Supporting Information:

The Hyper-inverse Wischart spatial autoregressive hidden markov model (HIW-spatial AR-HMM) model involves a number of correlated hidden Markov models (HMM), one for each EEG channel, and one *event* model across all of the channels. The channel-specific HMMs describe activity local to each channel, and event HMM describes the correlation structure between the individual channels. In this work, we primarily focus on the *event states* assigned at each time point the burst and seizure events, as they allow us to most easily compare different events to each other. See Wulsin et al, 2014 for full technical details of the model and its application to epileptic events on the EEG. In this work, we assumed a discrete set of 30 distinct event states and 50 channel states, as we found empirically that these captured the space of different event and channel activity. Bayesian posterior inference from the events of each dog was performed via 10 chains of Markov chain Monte Carlo (MCMC), which is a common method in Bayesian analysis to sampling from the posterior distribution. The first 2500 MCMC iterations were discarded as burn-in to ensure convergence. After 10-sample thinning to remove autocorrelation and ensure independence of samples, 750 MCMC samples were taken from each chain after 10-sample thinning. Each sample provides a specific state sequence that can be further assessed below.

Burst-Seizure Similarity Modeling

For a particular MCMC sample s, consider two events indexed by e_1 and e_2 with time lengths $T^{(e_1)}$ and $T^{(e_2)}$ and event state sequences $Z_{1:T^{(e_1)}}^{(e_1,s)}$ and $Z_{1:T^{(e_2)}}^{(e_2,s)}$, respectively. Over

S MCMC samples, the number of times $C_{t_1,t_2}^{(e_1,e_2)}$ a particular time point t_1 in event e_1 is assigned to the same event state as a particular time point t_2 in event e_2 is given by

$$C_{t_1,t_2}^{(e_1,e_2)} = \left|\left\{ s \,\middle|\, Z_{t_1}^{(e_1,s)} = Z_{t_2}^{(e_2,s)}, s = 1, \ldots, S\right\}\right|.$$

We often call these frequencies between two events the *co-states* between e_1 and e_2. The matrix $C^{(e_1,e_2)}$ gives the co-states over all the possible pairs of time points in events e_1 and e_2. For one of the events e_1, we are interested in the maximum co-states across the time points of the other event,

$$m_{t_1}^{(e_1,e_2)} = \max_{t_2=1,\ldots T^{(e_2)},} C_{t_1,t_2}^{(e_1,e_2)},$$

where the definition of $m_{t_2}^{(e_2,e_1)}$ is equivalent. The vector $\boldsymbol{m}^{(e_1,e_2)}$ thus gives the maximum co-state frequency in each time point of event e_1 over all the time points of event e_2. If event e_1 corresponds to a seizure, we can average these maximum co-state frequencies over the set $\boldsymbol{B}$ of all bursts to get the vector denoting the average similarity of each time point in seizure e_1 to those of all the bursts, which allows for visualization of seizures across all bursts.

$$\overline{\boldsymbol{m}}^{(e_1)} = \frac{1}{|\boldsymbol{B}|} \sum_{b\, \epsilon\, \boldsymbol{B}} \boldsymbol{m}^{(e_1,b)}.$$