Fully Bayesian estimation of virtual brain parameters with self-tuning Hamiltonian Monte Carlo

Jayant Jha, Meysam Hashemi, Anirudh Nihalani Vattikonda, Huifang Wang, Viktor Jirsa

Machine Learning: Science and Technology

Fully Bayesian estimation of virtual brain parameters with self-tuning Hamiltonian Monte Carlo

To cite this article before publication: Jayant Jha *et al* 2022 *Mach. Learn.: Sci. Technol.* in press https://doi.org/10.1088/2632-2153/ac9037

Fully Bayesian estimation of virtual brain parameters with self-tuning Hamiltonian Monte Carlo

Jayant Jha[1*], Meysam Hashemi[1], Anirudh Nihalani Vattikonda[1], Huifang Wang[1], Viktor Jirsa[1†]

Aix Marseille Université,
Institut National de la Santé et de la Recherche Médicale,
Institut de Neurosciences des Systèmes (INS), UMR1106
Marseille 13005, France[1]
E-mails: jayantjha@gmail.com*, viktor.jirsa@univ-amu.fr†

Abstract

Virtual brain models are data-driven patient-specific brain models integrating individual brain imaging data with neural mass modeling in a single computational framework, capable of autonomously generating brain activity and its associated brain imaging signals. Along the example of epilepsy, we develop an efficient and accurate Bayesian methodology estimating the parameters linked to the extent of the epileptogenic zone. State-of-the-art advances in Bayesian inference using Hamiltonian Monte Carlo (HMC) algorithms have remained elusive for large-scale differential-equations based models due to their slow convergence. We propose appropriate priors and a novel reparameterization to facilitate efficient exploration of the posterior distribution in terms of computational time and convergence diagnostics. The methodology is illustrated for in-silico dataset and then, applied to infer the personalized model parameters based on the empirical stereotactic electroencephalography (SEEG) recordings of retrospective patients. This improved methodology may pave the way to render HMC methods sufficiently easy and efficient to use, thus applicable in personalized medicine.

Some keywords: Bayesian inference, Brain-network modelling, Epilepsy, Hamiltonian Monte Carlo, Personalized medicine.

1 Introduction

The activity in the brain for cognition, perception and consciousness are governed by the fundamental mechanism of neural network oscillations. Consequently, perturbations of network activity play an important role in the pathophysiology of brain disorders such as epilepsy [1, 2, 3]. When individual structural information from non-invasive brain imaging is merged with mathematical modelling, then generative brain network models constitute personalized in-silico platforms for the exploration of causal mechanisms of brain dysfunction and clinical hypothesis testing [4]. Such data-driven computational whole-brain models are referred to as virtual brain models in which a patient's structural brain imaging data derived from non-invasive magnetic resonance imaging (MRI) constrain the latent space trajectories of the brain network model. Simple activation paradigms lack the functional complexity to explain the richness of observed spatiotemporal behaviours linked to these brain dynamics [5], leaving it essentially to non-linear network processes to explain the origin of the emergent functional and pathological spatiotemporal patterns. The virtual brain models emphasize the network character of the brain and bring dynamical properties to the structural data of individual brains [6]. In the virtual brain models, the dynamics of a network node is given by a neural mass model of lumped neuronal activity and is connected to other nodes via a connectivity matrix representing white matter fibre tracts derived from diffusion tractography of the brain [7, 1, 8]. This form of connectome-based virtual brain modelling [9, 10] exploits the explanatory power of measured network connectivity imposed as a constraint upon network dynamics and has provided important insights into the mechanisms underlying the emergence of the resting-state networks dynamics [11, 7, 12] of healthy subjects, stroke [13], schizophrenic patients [14] and epilepsy [15]. So far, neural mass models have proven successful in explaining the biophysical and dynamical nature of seizure onsets and offsets [16].

In partial epilepsy, seizures originate in a local network, the so-called epileptogenic zone (EZ), before recruiting other close or distant brain regions, the so-called propagation zone (PZ). Partial seizures have been reported to propagate through large-scale brain network in humans [17, 18, 15] and animals [19]. A possible treatment for the partial epileptic patients is the surgical resection of the EZ, a localized region or network where seizures arise, before recruiting secondary networks, PZ [20, 21, 22]. As a part of the standard

presurgical evaluation, stereotactic electroencephalography (SEEG) recordings are used to help correctly delineate the EZ. [23] suggests the calculation of Epileptogenicity Indices corresponding to each of the regions in the brain based on the SEEG recordings. However, in practice (surgery), the electrodes are placed sparsely with respect to the large and complex structure of the brain. Therefore, such a model-free approach cannot identify the activities in the brain regions which are far from the electrodes. Also, the SEEG recordings are based on the unknown combination of the activities at different hidden sources. This is also not captured in the calculation of Epileptogenicity Indices. Moreover, the seizure in partial epilepsy propagates by activating different regions in the brain due to the connectivity of the EZ(s) with other regions. The consideration of the virtual brain models enables the prediction of the activity in all the brain regions simultaneously based on the measurements at senors, which are placed sparsely, by taking the linear combination of the time-series of different brain regions as the generating mechanism for the observed sensor activity.

We consider a Bayesian model inversion methodology to infer the activity of different brain regions based on the SEEG measurements at the sparsely placed sensors in this paper. Bayesian framework offers powerful and principled method for parameter inference, dealing with uncertainty in estimation, and out-of-sample model prediction from experimental data with a broad range of applications [24, 25, 26, 24, 27]. Within a neuroimaging context, the Bayesian approach has been widely used for inference of neuronal population's intrinsic parameters and/or interactions between neuronal populations (effective connectivity) in a pre-specified neuronal network from neurophysiological data [28, 29, 30, 31, 32, 33]. The sampling from the posterior distribution provides the neuroscientist with an estimation of parameter distributions over plausible ranges of the parameter space consistent with the observed data, which makes the link to the underlying physiological mechanisms employed in the network model and, thus, enable its scientific interrogation, for instance for resting state and task paradigms [30, 29], and clinical translation, for instance in epilepsy [1, 34]. However, in previous studies, either the model considered is very simplistic i.e. linear models [35] or the number of nodes considered in the brain is very small or the data is assumed to be observed at the source level [34]. These considerations were possibly driven by the fact that the Bayesian inference for high-dimensional non-linear models require huge computational effort making them infeasible for application purpose.

The Virtual Epileptic Patient (VEP; [1]) model as a personalized virtual brain model, captures the activity of the brain regions during epilepsy. The measurements at the sensors are considered to be linear combination of the activities generated at the brain regions. The linear combination is determined by a known gain matrix considered in the model. The gain matrix represents the contribution of different source activities at the sensors. The i, j-th element of the matrix represents the inverse squared distance between the i-th sensor and the j-th node. The matrix is not of full rank in practice as the number of sources are generally more than the number of sensors and also, due to the sparse placement of SEEG electrodes. The presence of the gain matrix as well as the non-linear hidden dynamics at brain regional level, with large number of nodes and connectivity between the nodes induce correlations in the posterior distributions of the model parameters making inference difficult. Moreover, the activities of the brain regions during epilepsy are governed by the steady state behaviour of the time-series given by the model equations.

An attempt at inversion for the VEP model was considered in [34] under the assumption that the data are observed at source level for the different brain regions. The method cannot be directly applied to the empirical data as the measurements are taken at different SEEG electrodes and the observed activity at the electrodes are the combination of the activities generated at different brain regions. In our model, we have accounted for this using the gain matrix which maps the activities of different brain regions to the measurements at SEEG electrodes. We propose an efficient method incorporating the steady state behaviour and sparsity in the number of brain regions using the placement of appropriate priors on the model parameters. We also propose a novel reparameterization of the model configuration space based on the dominant linear combinations derived from the gain matrix making the inference procedure computationally feasible.

It is well-known that gradient-free sampling algorithms such as Metropolis-Hastings, Gibbs sampling and slice-sampling [27] generally fail to explore the parameter space efficiently when applied to large-scale inverse problems [36, 37, 38], as often encountered in the application of whole-brain imaging for clinical diagnoses. In particular, these traditional Markov chain Monte Carlo (MCMC) algorithms mix poorly in high-dimensional parameter spaces involving correlated variables [36, 37, 39]. In contrast, gradient-based algorithms such as Hamiltonian Monte Carlo (HMC) [40, 41], although computationally expensive,

they are far superior to gradient-free sampling algorithms in terms of the number of independent samples produced per unit computational time [42, 43]. This class of sampling algorithms provides efficient convergence and exploration of parameter space even in very high-dimensional spaces that may exhibit strong correlations between samples [36, 37, 39]. Nevertheless, the efficiency of gradient-based sampling methods such as HMC is highly sensitive to the user-specified algorithm parameters [36, 37]. More advanced MCMC sampling algorithms such as No-U-Turn Sampler (NUTS; [37]), a self-tuning variant of HMC [44] solve these issues by adaptively tuning the algorithm parameters. It has been shown that these algorithms efficiently sample from high-dimensional target distributions that allow us to solve complex inverse problems conditioned on massive data sets as the observation [45, 46, 47]. MCMC has the advantage of being non-parametric and asymptotically exact in the limit of long/infinite runs [24]. Of the other alternatives, Variational Inference (VI; [48, 49]) turns the Bayesian inference into an optimization problem, which typically results in much faster computation than MCMC methods [24, 47]. However, the classical derivation of VI requires a major model-specific work on defining a variational family appropriate to the probabilistic model, computing the corresponding objective function, computing gradients, and running a gradient-based optimization algorithm [50]. Probabilistic programming languages (PPLs; [44, 51]) provide efficient implementation for automatic Bayesian inference on user-defined probabilistic models by featuring the next generation of MCMC sampling such as NUTS [44, 52]. In particular, Stan [53] is a high-level statistical modeling tool for Bayesian inference and probabilistic machine learning, which provide the advanced inference algorithms, enriched with extensive and reliable diagnostics. Although PPLs allow for automatic inference, the performance of these algorithms can be sensitive to the form of parameterization [54, 39, 34]. An appropriate form of reparameterization in the probabilistic models to improve the inference efficiency of system dynamics (governed by a set of nonlinear differential equations) remains a challenging problem. We propose an approach to reparameterize the model configuration space based on the correlations between the parameters induced by the measurement function maps (gain matrix) so that the posterior distributions of the model parameters are explored efficiently. The efficiency is shown in terms of inference diagnostics and computational time. Moreover, we propose sparse priors on the model parameters to consider the small number of EZ's in the model and we also consider priors based on the

behaviour of the model equations. We focussed on the self-tuning variant of HMC called NUTS algorithm in this paper and the convergence diagnostics are monitored carefully to ensure the reliability and accuracy of the estimations.

In Sec. 2, we illustrate the probabilistic model and discuss the proposed prior as well as the reparameterization based approach. We show the results obtained from the NUTS on synthetic datasets in Sec. 3. The analyses for 19 patients and the comparison of the results with the clinical diagnostics are shown in the same section. We discuss the utility and extension of the work in Sec. 4.

2 Materials and methods

The body of the work is based on the virtual brain models and Bayesian inference as schematically illustrated in Fig 1. The workflow to build the Bayesian VEP (BVEP) consists of two main steps: constructing the VEP, a personalized virtual brain model of epilepsy spread [1], and then embedding the VEP model in a Bayesian framework to infer the model parameters with NUTS algorithm. We show that the proposed probabilistic framework in BVEP is able to efficiently invert the nonlinear state-space equations to infer the hidden system dynamics. This approach allows us to accurately estimate the spatial map of epileptogenicity in a personalized virtual brain model of epilepsy spread by taking advantage of flexible probabilistic inference in PPLs such as Stan [39]. The brief description of the empirical data collection process and the pre-processing steps are provided in Supplementary Sec. 1.1 and Sec. 1.2, respectively. The detailed description of the whole-brain model is provided in Supplementary Sec. 1.3 and the description of the gain matrix is provided in Supplementary Sec. 1.4.

2.1 Whole-brain model of epilepsy spread

Typically, to build a virtual brain model, the brain regions are defined using a parcellation scheme and then placing a set of mathematical equations (here Epileptor model, [16]) to capture the dynamics of the regional brain activity ([6, 1]). Taking such a data-driven approach to incorporate the subject-specific brain's anatomical information, the network edges are then represented by structural connectivity of the brain, which are obtained from non-invasive imaging data of individual patients ([1, 8]). The brain regions are defined according

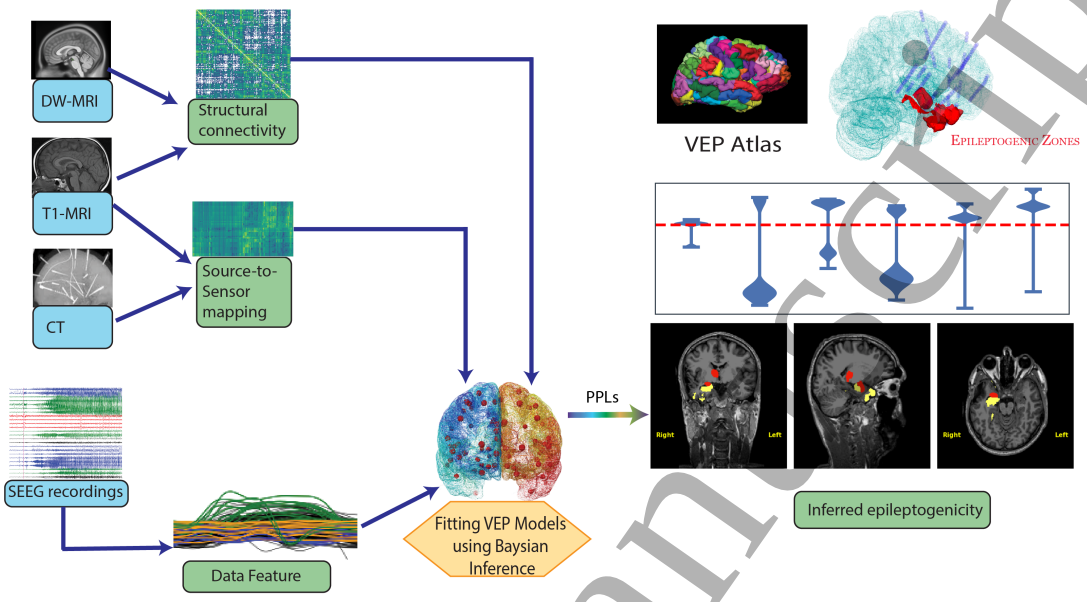

Figure 1: Schematic illustration of workflow in the BVEP. The approach to build the BVEP comprises two main steps: constructing the VEP model, and then embedding VEP in a PPL tool to infer and validate the model parameters. To build the VEP model, we take the following steps: First, the patient undergoes non-invasive brain imaging (MRI, diffusion tractography, computed tomography (CT)). Based on these images, the brain network anatomy including brain parcellation and the patient's connectome are provided from the reconstruction pipeline. Then, a neural population model is selected for each brain region to define the network model. In VEP, the Epileptor model is defined on each network node that are connected through structural connectivity derived from diffusion tractography. The mapping from source to sensor space is encoded in the gain matrix, where the distance between the sensors and the brain regions are obtained using MRI and tomography. The SEEG recordings obtained from the patient are pre-processed to obtain the data feature, which is the envelope time-series, as the target of fitting. Then, model fitting is performed using NUTS algorithm within a PPL tool. Finally, the posterior distribution of excitabilities are obtained for all the brain regions which help in the identification of EZ/PZ which could be the target of resection.

to the VEP atlas ([55]), a cortical and sub-cortical parcellation of the brain developed specifically for the use in domains of epileptology and functional neurosurgery. The atlas considers reasonable region sizes for improved performance of model inversion techniques ([55]). The VEP atlas indicating the indices and the corresponding brain regions is shown in Tables 1 and 2 of the Supplementary. The structural connectivity matrix based on VEP atlas for a randomly selected patient is shown in Fig 2A. In VEP model, the dynamics of brain network nodes are governed by Epileptor equations ([16]) that are coupled through the structural connectivity matrix. The Epileptor is a taxonomy of dynamical models of seizure evolution and is able to realistically reproduce the dynamics of onset, progression and offset of seizure-like events ([16, 18]). The 2D reduction of VEP model is an adiabatic expression of the time-scale separation which is obtained by averaging some of the variables in the detailed VEP model. The model equation for the 2D VEP are given as:

$$\begin{aligned} \dot{x_{1,i}} &= 1 - x_{1,i}^3 - 2x_{1,i}^2 - z_i + I_1 \\ \dot{z_i} &= \frac{1}{\tau_0}(4(x_{1,i} - \eta_i) - z_i - K\sum_{j=1}^{N} C_{ij}(x_{1,j} - x_{1,i})), \end{aligned} \tag{1}$$

Depending on the value of excitability parameter η, the 2D Epileptor exhibits different stability regimes. The details regarding linear stability analysis and parameter space exploration of 2D Epileptor are provided in [18] and [1]. For $\eta < \eta_C$, a trajectory in the phase plane is attracted to the single stable fixed-point of the system on the left branch of the cubic x-nullcline. In this regime, the Epileptor is said to be healthy, meaning not triggering epileptic seizure without external input. As the value of η increases, the z-nullcline moves down and a bifurcation occurs at $\eta = \eta_C$ corresponding to a seizure onset. For $\eta > \eta_C$, the system exhibits an unstable fixed-point allowing a seizure to happen (the Epileptor is said to be epileptogenic and corresponding brain region is referred to as EZ). Isolated nodes display a bifurcation at the critical value $\eta_C = -2.05$ ([18, 1]). In this study, we use the 2D VEP model for Bayesian inference of spatial map of epileptogenicity to reduce the computational cost associated with the model parameter estimation. The 2D reduction allows for faster inversion while enabling us to predict the envelope of fast discharges during the ictal seizure states ([18, 1]).

Based on the above dynamical properties, the spatial map of epileptogenicity across different brain regions comprises the excitability values of EZ (high value of excitability), PZ (smaller excitability value) and all other regions categorized as healthy zone (HZ). Note however, that an intermediate excitability value does not guarantee that the seizure recruits this area as part of the PZ, because the propagation is also determined by various other factors including connectivity and non-linear brain state dependence.

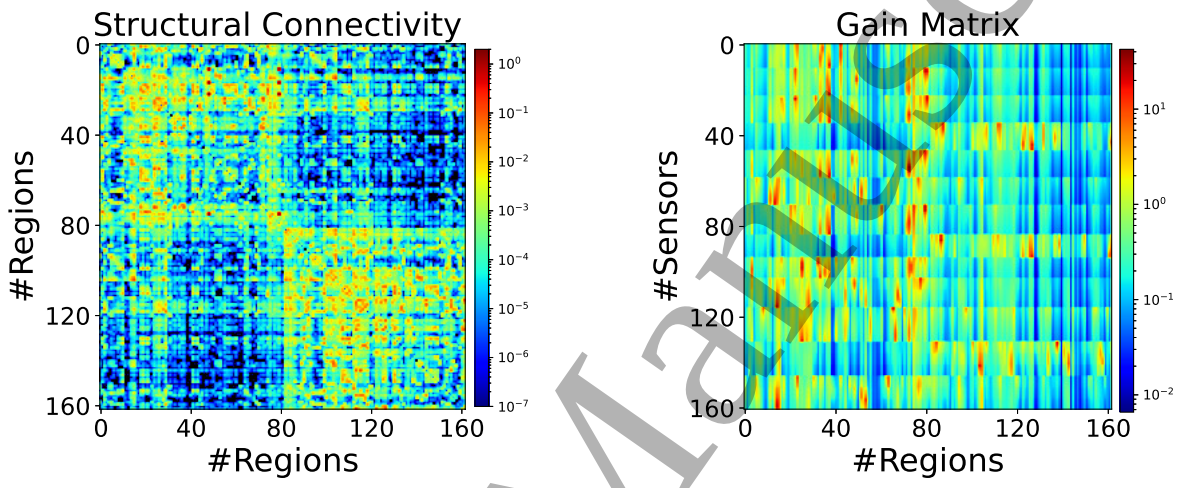

Figure 2: (A) Structural connectivity matrix of a patient, whose entries represent the connection strength between the brain regions, is derived from diffusion MRI tractography. Using VEP atlas, the brain of the patient is parcellated into 162 different regions and (B) Gain matrix of a patient. Each element represents the inverse-squared distance between the region and the sensor.

2.2 Probabilistic whole-brain model

The key component in constructing a probabilistic virtual brain model within a Bayesian framework is the generative model. Given a set of observations, the generative model is a probabilistic description of the mechanisms by which observed data are generated through some hidden states and unknown parameters [56, 57]. In this study, the generative model therefore has a mathematical formulation guided by the dynamical model that describes the evolution of model's state variables, given parameters, over time. This specification is necessary to construct the likelihood function [58, 59]. The full generative model is then

completed by specifying prior beliefs about the possible values of the unknown parameters which are the spatial map of epileptogenicity [29] and the other auxiliary parameters of the generative model. The BVEP brain model presented in this study is built upon two main steps. First, the VEP model equation that provides the basic form of the data generative process describing how the epileptic seizures are generated. Second, the linear combination of the source signals which are observed with independent measurement errors at the sensor level. The generative model in the BVEP is formulated based on a system of nonlinear ordinary differential equations of the form (so-called state-space representation):

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{\theta}), \\ \boldsymbol{y}(t) &= \boldsymbol{h}(\boldsymbol{x}(t), \boldsymbol{\theta}) + \boldsymbol{v}(t), \end{aligned} \tag{2}$$

where, N denotes the number of nodes, $\boldsymbol{x}(t) = (x_{t,1}, z_{t,1}, \ldots x_{t,N}, z_{t,N}) \in R^{2N}$ is a 2N-dimensional vector of system's states evolving overtime, $\boldsymbol{x}(t=0) = \boldsymbol{x}_{t_0}$ is the initial state vector at time $t = 0$. Here, the set of unknown parameters $\boldsymbol{\theta} \in R^p$, where $p = 3N + 2$ (N parameters corresponding to excitability, $2N$ parameters corresponding to $\boldsymbol{x}_{t_0}$, the coupling parameter K and the slow timescale parameter τ_0), contains all the unknown parameters of the VEP model to be estimated. In addition, $\boldsymbol{y}(t) \in R^M$ denotes the measured data subject to the measurement error $\boldsymbol{v}(t)$, where M denotes the number of electrodes. The measurement noise denoted by $\boldsymbol{v}(t) \sim N(0, \sigma^2 I)$ is assumed to follow a Gaussian distribution with mean zero and unknown variance σ^2. Moreover, $\mathbf{f}(.)$ is a vector function that describes the dynamical properties of the system and $\mathbf{h}(.)$ represents a measurement function which is represented as $\boldsymbol{h(x)} = a\, G\boldsymbol{x} + b$, where G represents the gain matrix of dimension $M \times N$, $a \in R^+$ and $b \in R^M$ represent scale and offset parameters, respectively. In this paper, we shall be considering the inference of the model given by the equations combining Eq. 1 and Eq. 2 and given as the following:

$$\begin{aligned} \dot{x_{1,i}} &= 1 - x_{1,i}^3 - 2x_{1,i}^2 - z_i + I_1 \\ \dot{z_i} &= \frac{1}{\tau_0}(4(x_{1,i} - \eta_i) - z_i - K\sum_{j=1}^{N} C_{ij}(x_{1,j} - x_{1,i})), \\ \boldsymbol{y(t)} &= G\boldsymbol{x_{1,t}} + \boldsymbol{\nu(t)}, \end{aligned} \tag{3}$$

where $x_{1,t} \in R^N$ represents the vector denoting the activity of the fast-variable x_1 at time t. Therefore, the total number of parameters to be estimated is $3N + M + 4$. In Bayesian inference, we seek the posterior density $P(\theta|Y)$, which is the conditional distribution of model parameters given the observations [27, 30]. Bayes's Theorem expresses this posterior density in terms of likelihood and prior as follows:

$$P(\theta|Y) = \frac{P(\theta)P(Y|\theta)}{P(Y)},$$

where the denominator $P(Y)$ represents the probability of the data and it is known as evidence or marginal likelihood (in practice amounts to simply a normalization term [24]). We employ the NUTS, a self-tuning variant of HMC algorithm in this paper to sample from posterior density $P(\theta|Y)$ of the model parameters. The performance of HMC is highly sensitive to the step size and the number of steps in leapfrog integrator for updating the position and momentum variables in Hamiltonian dynamic simulation [37]. If the number of steps in the leapfrog integrator is chosen too small, then HMC exhibits an undesirable random walk behaviour similar to Metropolis-Hastings algorithm, and thus algorithm poorly explores the parameter space. If the number of leapfrog steps is chosen too large, the associated Hamiltonian trajectories may loop back to a neighbourhood of the initial state, and the algorithm wastes computation efforts [37, 39]. NUTS extends HMC with adaptive tuning of both the step size and the number of steps in leapfrog integration to sample efficiently from posterior distributions [37, 54, 39].

2.2.1 Mathematical formulation of probabilistic model

We propose a Bayesian estimation of the model parameters involved in the model given by the VEP equations in its 2D formulation using the NUTS algorithm. Let us denote the observed data by $D = \{y_{i,t}, i = 1\ldots, M, t = 1\ldots, T\}$, and the parameter vector by Θ. We denote the matrix of eigenvectors of $G^\top G$ as V. Then, we consider the following independent priors on the regional parameters in the reparameterized space:

$$\begin{aligned} \eta_i^* &\sim N(\mu_{\eta^*}, \sigma_{\eta^*} I_N), \\ z_{init}^* &\sim N(\mu_{z^*}, \sigma_{z^*} I_N), \\ x_{init}^* &\sim N(\mu_{x^*}, \sigma_{x^*} I_N), \end{aligned} \tag{4}$$

and the following independent prior densities for the other auxiliary parameters in the model:

$$\begin{aligned} a &\sim N^+(\mu_a, \sigma_a), \\ b &\sim N(\mu_{\mathbf{b}}, \sigma_{\mathbf{b}} \mathbf{I_M}), \\ \sigma &\sim lognormal(\mu_\sigma, \tau), \\ K &\sim N^+(\mu_K, \sigma_K), \\ \frac{1}{\tau_0} &\sim N^+(\mu_{\tau_0}, \sigma_{\tau_0}). \end{aligned} \tag{5}$$

We consider the non-centered parameterization for all the parameters for efficient sampling from the respective posterior distributions. [60].

To incorporate the information that only few regions are epileptogenic, the mean and standard deviation of the prior distribution on η is considered in such a way that apriori all the brain regions are healthy. This choice of prior for the case of partial epilepsy is practically meaningful because with such a prior knowledge, a region is considered to be healthy unless there is a sufficient evidence based on the data contradicting this assumption and hence, it reduces the false positives. Also, the following relation in the prior means of the regional variables should be considered so that the HMC exploration happens near the stable fixed-points of the 2D Epileptor equations and the HMC sampler start sampling in a physically meaningful space of the high-dimensional parameter space. We consider $(V\mu_{z^*}, V\mu_{x^*})$ to be the fixed-points for a healthy region with parameter $V\mu_{\eta^*}$. In the analyses shown in the paper, we start the likelihood calculation after 10 time points which reduces the dependence between the regional variables x_{init}, z_{init} and η and shows better mixing of the HMC chain. The fixing of x_{init} has also been considered in some analyses to reduce computational time as we need to estimate N fewer parameters in that case. This is accurate as there is only

one stable fixed-point in the 2D epileptor phase-space for the healthy regions and hence, the stable fixed-point is reached quickly if the initialization is close to the fixed-point in the phase-plane. In the simulation studies, we found that the fixing of x_{init} does not introduce large bias and EZ, PZ and HZ are identified well. Also, in the data analysis, we take the number of timepoints in ictal seizure state in such a way that the prior density placed on τ_0, the time-scale of the system, is appropriate. For example, we consider the mean of the prior close to 10 with low standard deviation when approximately 150 timepoints are considered and mean of the prior close to 75 with low standard deviation when approximately 500 timepoints are considered. We perform prior predictive checks to fix the mean and standard deviation of τ_0. The posterior distribution of Θ can be written as:

$$\pi(\Theta|D) \propto \pi(\Theta) \times \prod_{i=1}^{M} \prod_{t=1}^{T} f(y_{i,t}|\Theta), \tag{6}$$

where $\pi(\Theta) = \pi(x_{init})\pi(z_{init})\pi(a)\pi(b)\pi(\sigma)\pi(K)\pi(\tau_0)$ denotes the prior density of the model parameters Θ.

3 Results

3.1 Validation on synthetic SEEG data

In order to study the performance of the proposed reparameterization approach using The Virtual Brain (TVB; [61]), we generate a dataset at the source level from the VEP model (see Eq 1) by considering the structural connectivity and the gain matrix of a real patient. While generating the data, we consider the Euler-Maryama integration with dt=0.1 and $\tau_0 = 10.0$, $a = 1$, $K = 1$. The offset parameters are generated from multivariate normal distribution with mean $\mathbf{10}$ and covariance matrix as I_M. The structural connectivity considered during the data generation corresponds to the anatomical information of a real patient with partial epilepsy and is shown in Fig 2A. Then, we consider the gain matrix of the same patient having dimension 161×162, and multiply it with this generated source level data (see Eq 3 and Supplementary Eq 1.3). The sensor level data is generated by adding an independent Gaussian noise with zero mean and standard deviation $\sigma = 3.0$ at every time point and for every sensor. We consider two of the regions to be EZ's and four regions to be PZ's while rest of the regions are considered to be HZ's. The data considered for the estimation

contains a seizure envelope and the number of timepoints considered are 130 for each sensor. The excitability parameters for the EZ, PZ and HZ are considered to be -1.6, -2.4 and -3.5 respectively. The prior mean and standard deviation of the excitability parameters for all the regions are considered to be -3.5 and 0.1, respectively. The parameters of the prior distribution considered for this analysis and as described by Eq. 4 and Eq. 5 are provided in Table 1. The performance of the method in terms of the fitted time-series are shown in the Fig 3. We show the generated time-series and the mean of the estimated time-series at the sensors. It can be seen from Fig 3 that the fitted time-series are able to closely replicate the true time-series at the sensors. We also show the violin plots based on the posterior distribution of the excitability parameters obtained from four HMC chains for all the 162 regions in Fig 4A. It can be seen that the estimated EZ's from the reparameterization method coincides precisely with the true EZ's. Also, all the PZ's are correctly identified. We also show the true source activity as well as the mean of the estimated source activities in Supplementary Fig 4. It can be seen that the EZ's and PZ's are well identified with little bias. There is no miscalssification of the HZ as EZ or PZ and vice versa. We also show the violin plots for η^* in Fig 4B. It can be seen that the reparameterization makes the HMC to explore the first few linear combinations while the rest of the weakly estimable combinations are very close to their prior distributions. This is due to the low-rank of the gain matrix and practically, this is because many of the regions are far from the sensors and are not anatomically connected to any region which is close to the sensors. We also show the posterior samples obtained from an MCMC chain for two regions: an EZ and a PZ in Fig 7 A and B. In terms of diagnostics, there are no divergences in the generated samples. The $\hat{R}$ values (the well-established diagnostics for convergence on Monte Carlo sampling) for the posterior samples corresponding to all the parameters are less than 1.1 indicating reliability of the estimates based on the convergence for the posterior distributions of the parameters and hidden-state variables. The run-time of the algorithm with 1000 warmup and 250 sampling iterations using a computer equipped with 8 GB RAM and 3.60 GHz Intel Core(TM) i7-7700 processor is approximately 12 hours. Therefore, by placing appropriate priors combined with the gain matrix based reparameterization, the HMC algorithm can efficiently and accurately discriminate between the epileptogenic and healthy regions.

In the right column of Fig 7, the posterior samples obtained without using the reparame-

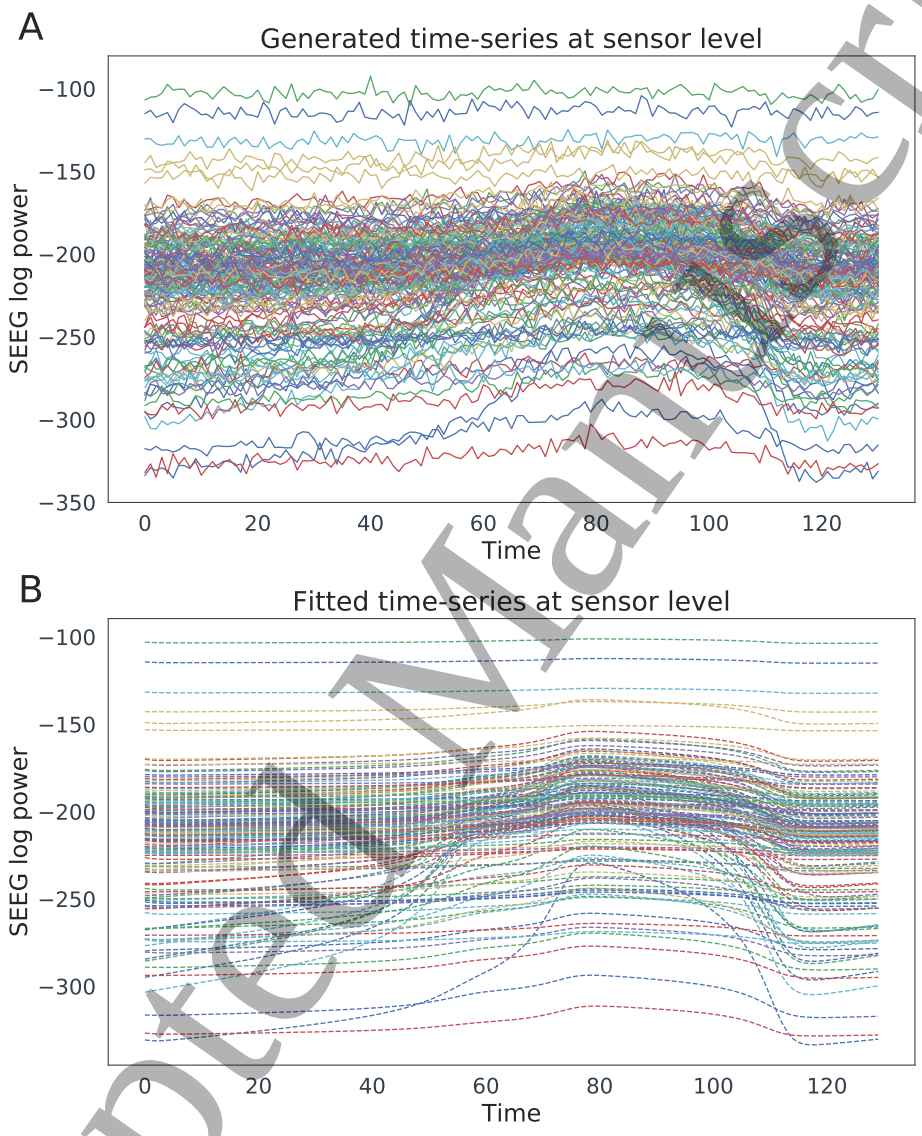

Figure 3: True (A) and Estimated (B) synthetic time-series measured at sensor level. The true time-series is generated using the 2D reduction of the VEP model considering the personalized structural connectivity and the gain matrix of a randomly selected patient.

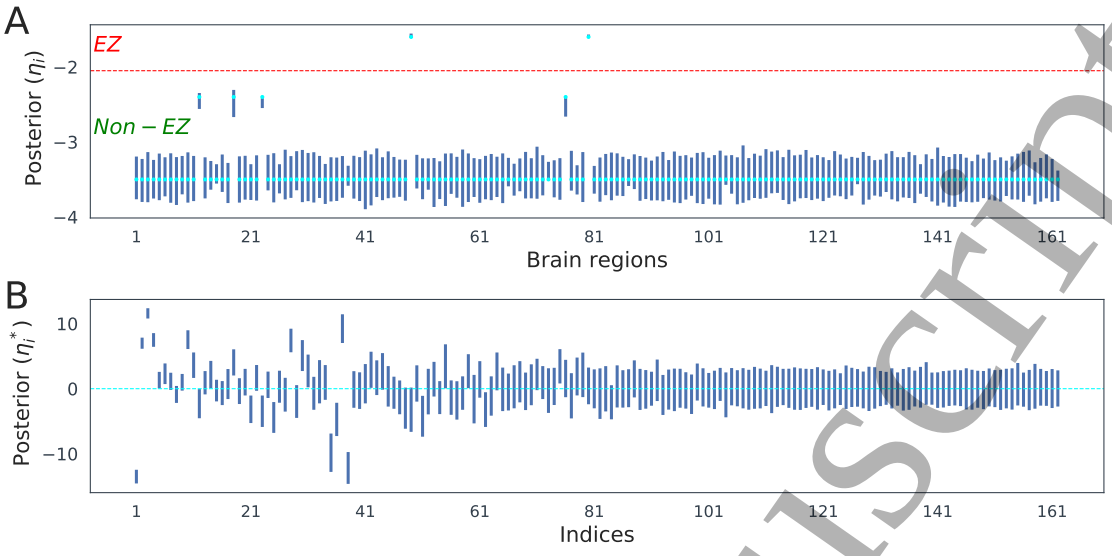

Figure 4: (A) Violin plot of the posterior distribution of η in the VEP model. The dots in cyan colour show the true values and the posterior distribution for different regions is shown in blue. The red dashed horizontal line shows the critical value of excitability for uncoupled Epileptor model. (B) Violin plot of the posterior distribution of η^* using reparameterization method. The dashed horizontal line in cyan colour shows the prior mean.

terization are shown. It can be seen from the posterior samples, that even with large warmup iterations, the mixing between the generated samples is very low as can be seen for an EZ, and for a PZ, many generated samples get stuck at the same value. This results in large autocorrelation between the samples rendering the effective sample size per iteration very low (e.g. it is less than 25 for more than 80% of the excitability parameters). The plots for log-probabilities are also shown in the Fig 7C and Fig 7F, and again it can be seen that the reparameterization based approach improves the mixing in the log-probabilities and shows faster convergence to the typical set.

It can be concluded that the reparameterization based approach solves the issues considered without reparameterization in terms of computational time as well as the diagnostics. Therefore, the proposed novel approach based on reparameterization should be considered for efficient estimation of the parameters of the VEP model. The consideration of priors also play a critical role in removing some inherent degeneracy in the model and makes the parameters identifiable. Some examples considering these issues are shown in the next subsections.

Table 1: Parameters of prior distribution considered for the analysis of simulated data in Sec. 3.1.

Parameter	Value	Parameter	Value
μ_{η^*}	-3.5	σ_{η^*}	0.1
μ_{z^*}	5.0	σ_{z^*}	0.1
μ_{x^*}	-2.0	σ_{x^*}	0.1
μ_a	0.0	σ_a	2.0
μ_b	0.0	σ_b	10.0
μ_σ	0.0	τ	1.0
μ_K	0.0	σ_K	10.0
μ_{τ_0}	0.1	σ_{τ_0}	0.01

3.1.1 Robustness with respect to Laplace prior

In Fig. 5, we show the violin plots of the posterior distribution of η when we consider the prior distribution of η as Laplace with mean -3.5 and scale parameter 0.1. The data considered for this analysis is the same as in Sec. 3.1 and the prior distribution for other parameters is the same as in Sec. 3.1. It can be seen that the posterior distribution of η contains the true values and none of the EZ, PZ or HZ are mis-classified. The diagnostics are also good with no divergences and none of the R-hat values are more than 1.1. This shows the robustness of the procedure with respect to different sparse priors.

3.1.2 Model identifiability using prior

We consider the effect on the posterior distribution of η due to the change in its prior mean. The model is non-identifiable for η for the regions which are HZ and the offset parameter b. This is due to the constant time-series of hidden states of HZ's when they are at their stable fixed-points. The Bayesian inference procedure makes the parameters identifiable by the consideration of the narrow prior on η, and none of the EZ, PZ and HZ are mis-classified. The posterior distribution of η based on the same data considered in Sec. 3.1 is shown in Fig 6. The mean of the prior on η for this analysis is considered to be -3.0. The prior distribution for all the other parameters are considered to be the same as in Sec. 3.1. It can

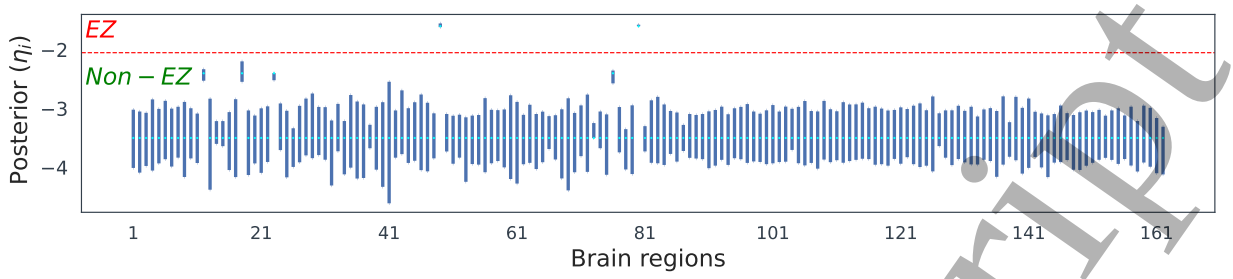

Figure 5: Violin plot of the posterior distribution of spatial parameter mask η in the VEP model. The dots in cyan colour show the true values and the posterior distribution for different regions is shown in blue. The red dashed horizontal line shows the critical value of excitability for uncoupled Epileptor model.

be seen that the posterior distribution of η for the HZ's are close to the prior mean showing that the consideration of the prior helps in removing degeneracy in η and none of the regions are mis-classified.

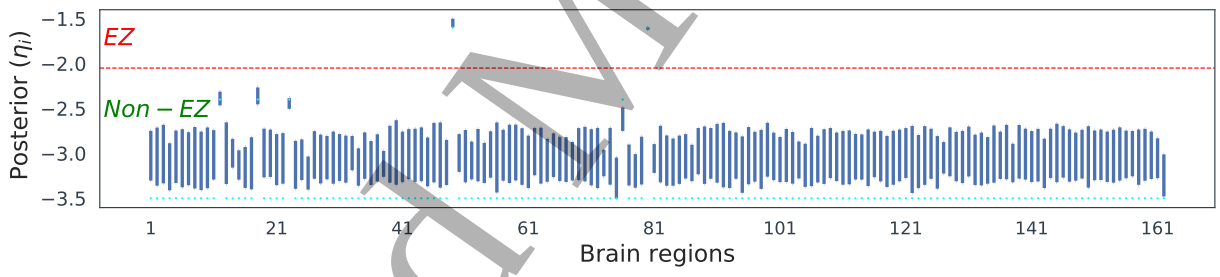

Figure 6: Violin plot of the posterior distribution of spatial parameter mask η in the VEP model. The dots in cyan colour show the true values and the posterior distribution for different regions is shown in blue. The red dashed horizontal line shows the critical value of excitability for uncoupled Epileptor model.

3.2 Bayesian inference of empirical SEEG data

While modelling the empirical SEEG data using the VEP model (Eq 1, Eq 3 and Supplementary Eq. 1.3), it should be noted that the dynamics explained by the model captures the envelope of the hidden-state variables and hence, the envelope of SEEG observations. Therefore, in the pre-processing step, we consider filtering of the data before analysing to

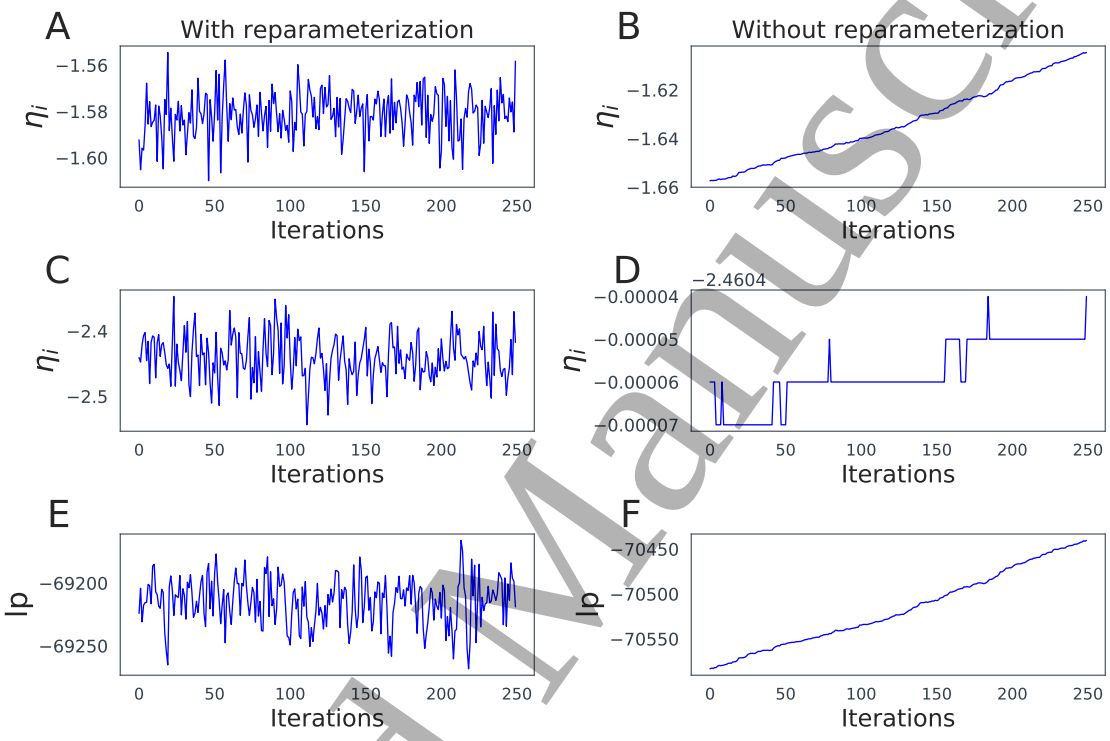

Figure 7: Monitoring the MCMC convergence in VEP model based on the trace plots obtained from the posterior samples using reparameterization (A, C, E) and without reprameterization (B, D, F) for an Epileptogenic zone (EZ) (A, D), a Propagation zone (PZ) (B, E) and log-probability (C, F). Using reparameterization (left column), MCMC samples show well mixing indicating that the HMC converges to the target distribution whereas without reparameterization (right column), HMC is still exploring the typical set and manifest high auto-correlation.

remove very high frequency and very low frequency signals. Then, we extract the envelop in the SEEG time-series and consider the logarithm of the time-series of the envelope as the target of estimation. We consider the connectivity and gain matrices for the patients as discussed in Supplementary Sec. 1.1 and 1.4. We illustrate the reparameterization method as applied to a real SEEG dataset obtained from a partial epilepsy patient. We run four parallel HMC chains to obtain the posterior samples in this case. For this particular analysis, we consider the prior distribution of η as Gaussian with mean -3.0 and sd 0.3.

The mean estimated source activities along with the violin plots of the posterior distribution of η for all the regions are shown in Fig 8. We next select the probable EZ's in this case by choosing only those regions for which $P(\eta \geq -2.05) > 0.25$ and show the posterior distributions of excitability parameters using violin plots for these selected regions. For this specific dataset, based on our methodology and the obtained HMC samples, there are six brain nodes which have at least 0.25 probability to be epileptogenic. Two of the identified regions (Right Amygdala and Right Thalamus) have more than 0.95 probability to be epileptogenic. There are three EZ's and eight PZ's identified by the clinicians for this patient. Among the selected six regions (Right F3 pars Orbitalis, Right Temporal Pole, Right Fusiform Gyrus, Right Hippocampus Anterior, Right Amygdala and Right Thalamus), Right Hippocampus anterior and Right Amygdala are also identified by the clinicians as EZ while two of the regions, Right Temporal Pole and Right Thalamus are identified by clinicians as PZ. Two regions which are identified clinically as HZ namely, Right F3 pars Orbitalis and Right Fusiform Gyrus are among the selected six regions as EZ and six of the clinically identified PZ's are shown to be HZ according to the proposed methodology based on four chains. The posterior distributions of η corresponding to three of the EZ's namely, Right Temporal Pole, Right Fusiform Gyrus and Right Hippocampus anterior have substantial posterior probabilities of $\eta < -2.05$, and hence, not to be EZ. The mean estimated phase-planes corresponding to an identified EZ and an identified HZ are shown in Fig 9. The practical application of the methodology creates additional issues. We discuss the issue of multi-modality in the posterior distribution of the parameters in Supplementary Sec 1.9. The consideration of sparsity in the number of EZ's using priors is discussed in Supplementary Sec. 1.8.

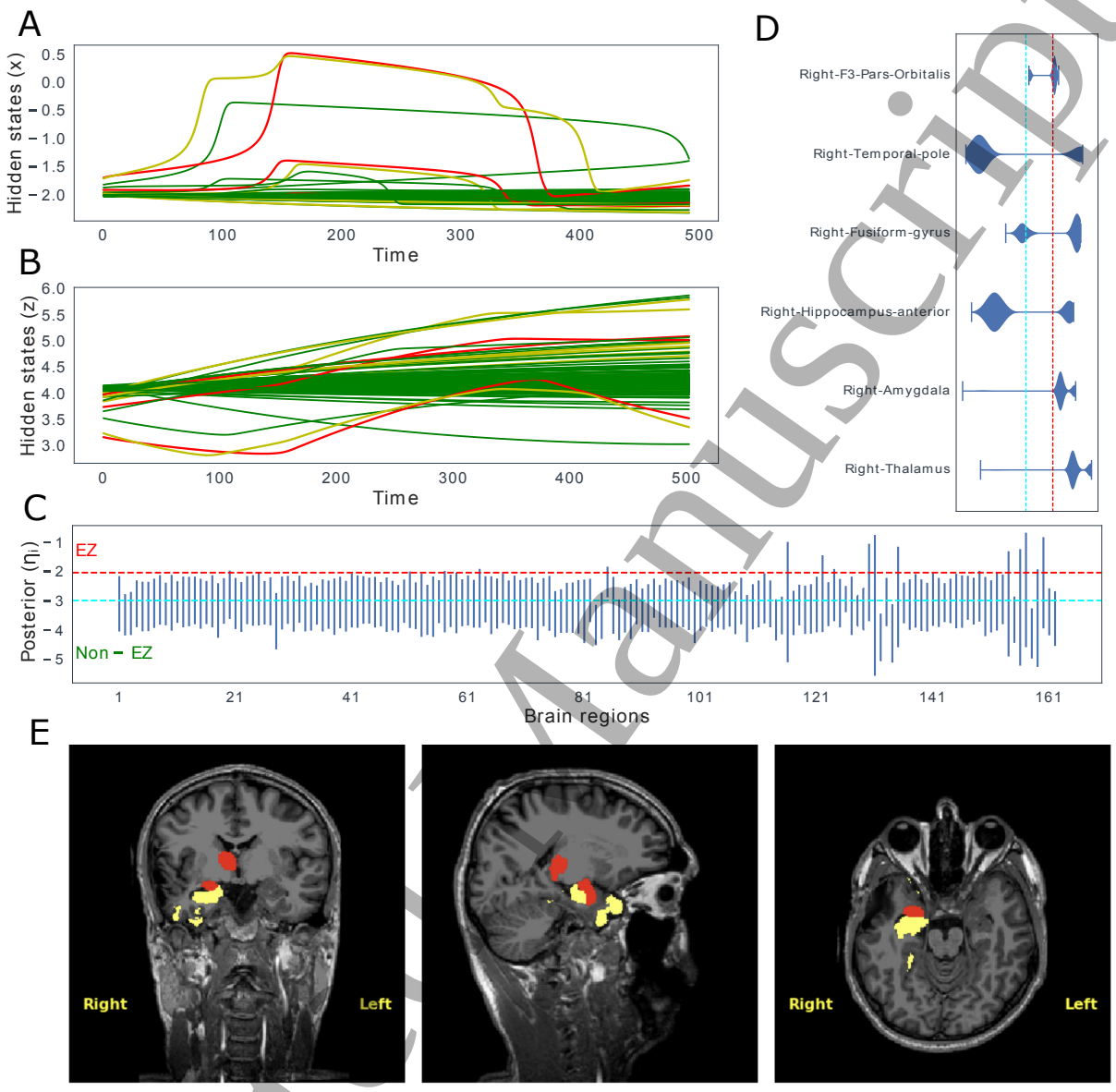

Figure 8: Bayesian estimation for empirical SEEG data of a selected patient: (A) Mean estimated hidden states for the fast variable and (B) slow variable, at source level. The brain regions identified by clinicians corresponding to EZ's, PZ's and HZ's are shown in red, yellow and green, respectively. (C) Estimated posterior distribution for all the brain regions (D) and the regions having highest probabilities to be identified as EZ by the VEP. (E) Heatmap of the predicted spatial map of epileptogenicity using VEP model reported to the clinicians. The regions identified as EZ's with more than 0.75 probability and with more than 0.25 probability are shown in red and yellow, respectively.

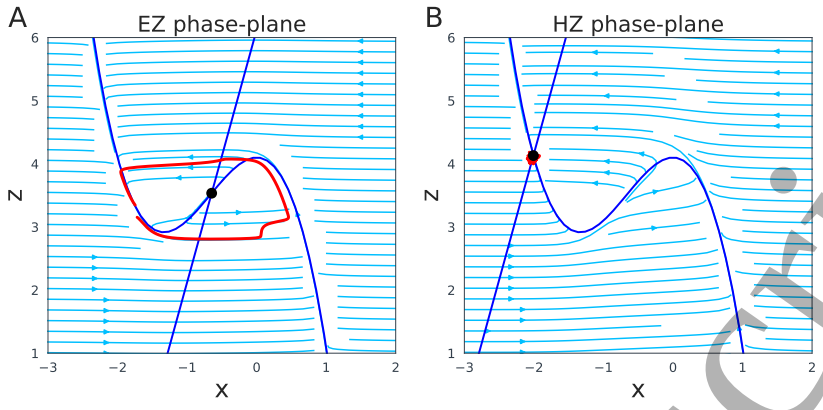

Figure 9: Estimated phase-planes: The phase-plane showing the hidden states corresponding to (A) a region identified as EZ and (B) a region identified as HZ. The black dot represents the fixed point based on the 2D Epileptor equations. The mean phase-plane shows limit-cycle for the estimated EZ, while the mean phase-plane shows all the values close to the stable fixed-point for the HZ.

3.3 Group analysis

We consider the SEEG data obtained from nineteen epileptic patients who underwent standard clinical evaluation at La Timone hospital in Marseille for the study. The details of the patients considered for the analysis are provided in Supplementary Table 3. We compare the results obtained from the proposed methodology with the clinical diagnostics. We consider the same prior density for the excitability parameters for all the regions meaning that there is no assumed difference between the excitabilities corresponding to different regions a priori. Also, the prior distribution for excitabilities in all the analyses are considered such that a priori all the regions are considered to be healthy with at least 0.95 probability. We run four parallel HMC chains for each dataset. The convergence diagnostics of the chains are assessed and no divergences are obtained in the considered chains. The within-chain split-$\hat{R}$ values [63] are also less than 1.5 for all η in most of the chains which is acceptable for such complicated multimodal posterior distributions [62]. To compare the results, we combine the obtained posterior distribution of the excitabilities and consider a region i to be epileptogenic for a patient if the probability of its excitability is given by $P(\eta_i > -2.05) > 0.25$. The cutoff -2.05 corresponds to the critical value of an isolated node given 2D Epileptor dynamics. We show the scatter-plot of the number of identified EZ's by the proposed methodology against

the number of identified EZ's based on the clinical diagnostics for each of the analysed patient in Fig 10A. It can be seen that generally the numbers of identified EZ's from our methodology are more than those based on the clinicians' diagnostics. This may be due to the consideration of all the regions in our estimation unlike the selection of the regions among only those few regions which are close to the electrodes by the clinicians. These estimates can help the clinicians in identifying some unexplored regions which could be epileptogenic and where the electrodes could not be implanted. In Fig 10B, we plot the proportion of the number of regions identified as EZ by our inference methodology among different groups of regions classified as EZ, PZ and HZ according to the clinical diagnostics for patients belonging to Engel I (n=10), Engel II (n=2), Engel III (n=4) and Engel IV (n=3). The Engel scores are based on the outcome of the surgery: (i) Engel I: patients who are seizure free, (ii) Engel II: patients who show rare disabling seizures, (iii) Engel III: patients with minimal improvement and (iv) Engel IV: patients with no improvement. Every index in the plot represents a patient and the plotted points represent the proportions. The round dots, the triangular dots and the '+' dots show the proportion for the clinically classified EZ's, PZ's and HZ's, respectively. It can be seen from the figure that in general, the proportion of regions identified as EZ based on the methodology which are clinically identified as HZ is very low showing that the consideration of high probability of the prior density helps in lowering the false positive rates. It can be observed that the proportion of EZ's identified from our methodology is slightly higher among the clinically identified EZ's for Engel-I patients with respect to other groups of patients but general statistical conclusions should not be drawn because of the low sample size of the number of patients in different groups. From Fig 10C, it can be observed that some of the clinically identified PZ's are identified as EZ's using the methodology. The proposed BVEP workflow with the reparameterization method can help the clinicians to examine these regions more closely before surgery.

It can be seen that obtained estimates match some of the clinically identified EZ's in most of the patients which shows that the personalized BVEP model can be considered to model the epilepsy propagation in the brain. Though, as a precaution, we recommend running multiple parallel chains and including some clinical information in the prior to make better use of the proposed methodology in real life therapeutic applications. The methodology may therefore help in identifying the epileptogenic regions for the patients based on the SEEG

data which in turn can be considered to be resected during surgery to stop the recurrence of seizures in the patients.

The results obtained from the group analysis show that the proposed methodology can be a highly valuable alternative to the model-free approach employed by the clinicians to identify the EZ's. In the model-free approach, the EZ's are identified as the brain regions which are close to the electrodes showing large activities during the seizure. The model-based approach used in BVEP is different from the model-free approach considered by the clinicians due to the following reasons: a) The VEP models the activity of all the regions in the brain and hence, even the subcortical EZ's can be identified in contrast to the model-free approach where only the brain regions which are close to the electrodes can be identified as EZ's. b) The Bayesian estimation technique is able to quantify the probability of being epileptogenic for each of the brain regions. This probabilistic quantification, which enables one to identify different modes of getting the same fit, is missing from the model free approach. Therefore, the regions identified by the clinicians and VEP estimates could be different and moreover, each of the brain regions is assigned a probability of being epileptogenic in the BVEP model. The BVEP estimation is also highly valuable when the surgery fails and VEP discovers more regions with associated probabilities of being EZ, rather than clinicians making binary estimation.

4 Discussion

This work is an attempt to merge the theoretical understanding of the whole-brain dynamics with the physically measured data using state-of-the-art advances in PPLs and Bayesian inference algorithms. We have provided an efficient probabilistic methodology to infer the posterior distribution of the model parameters involved in the biophysically realistic VEP model of epilepsy spread in the brain. The estimated time-series provides good fit on the extracted envelope from the raw SEEG data which makes the methodology more valuable. The prior distribution constraints the model parameters to be in a biophysically relevant space and the obtained posterior distribution quantifies the uncertainty in the regional excitability parameters. The proposed reparameterization method enables the accurate estimation of the uncertainty by efficiently generating samples from the complicated posterior distribution

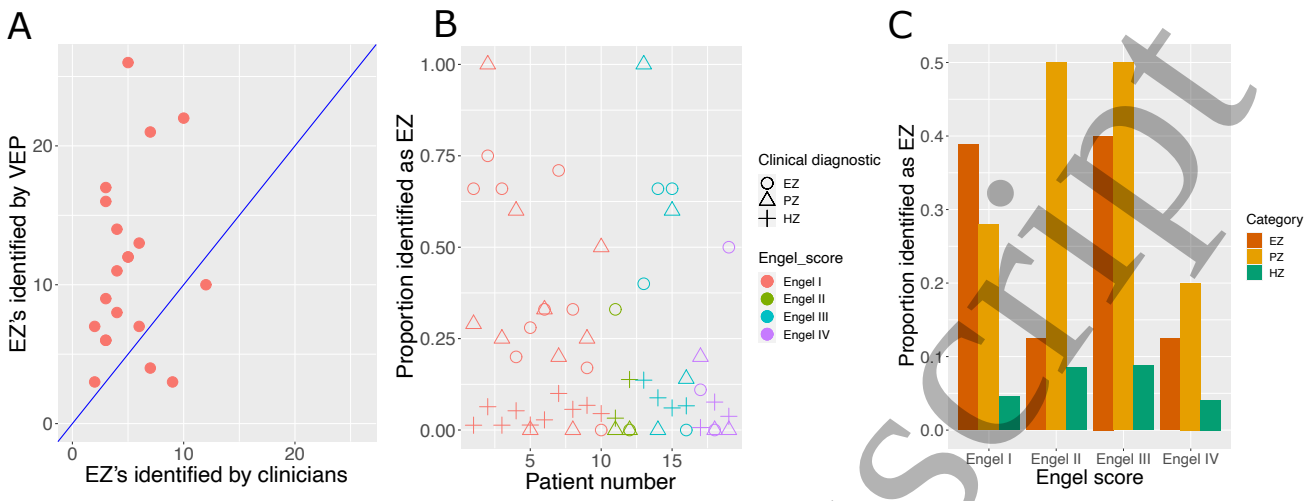

Figure 10: Comparison of VEP estimates with the clinical diagnostics: (A) The number of identified EZ's based on clinical diagnostics (x-axis) vs the number of identified EZ's based on VEP methodology (y-axis). The blue dashed line $y = x$ is drawn as a reference. (B) The proportion of EZ identified by the VEP among the clinically classified EZ, PZ and HZ for patients with different Engel scores for individual patients, (C) and by group, based on Engel scores.

of the model parameters. The estimated joint distribution of the excitability parameters provide the probabilities of different plausible sets of regions to be identified as epileptogenic together. The obtained estimates also inform about the nodes whose activities could not be observed due to their low connectivity with the regions close to the electrodes. The posterior distribution of the excitability parameters of these regions are similar to their corresponding prior distributions.

We provide a link between the virtual brain models, personalized treatment and systematic Bayesian inversion in PPLs. The accuracy and the reliability of the estimation is carefully investigated by the HMC diagnostics. The virtual brain models combine the anatomical connectivity with mathematical formulation of brain activity. However, the main challenge lies in inferring the system dynamics explained by the model during the activity. In this study, dynamical systems provided useful tools for the inversion of slow-fast system dynamics having incomplete observations in the phase-plane. For instance, η is the bifurcation parameter in the Epileptor model, i.e. small changes made to its value causes a sudden change in the system behaviour. This allows us to classify all the brain regions in three categories

based on the values of η and the structural connectivity. The knowledge that the considered system exhibits only one fixed-point in the phase-plane allows us to avoid the mismatch between the observed time-series and the estimated time-series without losing the biophysical relevance of the model. We incorporate this knowledge systematically in the model inversion by constraining the activity using the prior density over the regional parameters. The priors for initial values (x_{init}, z_{init}) and spatial parameter mask η are considered in such a way that the activity of the fast variable remains close to its stable fixed-point during the seizure for the healthy regions. The initialisation of the time-series of hidden state-space variables close to the steady-state avoids the initial transition in the time-series. This illustrates the important role which the system dynamics can play in guiding the estimation methodology for the physically relevant inversion of high-dimensional complex models.

The multimodality in the posterior distribution necessitates the application of MCMC methods compared to maximum a-posteriori estimation (MAP) and VI as shown in Fig 6C of the Supplementary. The application of other methodologies for identification of EZ/PZ such as Approximate Bayesian Computation (ABC)-related methods are not feasible for application in virtual brain models with large number of parcellations as these methods suffer from the curse of dimensionality which means that the number of simulated samples needed to provide a good estimate of the posterior can be prohibitively expensive. The reduction of the data to low-dimensional summary statistics and the distance tolerance discard some of the information about the parameters in the data which reduces the quality and accuracy of inference. Moreover, for such non-linear models, it is challenging to extract low-dimensional summary statistics which is sufficient statistic for the unknown parameters. Additionally, the ABC methods are sensitive to a threshold value to accept or reject samples. This makes the application of ABC related methods prohibitive in high-dimensional non-linear models. The NUTS algorithm considers the gradient information which avoids random-walk in high-dimensional space and allows it to converge to high-dimensional target distributions much more quickly. Therefore, the estimation using self-tuning and gradient-based NUTS algorithm should be considered when such non-linear large network models are applied to real datasets. However, as the proposed methodology is personalized and each seizure of every patient is analysed separately, group-level information regarding epilepsy is not incorporated in the method. Therefore, hierarchical Bayesian methods for group analysis should

be considered in future studies to combine information obtained from several patients across multiple seizures .

Acknowledgements

The authors also gratefully acknowledge the computing time granted through Swiss National Supercomputing Centre on the supercomputer Piz Daint at Zurich.

Competing interests

The authors declare that no competing interests exist.

Information sharing statement

The patient data sets cannot be made publicly available due to the data protection concerns regarding potentially identifying and sensitive patient information. Interested researchers may access the data sets by contacting Clinical Data Manager Aurélie Ponz (aurelie.ponz@univ-amu.fr) at the Institut de Neurosciences des Systèmes, Aix-Marseille Université. The main source codes needed to reproduce the presented results are available on GitHub (https://github.com/ins-amu/BVEP).

Funding

This work was funded by the French National Research Agency (ANR) as part of the second "Investissements d'Avenir" program, ANR-17-RHUS-0004, EPINOV (https://anr.fr), by European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 785907 and 945539, Human Brain Project SGA2 and SGA3 (https://ec.europa.eu/programmes/horizon2020), by European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement and No. 826421, VirtualBrainCloud (https://ec.europa.eu/programmes/horizon2020), PHRC-I 2013 EPISODIUM (grant number 2014-27), the Fondation pour la Recherche Médicale (DIC20161236442), and by SATT Sud-Est, 827-SA-16-UAM (https://www.sattse.com) to

Viktor Jirsa. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

[1] Jirsa VK, Proix T, Perdikis D, Woodman MM, Wang H, Gonzalez-Martinez J, et al. The Virtual Epileptic Patient: Individualized whole-brain models of epilepsy spread. NeuroImage. 2017;145:377–388. doi:https://doi.org/10.1016/j.neuroimage.2016.04.049.

[2] Bernard C, Jirsa VK. Virtual Brain for neurological disease modeling. Drug Discovery Today: Disease Models. 2016;19:5–10. doi:https://doi.org/10.1016/j.ddmod.2017.05.001.

[3] McIntosh AR, Jirsa VK The hidden repertoire of brain dynamics and dysfunction. Network Neuroscience. 2019;3(4):994–1008. doi:https://doi.org/10.1162/netn_a_00107.

[4] Hashemi M, Vattikonda AN, Sip V, Diaz-Pier S, Peyser A, Wang H, et al. On the influence of prior information evaluated by fully Bayesian criteria in a personalized whole-brain model of epilepsy spread. PLOS Computational Biology. 2021;17(7):e1009129. doi:https://doi.org/10.1371/journal.pcbi.1009129.

[5] Bressler SL. Large-scale cortical networks and cognition. Brain Research Reviews. 1995;20(3):288–304. doi:https://doi.org/10.1016/0165-0173(94)00016-I.

[6] Sanz-Leon P, Knock SA, Spiegler A, Jirsa VK. Large-scale cortical networks and cognition. Neuroimage. 2015;111:385–430. doi:https://doi.org/10.1016/j.neuroimage.2015.01.002.

[7] Deco G, Jirsa VK, McIntosh AR. Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat Rev Neurosci. 2011;12:43–56. doi:https://doi.org/10.1038/nrn2961.

[8] Bansal K, Nakuci J, Muldoon SF. Personalized brain network models for assessing structure-function relationships. Current Opinion in Neurobiology. 2018;52:42–47. doi:https://doi.org/10.1016/j.conb.2018.04.014.

[9] Jirsa VK, Jantzen KJ, Fuchs A and Kelso JAS. Spatiotemporal forward solution of the EEG and MEG using network modeling. IEEE Transactions on Medical Imaging. 2002;21(5):493–504. doi:https://doi.org/10.1109/TMI.2002.1009385.

[10] Jirsa VK, Sporns O, Breakspear M, Deco G, and McIntosh AR. Towards The Virtual Brain: network modeling of the intact and the damaged brain. Archives Italiennes de Biologie. 2010;148(3):189-205.

[11] Ghosh A, Rho Y, Mcintosh AR, Kotter R., Jirsa, VK. Noise during Rest Enables the Exploration of the Brain's Dynamic Repertoire. PLOS Computational Biology. 2008;4(10):e1000196. doi:https://doi.org/10.1371/journal.pcbi.1000196.

[12] Melozzi F, Bergmann E, Harris JA, Kahn I, Jirsa VK, Bernard C. Individual structural features constrain the mouse functional connectome. PNAS. 2019;116(52):26961–26969. doi:https://doi.org/10.1073/pnas.1906694116.

[13] Falcon M, Riley J, Jirsa VK, McIntosh A, Chen EE, Solodkin A. Functional mechanisms of recovery after chronic stroke: Modeling with the virtual brain. eNeuro. 2016;3:ENEURO.0158. doi:https://doi.org/10.1523/ENEURO.0158-15.2016.

[14] Cabral J, Fernandes HM, Hartevelt TJV, James AC, Kringelbach ML, Deco G. Structural connectivity in schizophrenia and its impact on the dynamics of spontaneous functional networks. Chaos. 2013;23(4):046111. doi:https://doi.org/10.1063/1.4851117.

[15] Proix T, Bartolomei F, Guye M, Jirsa VK. Individual brain structure and modelling predict seizure propagation. Brain. 2017;140(3):641–654. doi:https://doi.org/10.1093/brain/awx004.

[16] Jirsa VK, Stacey WC, Quilichini PP, Ivanov AI, Bernard C. On the nature of seizure dynamics. Brain. 2014;137(8):2210–2230. doi:https://doi.org/10.1093/brain/awu133.

[17] Bartolomei F, Guye M, Wendling F. Abnormal binding and disruption in large scale networks involved in human partial seizures. EPJ Nonlinear Biomed Phys. 2013;1(4):2210–2230. doi:https://doi.org/10.1140/epjnbp11.

[18] Proix T, Bartolomei F, Chauvel P, Bernard C, Jirsa VK. Permittivity coupling across brain regions determines seizure recruitment in partial epilepsy. Journal of Neuroscience. 2014;34(45):15009–15021. doi:https://doi.org/10.1523/JNEUROSCI.1570-14.2014.

[19] Toyoda I, Bower MR, Leyva F, Buckmaster PS. Early activation of ventral hippocampus and subiculum during spontaneous seizures in a rat model of temporal lobe epilepsy. Journal of Neuroscience. 2013;33:11100—11115. doi:https://doi.org/10.1523/JNEUROSCI.0472-13.2013.

[20] Talairach J, Bancaud J. Lesions, irritative zone and epileptogenic focus. Confin Neurol. 1966;27:91–94. doi:https://doi.org/10.1159/000103937.

[21] Bartolomei F, Wendling F, Bellanger JJ, Regis J, Chauvel P. Neural networks involving the medial temporal structures in temporal lobe epilepsy. Clin Neurophysiol. 2001;112:1746–1760. doi:https://doi.org/10.1016/s1388-2457(01)00591-0.

[22] Spencer, S. Neural Networks in Human Epilepsy: Evidence of and Implications for Treatment. Eplilepsia. 2002;43:219–227. doi:https://doi.org/10.1046/j.1528-1157.2002.26901.x.

[23] Bartolomei F, Chauvel P, Wendling F. Epileptogenicity of brain structures in human temporal lobe epilepsy: a quantified study from intracerebral EEG. Brain. 2008;131(7):1818–1830. doi:https://doi.org/10.1093/brain/awn111.

[24] Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis. Chapman and Hall/CRC. 2013.

[25] Aster RC, Borchers B, Thurber CH. Parameter Estimation and Inverse Problems. Elsevier 2018.

[26] Groetsch C. Inverse Problems: Activities for Undergraduates. Mathematical Association of America 1999.

[27] Bishop CM. Pattern Recognition and Machine Learning. Springer 1999.

[28] Friston KJ, Harrison L, Penny W. Dynamic causal modelling. Neuroimage. 2003;19(4):1273–302. doi:https://doi.org/10.1016/s1053-8119(03)00202-7.

[29] Friston KJ, Kahan J, Biswal B, Razi A. A DCM for resting state fMRI. Neuroimage. 2014;94:396–407. doi:https://doi.org/10.1016/j.neuroimage.2013.12.009.

[30] David O, Kiebel SJ, Harrison LM, Mattout J, Kilner JM, Friston KJ. Dynamic causal modeling of evoked responses in eeg and meg. Neuroimage. 2006;30:1255–1272. doi:https://doi.org/10.1016/j.neuroimage.2005.10.045.

[31] Daunizeau J, Stephan KE, Friston KJ. Stochastic dynamic causal modelling of fMRI data: should we care about neural noise? Neuroimage. 2012;62(1):464–481. doi:https://doi.org/10.1016/j.neuroimage.2012.04.061.

[32] Boly M, Moran R, Murphy M, Boveroux P, Bruno M-A, Noirhomme Q, Ledoux D, et al. Connectivity changes underlying spectral EEG changes during propofol-induced loss of consciousness. Journal of Neuroscience. 2012;32(20): 7082–7090. doi:https://doi.org/10.1523/JNEUROSCI.3769-11.2012.

[33] Frassle S, Lomakina EI, Kasper L, Manjaly ZM, Leff A, Pruessmann KP, et al. A generative model of whole-brain effective connectivity. Neuroimage. 2018;179:505-529. doi:https://doi.org/10.1016/j.neuroimage.2018.05.058.

[34] Hashemi M, Vattikonda AN, Sip V, Guye M, Bartolomei F, Woodman MM, Jirsa VK. The Bayesian Virtual Epileptic Patient: A probabilistic framework designed to infer the spatial map of epileptogenicity in a personalized large-scale brain model of epilepsy spread. Neuroimage. 2020;217:116839. doi:https://doi.org/10.1016/j.neuroimage.2020.116839.

[35] Sip V, Hashemi M, Vattikonda AN, Woodman MM, Wang H, Scholly J, Villalon SM, Guye M, Bartlomei F, Jirsa VK. Data-driven method to infer the seizure propagation patterns in an epileptic brain from intracranial electroencephalography. PLOS Computational Biology. 2021;17(2):e1008689. doi:https://doi.org/10.1371/journal.pcbi.1008689.

[36] Girolami M, Calderhead B. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2011;73(2):123–214. doi:https://doi.org/10.1111/j.1467-9868.2010.00765.x.

[37] Hoffman MD, Gelman A. The No-U-turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. J Mach Learn Res. 2014;15(1):1593–1623.

[38] Gelman A, Hwang J, Vehtari A. Understanding predictive information criteria for Bayesian models. Statistics and Computing. 2014;24(6):997–1016. doi:10.1007/s11222-013-9416-2.

[39] Betancourt M, Byrne S, Livingstone S, Girolami M. The geometric foundations of Hamiltonian Monte Carlo. arXiv:14105110. 2014;.

[40] Duane S, Kennedy AD, Pendleton BJ, Roweth D. Hybrid Monte Carlo. Phys Lett B. 1987;195. doi:10.1016/0370-2693(87)91197-X.

[41] Neal A. MCMC using Hamiltonian dynamics. London, UK: Chapman and Hall/CRC; 2011.

[42] Sengupta B, Friston KJ, Penny WD. Gradient-free MCMC methods for dynamic causal modelling. NeuroImage. 2015;112:375 – 381. doi:https://doi.org/10.1016/j.neuroimage.2015.03.008.

[43] Sengupta B, Friston KJ, Penny WD. Gradient-based MCMC samplers for dynamic causal modelling. NeuroImage. 2016;125:1107 –1118. doi:https://doi.org/10.1016/j.neuroimage.2015.07.043.

[44] Carpenter B, Gelman A, Hoffman M, Lee D, Goodrich B, Betancourt M, et al. Stan: A Probabilistic Programming Language. Journal of Statistical Software, Articles. 2017;76(1):1–32. doi:https://doi.org/10.18637/jss.v076.i01.

[45] Prem K. Gopalan, David M. Blei. Efficient discovery of overlapping communities in massive networks. PNAS. 2013;10(36):14534–14539. doi:https://doi.org/10.1073/pnas.1221839110.

[46] Gopalan, P., Hao, W., Blei, D. Scaling probabilistic models of genetic variation to millions of humans. Nature Genetics. 2016; 48: 1587—1590 doi:https://doi.org/10.1038/ng.3710.

[47] Kucukelbir A, Tran D, Ranganath R, Gelman A, Blei D. Automatic differentiation variational inference. Journal of Machine Learning Research. 2017; 18(14):1–45.

[48] Jordan MI, Ghahramani Z, Jaakkola TS, Saul, LK. An introduction to variational methods for graphical models. Machine Learning. 1999; 37: 183—233.

[49] Wainwright MJ, Jordan MI. Graphical models, exponential families, and variational inference. Foundations and Trends in Machine Learning. 2008; 1: 1–305. doi:http://dx.doi.org/10.1561/2200000001.

[50] Kucukelbir A, Rajesh R Gelman A, Blei D. Automatic Variational Inference in Stan. Advances in Neural Information Processing Systems. 2015. url:https://proceedings.neurips.cc/paper/2015/file/352fe25daf686bdb4edca223c921acea-Paper.pdf.

[51] Bingham E, Chen JP, Jankowiak M, Obermeyer F, Pradhan N, Karaletsos T, et al. Pyro: Deep Universal Probabilistic Programming. Journal of Machine Learning Research. 2018.

[52] Salvatier J, Wiecki TV, Fonnesbeck C. Probabilistic programming in Python using PyMC3. PeerJ Computer Science. 2016;doi:10.7717/peerj-cs.55.

[53] Stan Development Team. Stan Modeling Language Users Guide and Reference Manual, Version 2.18.0.; 2018. Available from: http://mc-stan.org/.

[54] Betancourt M, Girolami M. Hamiltonian Monte Carlo for hierarchical models. arXiv:13120906. 2013.

[55] Wang HE, Scholly J, Triebkorn P, Sip V, Medina Villalon S, Woodman MM, Le Troter A, Guye M, Bartolomei F, Jirsa VK. VEP atlas: An anatomic and functional human brain atlas dedicated to epilepsy patients. J Neurosci Methods. 2021; 348:108983.

[56] Daunizeau J, Friston K, Kiebel S. Variational Bayesian identification and prediction of stochastic nonlinear dynamic causal models. Physica D: Nonlinear Phenomena. 2009; 238: 2089–2118.

[57] Daunizeau J, Adam V, Rigoux L. Vba: A probabilistic treatment of nonlinear models for neurobiological and behavioural data. PLOS Computational Biology. 2014; 10: 1–16.

[58] Cooray GK, Sengupta B, Douglas P, Englund M, Wickstrom R, Friston K. Characterising seizures in anti-NMDA-receptor encephalitis with dynamic causal modelling. NeuroImage. 2015;118:508 – 519. doi:https://doi.org/10.1016/j.neuroimage.2015.05.064.

[59] Hashemi M, Hutt A, Buhry L, Sleigh J. Optimal Model Parameter Estimation from EEG Power Spectrum Features Observed during General Anesthesia. Neuroinformatics. 2018;16(2):231–251. doi:10.1007/s12021-018-9369-x.

[60] Betancourt M. A Conceptual Introduction to Hamiltonian Monte Carlo. arXiv:70102434. 2014.

[61] Sanz Leon P, Knock S, Woodman M, Domide L, Mersmann J, McIntosh A, et al. The Virtual Brain: a simulator of primate brain network dynamics. Frontiers in Neuroinformatics. 2013;7:10. doi:http://doi.org/10.3389/fninf.2013.00010.

[62] Yao Y, Vehtari A, Gelman A Stacking for non-mixing Bayesian computations: the curse and blessing of multimodal posteriors. arXiv:2006.12335 2020.

[63] Vehtari A, Gelman A, Simpson D, Carpenter B, and Burkner PC Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC. Bayesian Analysis. 2020; 16(2): 667–718.