, Hirose, et al., 2014; Watanabe, Kan, et al., 2014; Watanabe, Masuda, et al., 2014; Park and Kang, 2021).

Note that this nonlinear state dynamics in the current study differs from dynamic functional connectivity researched in recent studies

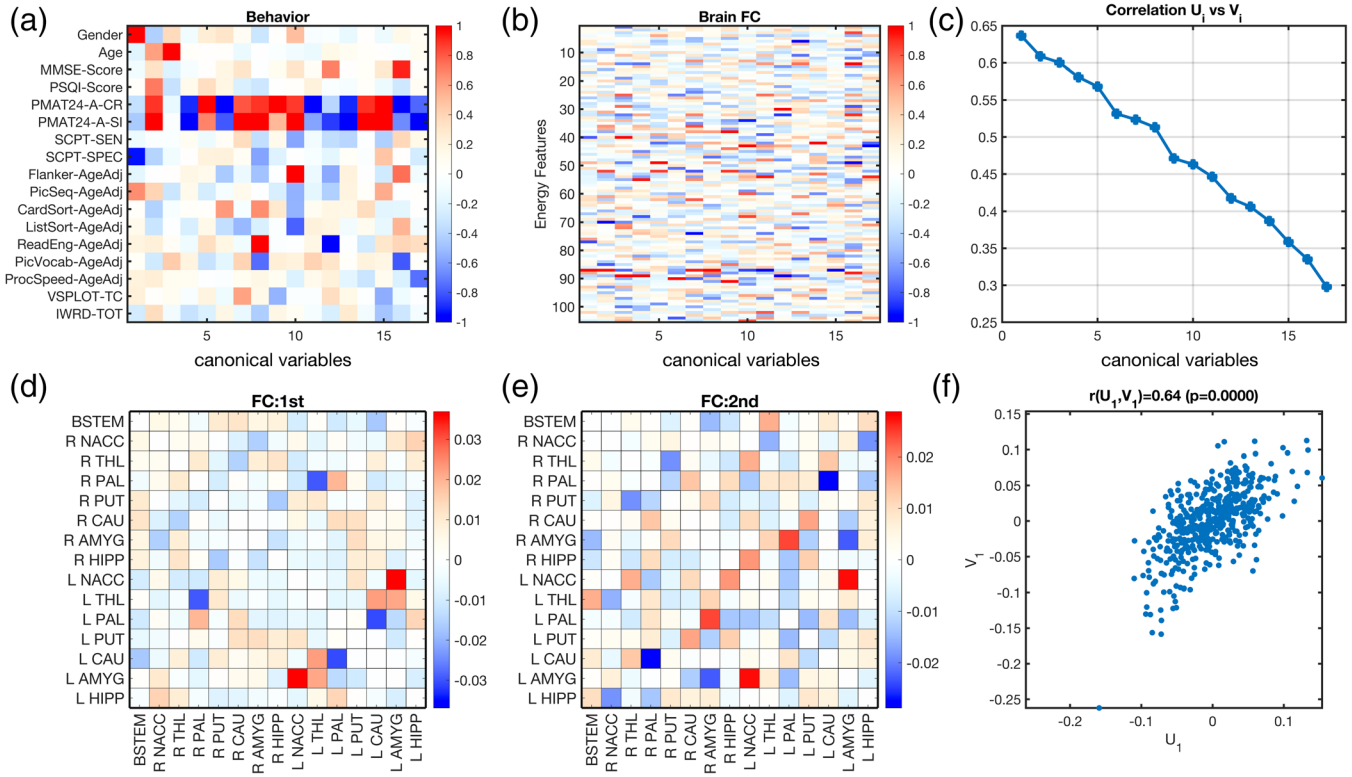

FIGURE 8 sCCA results between cognitive scores and functional connectivity matrix (FC). (a) Canonical variables for cognitive scores (with gender and age), (b) canonical variables for FC, and (c) canonical correlations between every pairs of canonical variables for cognitive scores and FC are displayed. (d) The first FC canonical variable and (e) the second FC canonical variable are shown. (f) The individual samples of cognitive scores and FC projected on to the first canonical variable for cognition scores (U1) and the first FC canonical variable (V1) have a canonical correlation of 0.64 between the two canonical variables. In (a) and (b), for a better display purpose, each row was scaled to its maximum value

(Allen et al., 2014; Calhoun, Miller, Pearlson, & Adali, 2014; Chang & Glover, 2010; Cribben, Haraldsdottir, Atlas, Wager, & Lindquist, 2012; Handwerker, Roopchansingh, Gonzalez-Castillo, & Bandettini, 2012; Hutchison, Womelsdorf, Gati, Everling, & Menon, 2013; Jeong, Pae, & Park, 2016; Monti et al., 2014; H.-J. Park, Friston, Pae, Park, & Razi, 2017; Preti, Bolton, & Van De Ville, 2017), which explains the brain's dynamics in terms of temporal connectivity changes. In contrast, state dynamics in the current study is considered an emergent property of nonlinear interactions among nodes in the stabilized system in equilibrium without any changes in their interaction.

Using energy landscape analysis, we can explore the multistability and state transitional properties of the human brain. For example, in Kang et al. (2017), energy landscape analysis revealed that the subcortical brain system at rest exhibits the maximal number of stable states and small sets of stable states show most of the occupation times. They also explored perturbation effects on the energy landscape at hub nodes and at edges interconnecting nodes with relatively higher node strength. Beyond the multistability of the resting-state brain system, the architecture of state transition processes was analyzed by applying a graph-theoretic analysis to the energy landscape (Kang et al., 2019), which revealed a hub-like state transition organization embedded in the resting-state human brain. These pieces of knowledge were obtained from the group-level fixed-effect analysis by concatenating individual data together. However, the application of state transition analysis to individuals is limited in most studies with small sample data sets. The clinical acquisition setting of rsfMRI data, for which 200 $\sim$ 500 samples are typically used, may not be sufficient to estimate a model of a network size of 10 (55 parameters). Moreover, in the MEM estimation, the temporal dependency (redundancy) of rsfMRI samples due to prolonged hemodynamic responses adds the need for more data to fit the model reliably. To resolve this practical limitation, MEM model fitting has been done in two directions: to increase the size of samples by concatenating individual data within a group (by sacrificing individual variability) and/or to decrease the number of parameters by reducing the number of brain regions of interest. For example, relatively small brain networks were evaluated in previous studies (Watanabe et al., 2013; Watanabe, Hirose, et al., 2014). Kang et al. (2017) assessed the dynamics of a reduced brain system at each hemisphere (mainly focusing on the left hemisphere) to reduce the number of regions even with a group concatenated data set. This makes energy landscape analysis of nonlinear dynamics impractical to use in clinical studies or individual characterization studies.

In this regard, BMEM is of high value in its applicability to individuals. Compared to conventional MLE estimation, BMEM shows a reliable performance in estimating model parameters for a clinically

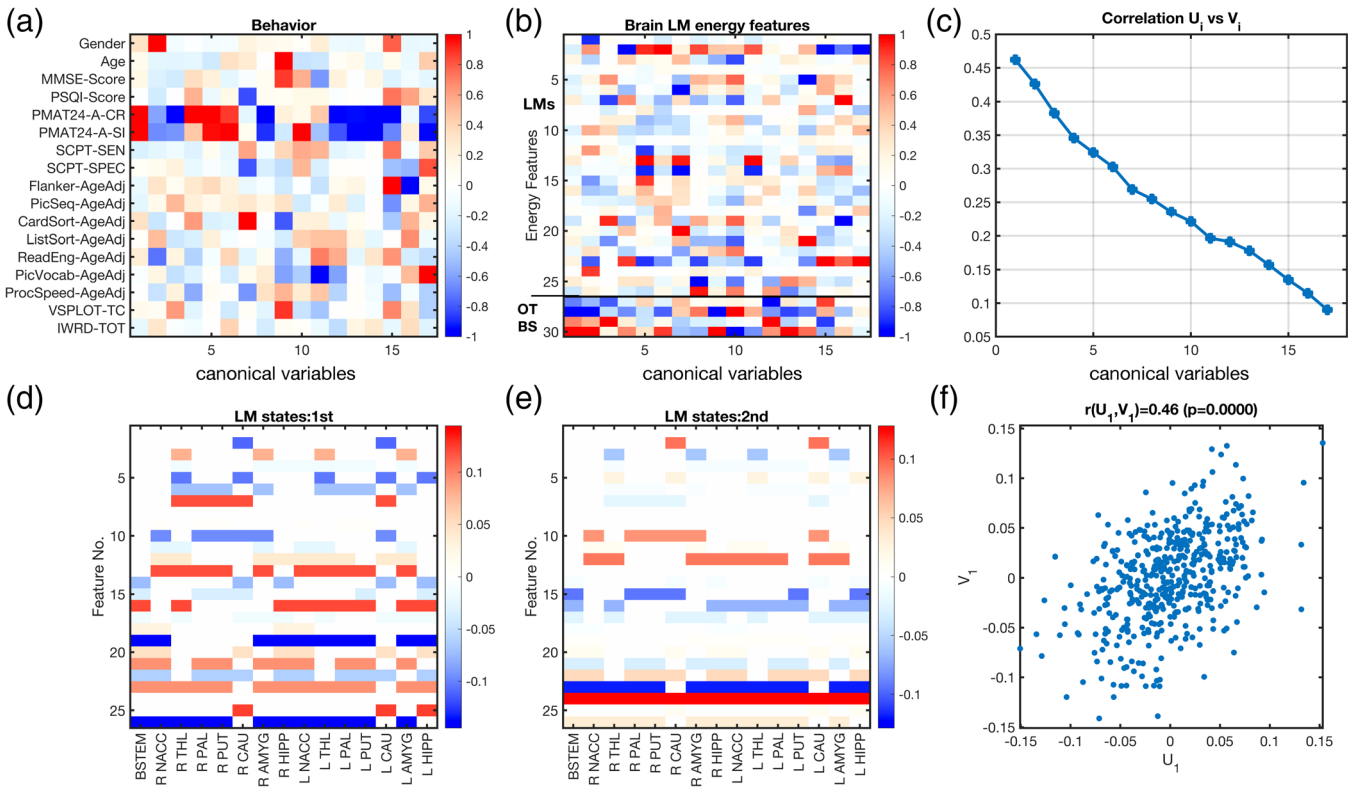

FIGURE 9 sCCA results between cognitive scores and energy landscape features. (a) Canonical variables for cognitive scores (with gender and age) and (b) canonical variables for energy landscape features (upper part: energies at local minima; lower part: the first and second occupation times (OT) and basin sizes (BS) of the lowest local minima for each individual) are displayed. In (c), canonical correlations between every pair of canonical variables for cognitive scores and energy landscape features are shown. In (d) and (e), each row indicates a binarized state vector, weighted by the first and second canonical weights. (f) The individual samples of cognitive scores and energy landscape features projected onto the first canonical variable for cognition scores (U1) and the first canonical variable for energy landscape feature (V1). In (a) and (b), each row was scaled to its maximum value for a better display purpose

achievable data size from an individual due to its capability of utilizing a priori information. When suitable empirical priors were used, the accuracy of BMEM increased to a reasonable level for the small sample data. In particular, empirical priors derived from group-concatenated data improved the parameter estimation of each individual. As a Bayesian method, it is reasonable to expect better performance for better priors that reflect the underlying parameter distribution. In this study, priors with group-concatenated data showed better performance than group average parameters or zero-mean priors. It is speculated that model estimation for a small sample dataset from an individual does not cover all the details of the human brain's energy landscape. Simply averaging model parameters across subjects within a group may not produce a better approximation of priors for individuals than those estimated from a sufficiently large dataset (even if concatenated). The other advantage of utilizing the priors for concatenated data is its simplicity in the calculation compared to using a group average's priors, which requires iterations of model estimation for an individual twice.

It should be noted that BMEM showed empirically desirable statistical properties. As shown in Figure 4, the average of individually estimated model parameters approximated the group ground-truth parameters using the assumption that an individual's model parameters were Gaussian variations from the group ground-truth. If this assumption is valid in real human studies, we may consider the average model parameters across individuals as a reasonable surrogate for the group model. This differs from the conventional way of group model estimation using group-concatenated data. As shown in Figure 5, there exist diverse local minima across individuals, and some of the local minima frequently found in many individuals are not detected as local minima in the model of group concatenated data. Although the group model with the group-concatenated data grossly reflects the group population's brain states, group local minima are limited in representing details of subject-specific local minima. This may be attributable to a nonlinear relationship between a set of model parameters and the local minima in the energy landscape. In summary, the energy landscape analysis may well be done with individual models rather than with a group-level model.

BMEM and its energy landscape show subject specificity in characterizing individuals comparable to subject specificity of functional connectivity. The subject specificity of MEM parameters and their energy landscape features were highly comparable to that of the same data set's functional connectivity. If two individuals have similar functional connectivity, they will have similar MEM parameters and energy landscapes. Although the similarity between the inter-individual

similarity of functional connectivity and that of the energy landscape features is high, energy landscape features also contain subject-specific information different from the functional connectivity in characterizing each individual. This is evident in the correlation analysis between the similarity matrix of the functional connectivity and the similarity matrix of the energy levels at the group model's local minima. The inter-individual similarity of energy levels at the whole states (which contain zero-mean states) are highly correlated with the inter-individual similarity of functional connectivity (r = .86). However, if we look at the details, energies at the group model's local minima show divergence (relatively lower correlation) from functional connectivity across individuals (r = .57). In line with the discussion above regarding Figure 5, this lower correlation indicates that the local minima information of individuals can be lost when we construct a group model using concatenated data across individuals.

sCCA analysis between the cognitive scores and neurobiological features (from fMRI brain signals) in the HCP database suggests the effectiveness of BMEM in identifying neural correlates of cognitive scores and thus in characterizing individuals. We used three sets of neurobiological features from individuals: (a) MEM parameters; (b) energy landscape features; and (c) functional connectivity matrices. Although these features are estimated from the same rsfMRI data set, the implication of each feature differs. Functional connectivity matrices have been the most common neurobiological markers in many brain studies, including analysis of brain diseases using functional connectivity (Drysdale et al., 2017; Jang, Park, Chang, Pae, & Chang, 2016; Lee et al., 2017; Yahata et al., 2016), exploration of neural correlates of behaviors (Kyeong, Kim, Park, & Hwang, 2014; Smith et al., 2015), identification of individuals (Finn et al., 2015), analysis of inter-individual variability (Jang et al., 2017), and features in machine learning (Dosenbach et al., 2010; Drysdale et al., 2017; Finn et al., 2015; Yahata et al., 2016). The correlation-based functional connectivity matrix reflects the brain as a stable, linearly interacting system among the brain regions. In contrast, exploration of the brain using MEM and energy landscape features is based on the brain state's nonlinear dynamics. MEM parameters indicate nonlinear interactions among the brain regions that generate nonlinear state transitions. Using sCCA, we found that the neurobiological features using MEM parameters deliver information in explaining human cognitions and thus in characterizing individuals (Figures 7 and 8).

The energy landscape features (such as energies at the local minima, their occupation times, and basin sizes) are extracted from the MEM parameters, which indicate interactions among brain regions or self-excitability. Meanwhile, the energy landscape features directly reflect information about frequent brain states and their properties in the dynamics, which cannot be explored from the interaction parameters in FC or MEM. Our preliminary study with sCCA suggests the potential of the energy landscape features in explaining human cognitions (r = .46), if not as high as MEM parameters (Figure 9). For example, the basin sizes of the two lowest local minima for each individual are highly associated with the first canonical variable for cognition scores composed of high fluid intelligence scores (Figure 9). These results indicate the importance of nonlinear dynamic properties of brain states, and BMEM provides a critical method to explore those properties at the individual level. It should be noted that the relationship between MEM parameters and energy landscapes (probability distribution) is nonlinear but is uniquely determined. Thus, the correlation with inter-individual similarity matrix of functional connectivity does not show a statistical difference between them. However, the relationship between MEM parameters and energy landscape features is highly complex and shows different aspects of the human brain dynamics; the former focus on the amount of interaction the latter the state dynamics.

We evaluated nonlinear dynamics at the subcortical and (part of) limbic subsystem as an example to present a method for estimating model parameters for a given subsystem. We chose this subcortical system based on the previous study to show that MEM parameters and energy landscapes can be used as biological markers in representing individual characteristics, characterizing individual or biological subgroups in a meaningful way. The characteristics of nonlinear dynamics would differ according to different ROIs. Therefore, one needs a hypothesis or question of interest when choosing a set of ROIs, either the whole brain system or some specific subnetworks such as DMN.

The validity of the proposed approach could be heavily dependent upon the relevance of the prior, mainly when the length of the time series is not long enough. If the length of time series is short, the posterior mean vector is forced to shrink to the mean vector of the prior distribution. Hence, if the hyperparameters of the prior are not good, the resulting posterior for an individual cannot give reasonable estimates. To circumvent this issue, we considered a group averaging and a concatenating strategy to determine the mean vector of the prior distribution, and used an arbitrarily small positive number for the prior precision, say 6.67 as done in our experiments, not to make the resulting posterior mean vector shrink to the prior mean vector too much. A further study for any novel choice of the hyperparameters is left for future research. Note that if the group size (number of subjects) is high, we expect a variation of the group prior may be slight. In this case, the group prior may not be highly dependent on how we choose a group.

In the Bayesian formulation of the model parameter estimation, one may question not using the conjugate prior as the Boltzmann distribution is an exponential family and has a conjugate prior. We presented a method of using conjugate prior in the MEM parameter estimation at the Appendix C. As explained in Appendix C, the posterior of an individual's model parameter is determined too much by the group data, and the individual's data hardly contributes to the calculation of the individual's posterior. Thus, we decided to use a normal prior, which enables us to regularize the effect of the group prior directly by controlling the prior precision.

The current study used the HCP rsfMRI database since it contains well-documented cognitive performance scores and high-quality rsfMRI data from a large population. The database is advantageous to test individual variations and the use of individualized MEM parameters and energies in representing individual variations of cognitive performance before application to clinical data. One drawback of using

rsfMRI of the HCP database is that it does not reflect the conventional clinical setting. The sample size of HCP rsfMRI (1,200 sample points per session with a short TR = 0.72) is generally large compared to the conventional clinical setting. However, the fMRI sample acquisition duration (14.4 min per session) is not sufficient for traditional MEM estimation for a large network that requires independent samples due to the prolonged hemodynamics response (peaks at around 6 s). As shown in Appendix B, the time lags of 6 ∼ 8 s to avoid the temporal dependence results in much reduced independent sample numbers. As we confirmed the usefulness of MEM-based features in characterizing individual differences in cognition, we expect that the current method can be used to analyze data samples acquired from the conventional environment in identifying brains and characterizing brain diseases.

In the current study, we have primarily focused on explaining BMEM and an evaluation of its basic performance. Thus, the detailed characterization of individuals using various energy landscape properties and their behavioral implications remain to be further researched. In particular, dynamic properties in brain diseases according to symptom levels could be studied using BMEM, which was not possible in previous energy landscape analyses of autism spectrum disorders (Watanabe & Rees, 2017) and schizophrenia (Joana Cabral et al., 2013; Loh et al., 2007). We expect that more studies using basic and clinical data will validate the usefulness of the proposed approach and expedite our understanding of the brain states' dynamics.

ACKNOWLEDGMENT

This research was supported by the Brain Research Program and the Brain Pool Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017H1D3A1A01053094, and NRF-2017M3C7A1030750). Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centres that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Centre for Systems Neuroscience at Washington University.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

Data availability All fMRI data analyzed in the present study were obtained from the Human Connectome Project database, which is publicly accessible at: http://www.humanconnectome.org.

ORCID

Jiyoung Kang https://orcid.org/0000-0003-2893-1512
Seok-Oh Jeong https://orcid.org/0000-0002-7073-7599
Chongwon Pae https://orcid.org/0000-0002-9908-3437
Hae-Jeong Park https://orcid.org/0000-0002-4633-0756

REFERENCES

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cerebral Cortex*, *24*(3), 663–676. https://doi.org/10.1093/cercor/bhs352

Bishop, C. M. (2006). Chapter 10. Approximate inference. In Michael Jordan, Jon Kleinberg, & B. Scho¨lkopf (Eds.), *Pattern recognition and machine learning (information science and statistics)*: New York, NY: Springer-Verlag.

Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nature Neuroscience*, *20*(3), 340–352. https://doi.org/10.1038/nn.4497

Cabessa, J., & Villa, A. E. P. (2014). An attractor-based complexity measurement for Boolean recurrent neural networks. *PLoS One*, *9*(4), e94204. https://doi.org/10.1371/journal.pone.0094204

Cabral, J., Fernandes, H. M., Van Hartevelt, T. J., James, A. C., Kringelbach, M. L., & Deco, G. (2013). Structural connectivity in schizophrenia and its impact on the dynamics of spontaneous functional networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *23*(4), 046111. https://doi.org/10.1063/1.4851117

Cabral, J., Kringelbach, M. L., & Deco, G. (2014). Exploring the network dynamics underlying brain activity during rest. *Progress in Neurobiology*, *114*, 102–131. https://doi.org/10.1016/j.pneurobio.2013.12.005

Cabral, J., Kringelbach, M. L., & Deco, G. (2017). Functional connectivity dynamically evolves on multiple time-scales over a static structural connectome: Models and mechanisms. *NeuroImage*, *160*, 84–96. https://doi.org/10.1016/j.neuroimage.2017.03.045

Calhoun, V. D., Miller, R., Pearlson, G., & Adali, T. (2014). The chronnectome: Time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron*, *84*(2), 262–274. https://doi.org/10.1016/j.neuron.2014.10.015

Chang, C., & Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *NeuroImage*, *50*(1), 81–98. https://doi.org/10.1016/j.neuroimage.2009.12.011

Chu, D., Liao, L. Z., Ng, M. K., & Zhang, X. (2013). Sparse canonical correlation analysis: New formulation and algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(12), 3050–3065. https://doi.org/10.1109/TPAMI.2013.104

Cribben, I., Haraldsdottir, R., Atlas, L. Y., Wager, T. D., & Lindquist, M. A. (2012). Dynamic connectivity regression: Determining state-related changes in brain connectivity. *NeuroImage*, *61*(4), 907–920. https://doi.org/10.1016/j.neuroimage.2012.03.070

Deco, G., & Jirsa, V. K. (2012). Ongoing cortical activity at rest: Criticality, multistability, and ghost attractors. *The Journal of Neuroscience*, *32*(10), 3366–3375. https://doi.org/10.1523/JNEUROSCI.2523-11.2012

Deco, G., Jirsa, V. K., & McIntosh, A. R. (2013). Resting brains never rest: Computational insights into potential cognitive architectures. *Trends in Neurosciences*, *36*(5), 268–274. https://doi.org/10.1016/j.tins.2013.03.001

Deco, G., Senden, M., & Jirsa, V. (2012). How anatomy shapes dynamics: A semi-analytical study of the brain at rest by a simple spin model. *Frontiers in Computational Neuroscience*, *6*, 68. https://doi.org/10.3389/fncom.2012.00068

Deco, G., Tononi, G., Boly, M., & Kringelbach, M. L. (2015). Rethinking segregation and integration: Contributions of whole-brain modelling. *Nature Reviews. Neuroscience*, *16*(7), 430–439. https://doi.org/10.1038/nrn3963

Diaconis, P., & Ylvisaker, D. (1979). Conjugate priors for exponential families. *Annals of Statistics*, *7*, 269–281.

Dosenbach, N. U., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., … Schlaggar, B. L. (2010). Prediction of individual brain maturity using fMRI. *Science*, *329*(5997), 1358–1361. https://doi.org/10.1126/science.1194144

Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., … Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, *23* (1), 28–38. https://doi.org/10.1038/nm.4246

Ezaki, T., Fonseca Dos Reis, E., Watanabe, T., Sakaki, M., & Masuda, N. (2020). Closer to critical resting-state neural dynamics in individuals with higher fluid intelligence. *Communications Biology*, *3*(1), 52. https://doi.org/10.1038/s42003-020-0774-y

Ezaki, T., Sakaki, M., Watanabe, T., & Masuda, N. (2018). Age-related changes in the ease of dynamical transitions in human brain activity. *Human Brain Mapping*, *39*(6), 2673–2688. https://doi.org/10.1002/hbm.24033

Ezaki, T., Watanabe, T., Ohzeki, M., & Masuda, N. (2017). Energy landscape analysis of neuroimaging data. *Philosophical Transaction A Math Phys Eng Sci*, *375*(2096), 20160287. https://doi.org/10.1098/rsta.2016.0287

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., … Constable, R. T. (2015). Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, *18*(11), 1664–1671. https://doi.org/10.1038/nn.4135

Freyer, F., Roberts, J. A., Becker, R., Robinson, P. A., Ritter, P., & Breakspear, M. (2011). Biophysical mechanisms of multistability in resting-state cortical rhythms. *The Journal of Neuroscience*, *31*(17), 6353–6361. https://doi.org/10.1523/JNEUROSCI.6693-10.2011

Freyer, F., Roberts, J. A., Ritter, P., & Breakspear, M. (2012). A canonical model of multistability and scale-invariance in biological systems. *PLoS Computational Biology*, *8*(8), e1002634. https://doi.org/10.1371/journal.pcbi.1002634

Golos, M., Jirsa, V., & Dauce, E. (2015). Multistability in large scale models of brain activity. *PLoS Computational Biology*, *11*(12), e1004644. https://doi.org/10.1371/journal.pcbi.1004644

Gu, S., Cieslak, M., Baird, B., Muldoon, S. F., Grafton, S. T., Pasqualetti, F., & Bassett, D. S. (2018). The energy landscape of neurophysiological activity implicit in brain network structure. *Scientific Reports*, *8*(1), 2507. https://doi.org/10.1038/s41598-018-20123-8

Handwerker, D. A., Roopchansingh, V., Gonzalez-Castillo, J., & Bandettini, P. A. (2012). Periodic changes in fMRI connectivity. *NeuroImage*, *63*(3), 1712–1719. https://doi.org/10.1016/j.neuroimage.2012.06.078

Harrell, F. E., Jr., Lee, K. L., Califf, R. M., Pryor, D. B., & Rosati, R. A. (1984). Regression modelling strategies for improved prognostic prediction. *Statistics in Medicine*, *3*(2), 143–152.

Harrell, F. E., Jr., Lee, K. L., & Mark, D. B. (1996). Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, *15*(4), 361–387. https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4

Hutchison, R. M., Womelsdorf, T., Gati, J. S., Everling, S., & Menon, R. S. (2013). Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Human Brain Mapping*, *34*(9), 2154–2177. https://doi.org/10.1002/hbm.22058

Jang, C., Knight, E. Q., Pae, C., Park, B., Yoon, S. A., & Park, H. J. (2017). Individuality manifests in the dynamic reconfiguration of large-scale brain networks during movie viewing. *Scientific Reports*, *7*, 41414. https://doi.org/10.1038/srep41414

Jang, C., Park, H. J., Chang, W. S., Pae, C., & Chang, J. W. (2016). Immediate and longitudinal alterations of functional networks after Thalamotomy in essential tremor. *Frontiers in Neurology*, *7*, 184. https://doi.org/10.3389/fneur.2016.00184

Jeong, S. O., Pae, C., & Park, H. J. (2016). Connectivity-based change point detection for large-size functional networks. *NeuroImage*, *143*, 353–363. https://doi.org/10.1016/j.neuroimage.2016.09.019

Kang, J., Pae, C., & Park, H. J. (2017). Energy landscape analysis of the subcortical brain network unravels system properties beneath resting state dynamics. *NeuroImage*, *149*, 153–164. https://doi.org/10.1016/j.neuroimage.2017.01.075

Kang, J., Pae, C., & Park, H. J. (2019). Graph-theoretical analysis for energy landscape reveals the organization of state transitions in the resting-state human cerebral cortex. *PLoS One*, *14*(9), e0222161. https://doi.org/10.1371/journal.pone.0222161

Kelso, J. A. (2012). Multistability and metastability: Understanding dynamic coordination in the brain. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1591), 906–918. https://doi.org/10.1098/rstb.2011.0351

Knierim, J. J., & Zhang, K. (2012). Attractor dynamics of spatially correlated neural activity in the limbic system. *Annual Review of Neuroscience*, *35* (1), 267–285. https://doi.org/10.1146/annurev-neuro-062111-150351

Kringelbach, M. L., McIntosh, A. R., Ritter, P., Jirsa, V. K., & Deco, G. (2015). The rediscovery of slowness: Exploring the timing of cognition. *Trends in Cognitive Sciences*, *19*(10), 616–628. https://doi.org/10.1016/j.tics.2015.07.011

Kyeong, S., Kim, E., Park, H. J., & Hwang, D. U. (2014). Functional network organizations of two contrasting temperament groups in dimensions of novelty seeking and harm avoidance. *Brain Research*, *1575*, 33–44. https://doi.org/10.1016/j.brainres.2014.05.037

Lee, D., Pae, C., Lee, J. D., Park, E. S., Cho, S. R., Um, M. H., … Park, H. J. (2017). Analysis of structure-function network decoupling in the brain systems of spastic diplegic cerebral palsy. *Human Brain Mapping*, *38* (10), 5292–5306. https://doi.org/10.1002/hbm.23738

Loh, M., Rolls, E. T., & Deco, G. (2007). A dynamical systems hypothesis of schizophrenia. *PLoS Computational Biology*, *3*(11), e228.

Monti, R. P., Hellyer, P., Sharp, D., Leech, R., Anagnostopoulos, C., & Montana, G. (2014). Estimating time-varying brain connectivity networks from functional MRI time series. *NeuroImage*, *103*, 427–443. https://doi.org/10.1016/j.neuroimage.2014.07.033

Park, B., Ko, J. H., Lee, J. D., & Park, H. J. (2013). Evaluation of node-inhomogeneity effects on the functional brain network properties using an anatomy-constrained hierarchical brain parcellation. *PLoS One*, *8*(9), e74935. https://doi.org/10.1371/journal.pone.0074935

Park, H.-J., Friston, K. J., Pae, C., Park, B., & Razi, A. (2017). Dynamic effective connectivity in resting state fMRI. *NeuroImage*, *180*, 594–608. https://doi.org/10.1016/j.neuroimage.2017.11.033

Park, H.-J., & Kang, J. (2021). A computational framework for controlling the self-restorative brain based on the free energy and degeneracy principles. *Frontiers in Computational Neuroscience*. https://doi.org/10.3389/fncom.2021.590019.

Preti, M. G., Bolton, T. A., & Van De Ville, D. (2017). The dynamic functional connectome: State-of-the-art and perspectives. *NeuroImage*, *160*, 41–54. https://doi.org/10.1016/j.neuroimage.2016.12.061

Rabinovich, M. I., Huerta, R., Varona, P., & Afraimovich, V. S. (2008). Transient cognitive dynamics, Metastability, and decision making. *PLoS Computational Biology*, *4*(5), e1000072. https://doi.org/10.1371/journal.pcbi.1000072

Rabinovich, M. I., & Varona, P. (2011). Robust transient dynamics and brain functions. *Frontiers in Computational Neuroscience*, *5*, 24. https://doi.org/10.3389/fncom.2011.00024

Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E., Glasser, M. F., … Miller, K. L. (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nature Neuroscience*, *18*(11), 1565–1567. https://doi.org/10.1038/nn.4125

Tognoli, E., & Kelso, J. A. (2014). The metastable brain. *Neuron*, *81*(1), 35–48. https://doi.org/10.1016/j.neuron.2013.12.022

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E., Bucholz, R., … Consortium, W. U.-M. H. (2012). The human connectome project: A data acquisition perspective. *NeuroImage*, *62*(4), 2222–2231. https://doi.org/10.1016/j.neuroimage.2012.02.018

Watanabe, T., Hirose, S., Wada, H., Imai, Y., Machida, T., Shirouzu, I., … Masuda, N. (2013). A pairwise maximum entropy model accurately describes resting-state human brain networks. *Nature Communications*, *4*, 1370. https://doi.org/10.1038/ncomms2388

Watanabe, T., Hirose, S., Wada, H., Imai, Y., Machida, T., Shirouzu, I., … Masuda, N. (2014). Energy landscapes of resting-state brain networks. *Frontiers in Neuroinformatics*, *8*, 12. https://doi.org/10.3389/fninf.2014.00012

Watanabe, T., Kan, S., Koike, T., Misaki, M., Konishi, S., Miyauchi, S., … Masuda, N. (2014). Network-dependent modulation of brain activity during sleep. *NeuroImage*, *98*, 1–10. https://doi.org/10.1016/j.neuroimage.2014.04.079

Watanabe, T., Masuda, N., Megumi, F., Kanai, R., & Rees, G. (2014). Energy landscape and dynamics of brain activity during human bistable perception. *Nature Communications*, *5*, 4765. https://doi.org/10.1038/ncomms5765

Watanabe, T., & Rees, G. (2017). Brain network dynamics in high-functioning individuals with autism. *Nature Communications*. 8. https://doi.org/10.1038/ncomms16048

Yahata, N., Morimoto, J., Hashimoto, R., Lisi, G., Shibata, K., Kawakubo, Y., … Kawato, M. (2016). A small number of abnormal brain connections predicts adult autism spectrum disorder. *Nature Communications*, *7*, 11254. https://doi.org/10.1038/ncomms11254

Yeh, F. C., Tang, A. N., Hobbs, J. P., Hottowy, P., Dabrowski, W., Sher, A., … Beggs, J. M. (2010). Maximum entropy approaches to living neural networks. *Entropy*, *12*(1), 89–106. https://doi.org/10.3390/e12010089

How to cite this article: Kang J, Jeong S-O, Pae C, Park H-J. Bayesian estimation of maximum entropy model for individualized energy landscape analysis of brain state dynamics. *Hum Brain Mapp*. 2021;42:3411–3428. https://doi.org/10.1002/hbm.25442

APPENDIX A. MATHEMATICAL DETAILS OF BMEM FORMULA

This appendix explains the mathematical details of BMEM formulation. The evidence lower bound (ELBO) is defined as follows:

$$\mathcal{F}(q,\boldsymbol{\eta},\boldsymbol{\alpha}) = \int q(\boldsymbol{\theta})\log p\left(\frac{\mathcal{D},\boldsymbol{\theta}|\boldsymbol{\eta},\boldsymbol{\alpha})}{q(\boldsymbol{\theta})}d\boldsymbol{\theta} = T_1 + T_2 + T_3,$$

with $T_1 = \int q(\boldsymbol{\theta})\log p(\mathcal{D}|\boldsymbol{\theta})d\boldsymbol{\theta}$, $T_2 = \int q(\boldsymbol{\theta})\log p(\boldsymbol{\theta}|\boldsymbol{\eta},\boldsymbol{\alpha})d\boldsymbol{\theta}$ and $T_3 = -\int q(\boldsymbol{\theta})\log q(\boldsymbol{\theta})d\boldsymbol{\theta}$. Plugging $\log p(\mathcal{D}|\boldsymbol{\theta}) = \sum_{t=1}^{T}\boldsymbol{\theta}^\top\tilde{\boldsymbol{\sigma}}(t) - T\log\sum_{\boldsymbol{\sigma}\in\mathcal{S}}e^{\boldsymbol{\theta}^\top\tilde{\boldsymbol{\sigma}}}$ into T_1, we have

$$T_1 = \sum_{t=1}^{T}\boldsymbol{\mu}^\top\tilde{\boldsymbol{\sigma}}(t) - T\int q(\boldsymbol{\theta})\log\sum_{\boldsymbol{\sigma}\in\mathcal{S}}e^{\boldsymbol{\theta}^\top\tilde{\boldsymbol{\sigma}}}d\boldsymbol{\theta}.$$

The integration in the above is intractable, so we consider a Taylor series expansion near $\boldsymbol{\eta}$ for its integrand given by

$$\log\sum_{\boldsymbol{\sigma}\in\mathcal{S}}e^{\boldsymbol{\theta}^\top\tilde{\boldsymbol{\sigma}}} \approx \log\sum_{\boldsymbol{\sigma}\in\mathcal{S}}e^{\boldsymbol{\eta}^\top\tilde{\boldsymbol{\sigma}}} + \langle\tilde{\boldsymbol{\sigma}}\rangle_{\boldsymbol{\eta}}^\top(\boldsymbol{\theta}-\boldsymbol{\eta}) + \frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\eta})^\top C_{\boldsymbol{\eta}}(\boldsymbol{\theta}-\boldsymbol{\eta}),$$

where $\langle\tilde{\boldsymbol{\sigma}}\rangle_{\boldsymbol{\eta}}$ and $C_{\boldsymbol{\eta}} = \mathrm{Cov}_{\boldsymbol{\eta}}(\tilde{\boldsymbol{\sigma}})$ are the model mean vector and the variance–covariance matrix of $\tilde{\boldsymbol{\sigma}}$ when the model parameter is $\boldsymbol{\theta} = \boldsymbol{\eta}$, respectively. By plugging this approximation into T_1, we have

$$T_1 \approx \sum_{t=1}^{T}\boldsymbol{\mu}^\top\tilde{\boldsymbol{\sigma}}(t) - T\left[\log\sum_{\boldsymbol{\sigma}\in\mathcal{S}}e^{\boldsymbol{\eta}^\top\tilde{\boldsymbol{\sigma}}} + \langle\tilde{\boldsymbol{\sigma}}\rangle_{\boldsymbol{\eta}}^\top(\boldsymbol{\mu}-\boldsymbol{\eta}) + \frac{1}{2}\left\{\mathrm{tr}\left(\mathrm{diag}(\boldsymbol{\beta})^{-1}C_{\boldsymbol{\eta}}\right) + (\boldsymbol{\mu}-\boldsymbol{\eta})^\top C_{\boldsymbol{\eta}}(\boldsymbol{\mu}-\boldsymbol{\eta})\right\}\right].$$

It is rather simple to compute T_2:

$$T_2 = \frac{1}{2}\sum_{j=1}^{M}\log\alpha_j - \frac{1}{2}\sum_{j=1}^{M}\alpha_j\left\{(\mu_j-\eta_j)^2 + \beta_j^{-1}\right\} + \text{constant}.$$

The term T_3 is the entropy of q, so it holds that

$$T_3 = -\frac{1}{2}\sum_{j=1}^{M}\log\beta_j + \text{constant}.$$

Hence, we have an approximation for ELBO as follows:

$$\mathcal{F}(q,\boldsymbol{\eta},\boldsymbol{\alpha}) \approx \sum_{t=1}^{T}\boldsymbol{\mu}^\top\tilde{\boldsymbol{\sigma}}(t) - T\left[\log\sum\nolimits_{\boldsymbol{\sigma}\in\mathcal{S}}e^{\boldsymbol{\eta}^\top\tilde{\boldsymbol{\sigma}}} + \langle\tilde{\boldsymbol{\sigma}}\rangle_{\boldsymbol{\eta}}^\top(\boldsymbol{\mu}-\boldsymbol{\eta}) + \frac{1}{2}\left\{\mathrm{tr}\left(\mathrm{diag}(\boldsymbol{\beta})^{-1}C_{\boldsymbol{\eta}}\right) + (\boldsymbol{\mu}-\boldsymbol{\eta})^\top C_{\boldsymbol{\eta}}(\boldsymbol{\mu}-\boldsymbol{\eta})\right\}\right] + \frac{1}{2}\sum\nolimits_{j=1}^{M}\log\alpha_j - \frac{1}{2}\sum\nolimits_{j=1}^{M}\alpha_j\left\{(\mu_j-\eta_j)^2 + \beta_j^{-1}\right\} - \frac{1}{2}\sum\nolimits_{j=1}^{M}\log\beta_j + \text{constant}.$$

Now, we are to maximize the above approximate $\mathcal{F}$ with respect to q. For $j = 1, 2, \ldots, M$,

$$\frac{\partial\mathcal{F}}{\partial\mu_j} = T\left(\langle\tilde{\sigma}_j\rangle_{\text{empirical}} - \langle\tilde{\sigma}_j\rangle_{\boldsymbol{\eta}}\right) - T(\mu_j-\eta_j)C_{\boldsymbol{\eta},jj} - T\sum_{i\neq j}(\mu_i-\eta_i)C_{\boldsymbol{\eta},ij} - \alpha_j(\mu_j-\eta_j) = 0,$$

$$\frac{\partial\mathcal{F}}{\partial\beta_j} = \frac{TC_{\boldsymbol{\eta},jj}}{2\beta_j^2} + \frac{\alpha_j}{2\beta_j^2} - \frac{1}{2\beta_j} = 0,$$

where $C_{\boldsymbol{\eta},ij}$ is the (i,j) component of $C_{\boldsymbol{\eta}}$. Solving the above equation system, we have posterior mean $\boldsymbol{\mu}$ and its precision $\boldsymbol{\beta}$:

$$\boldsymbol{\mu} = \boldsymbol{\eta} + T\mathbf{A}_{\boldsymbol{\eta},\boldsymbol{\alpha}}^{-1}\left(\langle\tilde{\boldsymbol{\sigma}}\rangle_{\text{empirical}} - \langle\tilde{\boldsymbol{\sigma}}\rangle_{\boldsymbol{\eta}}\right), \boldsymbol{\beta} = \boldsymbol{\alpha} + T\mathbf{c}_{\boldsymbol{\eta}},$$

with $\mathbf{A}_{\boldsymbol{\eta},\boldsymbol{\alpha}} = \mathrm{diag}(\boldsymbol{\alpha}) + TC_{\boldsymbol{\eta}}$ and $\mathbf{c}_{\boldsymbol{\eta}}$ is the vector composed of the diagonal elements of $C_{\boldsymbol{\eta}}$.

APPENDIX B. TEMPORAL DEPENDENCY OF FMRI SIGNALS

The temporal dependency of fMRI signals was evaluated using autocorrelation analysis. The results for 15 ROIs are shown in Figure B1.

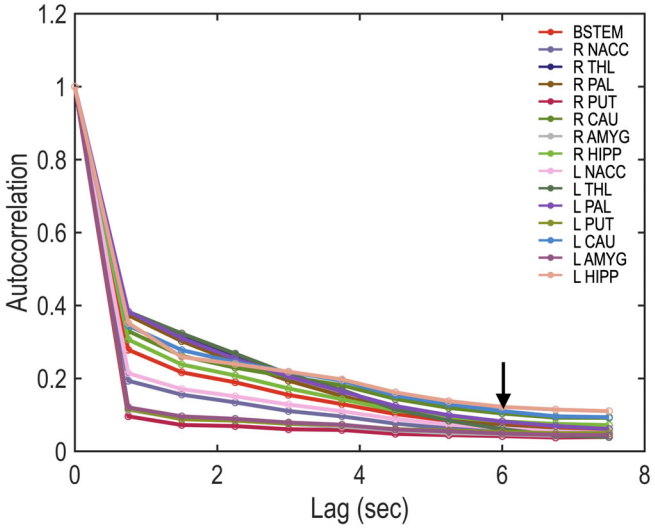

FIGURE B1 Autocorrelation function of fMRI time series with different time lags at 15 ROIs of a subject in the HCP database. The autocorrelation functions show that fMRI samples' temporal dependency is substantially weakened for time lags of 6 ∼ 8 s

APPENDIX C. CONJUGATE PRIOR BASED APPROACH

A Boltzmann distribution belongs to the exponential family, and every member of the exponential family has a conjugate prior, see Diaconis and Ylvisaker (1979). The canonical conjugate prior can be written as below.

$$p(\boldsymbol{\theta}|\,\lambda) \propto \exp\left\{\lambda_1^T\boldsymbol{\theta} + \lambda_2 a(\boldsymbol{\theta})\right\}, \lambda = (\lambda_1, \lambda_2),$$

of which hyperparameter vector is given in a form related to the group size:

$$\lambda_1 = \sum_{g=1}^{G}\sum_{t=1}^{T_g}\tilde{\boldsymbol{\sigma}}_g(t),\ \lambda_2 = \sum_{g=1}^{G} T_g,$$

where G is the number of subjects in the group concatenated data and T_g is the length of time series observed from the g-th subject, $g = 1, 2, \ldots, G$. With this prior, the resulting parameter vector of posterior distribution is given by the sum of the individual data vector and the hyperparameter vector (i.e., the sum of group data vector):

$$p(\boldsymbol{\theta}|\,\mathcal{D},\lambda) \propto \exp\left\{\left(\lambda_1 + T\langle\tilde{\boldsymbol{\sigma}}\rangle_{\text{empirical}}\right)^T\boldsymbol{\theta} + (\lambda_2 + T)a(\boldsymbol{\theta})\right\}.$$

Therefore, especially when the number of subjects (G) in the group concatenated data for the prior is large, λ_1 and λ_2 get very large, and the information from the individual of interest hardly contributes to constructing the posterior. One may select a subset from the database to form a "good" group concatenated data with "proper" size, but this arouses a new problem: which and how many subjects? Instead, we decided to use the normal prior, which enables us to regularize the effect of the prior (group) information in a direct fashion by controlling the prior precision.