Hidden Markov Models: A Breakdown of Random Matrix Universality & The Brain Criticality Hypothesis

Ryerson University
Dept. of Physics

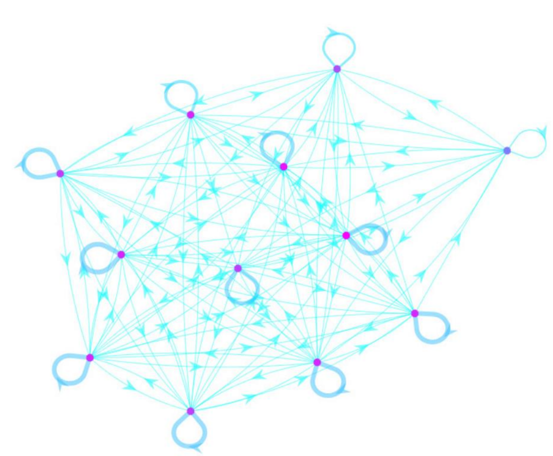
Faheem Mosam & Dr. Eric DeGiuli
Ryerson University

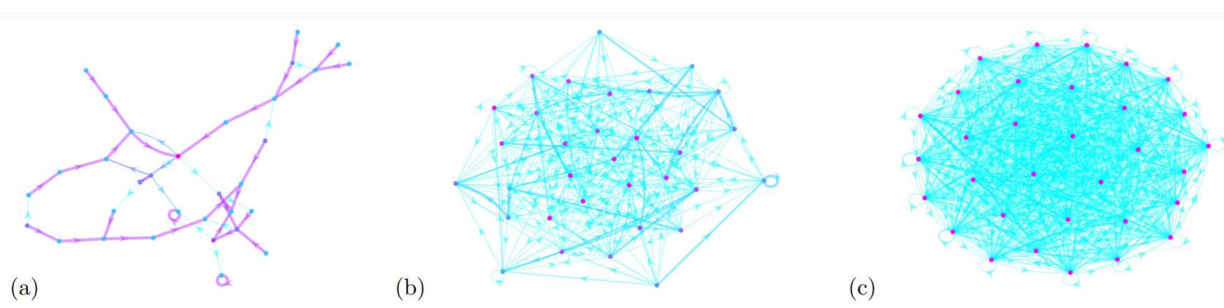
INTRODUCTION AND BACKGROUND

Biological systems need to react to stimuli over a broad spectrum of timescales, how these timescales can emerge without external fine-tuning is still a puzzle. We consider discrete Markovian systems which are governed by transition matrices, allowing us to leverage results from random matrix theory. An ensemble of transition matrices is considered, and we introduce and motivate a temperature-like parameter that controls the dynamic range of matrix elements. We show that a phase transition from full to sparse occurs whereby random matrix theory results breakdown. This phase transition signifies the emergence of a nontrivial Shannon entropy and is accompanied by a peak in complexity as measured by predictive information. The results are then applied to fMRI data of human subjects at wakeful rest and show that brain activity lies close to the phase transition when engaged in unconstrained, task-free cognition.

Hidden Markov Models (HMM)

A HMM relates a set containing N hidden states $\{x_N\}$ with a set containing T observable states $\{o_T\}$. This is done via a transition matrix M of size $N \times N$ with elements m_{ij} = P(x_{t+1} = j|x_t = i). Naturally the elements of M are normalized such that $\sum_{j=1}^{N} m_{ij} = 1$ and $m_{ij} \geq 0$. The out degree is the number of transitions to nodes out of a given node.





Illustrative networks at with N = 32 at temperatures (a) $\epsilon/\epsilon_c = 10^{-2.4}$; (b) $\epsilon/\epsilon_c = 10^{-0.6}$; (c) $\epsilon/\epsilon_c = 10^{0.6}$. Edges are defined for matrix elements $M_{ab} > m_o = 10^{-3}$and are coloured based on their weight. Nodes are coloured based on their out-degree.

Brain Criticality Hypothesis

HMM have shown success in modelling a variety of biological systems including brain network dynamics [1]. Similar to how a pile of sand will reach a critical slope at which an avalanche will occur, brains are hypothesized to self organize into a quasicritical region around a critical state [2]. Most research on the critical brain hypothesis has occurred at the neuronal level, so it would be interesting to see if brains express such criticality at the scale of whole brain network dynamics. In subcritical and supercritical phases, correlation lengths are finite. But if a brain has to react to a wide variety of stimuli on very short time scales, then it needs to have a large susceptibility to external stimulus; ie: a large correlation length. For these reasons, the brain is hypothesized to be at or near a critical state.

Key Parameters

Sparsity, Temperature, and Structure

- The matrix for a HMM can be expressed in terms of a more primitive matrix: $M_{ab} = \frac{Q_{ab}}{\sum_c Q_{ac}}$ such that a uniform matrix $\bar{Q}$ has elements equal to $\frac{1}{N}$. The sparsity can be defined then as: $s(Q) = \frac{1}{N^2}\sum_{a,b} log^2(\frac{Q_{ab}}{\bar{Q}})$
- A temperature parameter ϵ can be defined in terms of sparsity via: $\langle s \rangle = \frac{1}{2\epsilon}$ showing that 'cold' matrices will have large sparsity whereas 'hot' matrices will have very little sparsity.

Shannon Entropy (condense and discuss entropy rate at end)

- Shannon entropy is a quantity H that measures the amount of information that a variable can contain. When H=0 there is complete certainty of the outcome, whereas for a given N from a matrix, H is maximized at log(N) indicating that there is a uniform distribution of probabilities for a result.[3]
- Define the entropy for a sequence of a sequence of length t as $H(t) = H_\pi + (t-1)H_d$; then at large times H(t) is dominated by the entropy rate, shown as: $H_d = \lim_{t\to\infty} \frac{1}{t} H(t)$ where H_π is the entropy of the stationary distribution (for simplicity we only looked at ergodic samples)

Predictive Information

- Bialek, Nemenman, and Tishby [cite] formalized a notion from Grassberger [cite] that complexity can be measured by asking how quickly the entropy rate H_d reaches its asymptotic value through the introduction of predictive information.
- Predictive information can be measured through the following equation: $I_{pred}(t) = \frac{H_\pi - H_d}{\log 2}$ for discrete stationary Markov processes
- Predictive information can be understood as how well the past t states predict the entire future trajectory of the system

Relaxation Times

- Since $\lambda^t = e^{t\log(\lambda)}$, the relaxation times of a discrete stationary Markov process can be calculated via: $\tau = \frac{-1}{\log|\lambda|}$

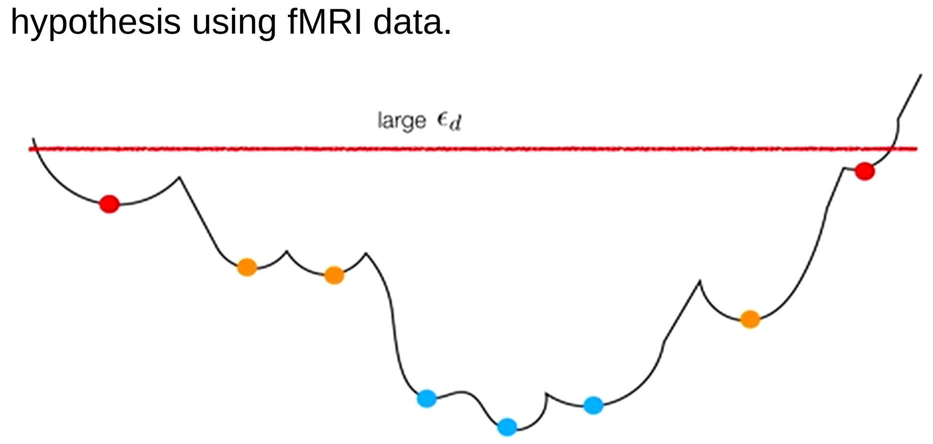
HYPOTHESIS AND SPECIFIC AIMS/OBJECTIVES

Model systems in a way that allows identification of the phase transition from noisy behaviour into an ordered phase. Moreover, show that this phase transition identifies a point of criticality in dynamical HMM systems.

Starting with M_{ab} which has universal features in the large N limit, we consider Q_{ab} that are identically and independently distributed with bounded density, mean $\mu = \bar{M}e^{N^2/4\epsilon}$ & finite variance $\sigma^2 = \bar{M}e^{\frac{N^2}{2\epsilon}}(e^{\frac{N^2}{2\epsilon}} - 1)$. Where $\bar{M}$ is irrelevant, as $N \to \infty$ the spectrum of M_{ab} converges to the uniform law on the disk $|\lambda| < \lambda_c$ in the complex plane, where $\lambda_c = \frac{\sigma}{\mu\sqrt{N}}$. The normalized spectrum for large N then allows us to predict a transition at $\epsilon \log(N+1)/N^2 = \frac{1}{2}$.

The predicted transition represents a point of criticality for dynamical HMM systems, by using this first-principles theoretical prediction we show a simple scenario for the emergence of long time-scales in discrete Markovian systems, by varying the dynamic range of matrix elements. The results are then used to test the brain criticality hypothesis using fMRI data.

By analogy, consider a particle with a large energy traversing a landscape with many minima, as temperature drops, it will eventually fall to rest in some particular minima.



MATERIALS AND METHODS

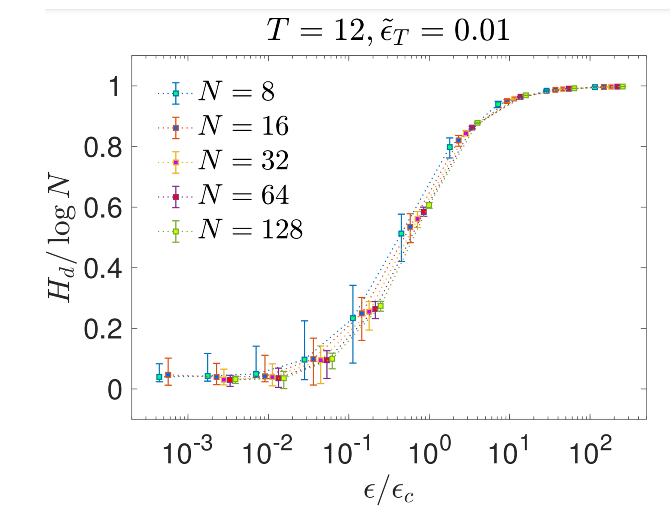
All code for this project is written in MATLAB, split between two main sections: one for the random HMM (RHMM) and one for the human data analysis (NHMM). The RHMM code is adapted from a previous project based off of DeGiuli(2019) [5]. The NHMM code was written from scratch with the exception of the generation of plots which used sections from the RHMM code.

The RHMM code consisted of two sets of data, each with 5 different sized matrices and 10 different temperature parameters. The first set had 300 replicas per matrix size and temperature with each replica being sampled 100 times with sequences of maximum length 4000. The second set of data had 1000 replicas per matrix size and no sampling occurred.

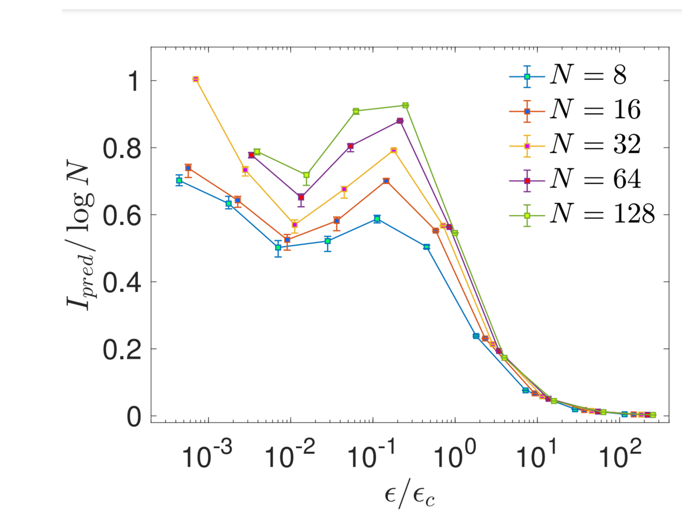
The NHMM code used data provided by Vidaurre [1]. It consisted of 820 subjects which each sat for four sessions for 1200 time steps each. The data was processed accordingly such that transition matrices could be built and measured.

RESULTS

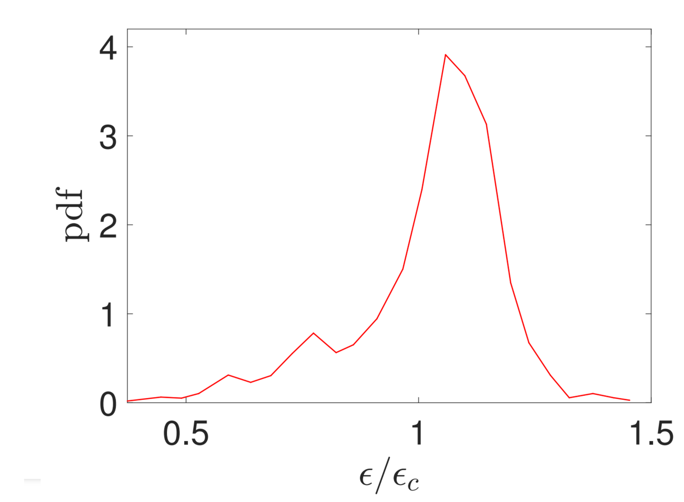
(1) Shannon Entropy rate (RHMM)



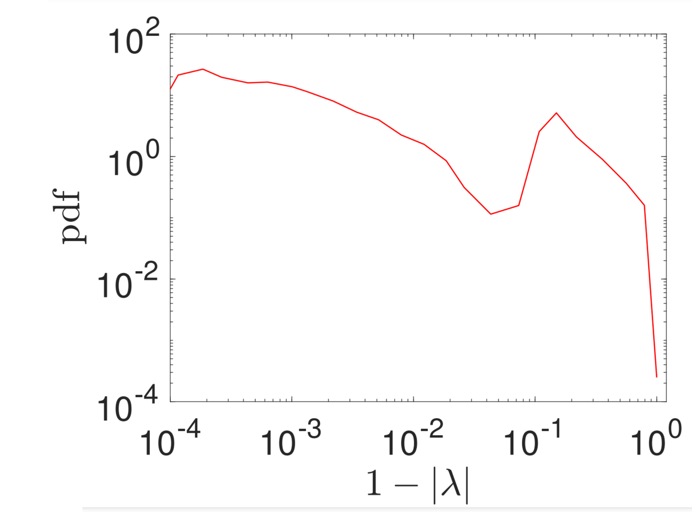
(2) Predictive Information (RHMM)



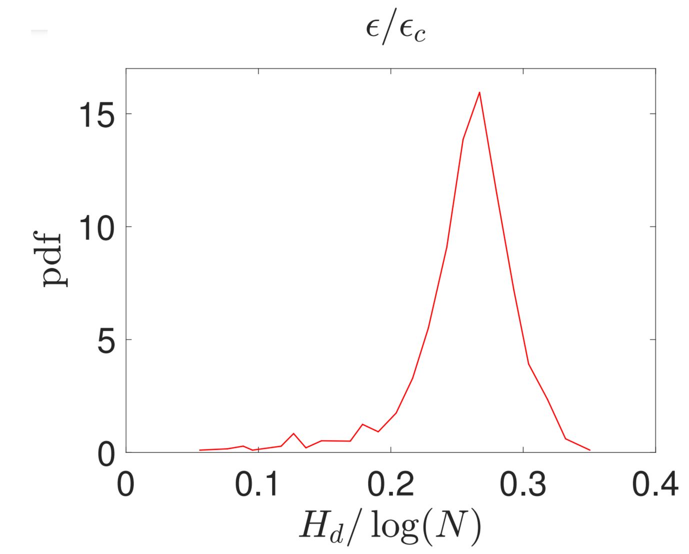
(3) Distribution of Temperatures (NHMM)



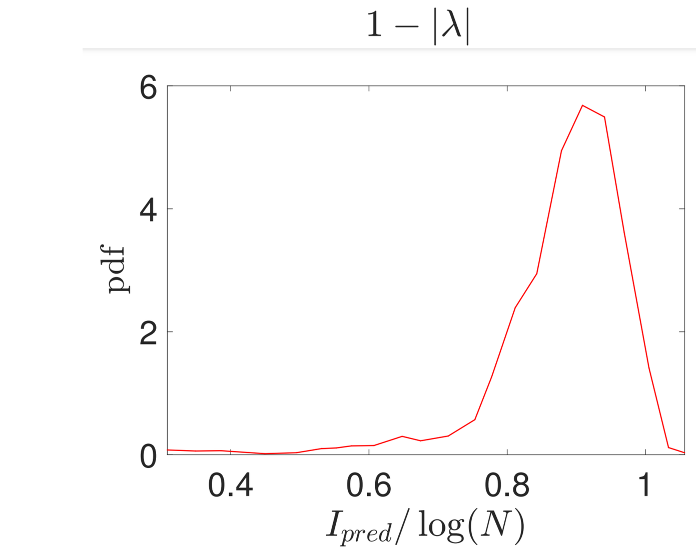
(4) Distribution of Eigenvalues (NHMM)



(5) Shannon Entropy rate (NHMM)



(6) Predictive Information (NHMM)



DISCUSSION AND CONCLUSIONS

- Figure 1 shows the Shannon entropy rate of RHMM data normalized by its maximum value log(N) plotted against the temperature parameter ϵ which is normalized by it's critical value. The normalized entropy rate tends to unity for large temperature, indicating that the transition sequences are indistinguishable from random noise, while low temperatures are nearly equal to 0, indicating that sequences are nearly deterministic
- Figure 2 plots the predictive information of RHMM data normalized by log(N), it peaks at an intermediate $\epsilon \sim 0.1\epsilon_c$, showing that high and low-temperature phases are separated by a phase transition.
- Figure 3 shows the distribution of temperatures for NHMM data, the mean value is indeed very near ϵ_c, a striking result given its status as a phase transition.
- Figure 4 is a probability distribution function of $1 - |\lambda|$ for the human data, which sheds light on the relaxation times. The distribution shows a a distribution of short and long relaxation times, with more values leaning towards longer relaxation times.
- Figure 5 shows the probability distribution of the normalized Shannon entropy rate, while the peak looks far off from the expected value based on figure 1, when taking into account the self-transitions, the expected value for Shannon entropy rate becomes $H_d/logN \approx 0.26$ which is exactly what is found
- Figure 6 shows the predictive information for human data, again the peak can be understood by taking into account the self transitions, leading to an expected value that is as follows: $I_{pred}/logN \approx 0.89$ which is remarkably close to the measured mean value of 0.85
- The phase transition discussed here can be interpreted as a full-to-sparse transition of random matrices. Applied to the fMRI data, we found that human data lies very near the transition, supporting the brain criticality hypothesis.
- Since the ensemble includes all discrete Markov models, it can unify the study of disparate systems, in the goals of seeking universal patterns. This may shed light on the origin and possible universality of criticality in biological systems

REFERENCES

1. S. M. Smith D. Vidaurre and M. W. Woolrich. Brain network dynamics are hierarchically organized in time . PNAS, 114(48), 2017.
2. Leandro J. Fosque, Rashid V. Williams-Garc´ıa, John M. Beggs, and Gerardo Ortiz. Evidence for quasicritical brain dynamics. Phys. Rev. Lett., 126:098101, Mar 2021.
3. C.E.Shannon. A mathematical theory of communication. The Bell System Technical Journal, 27:379–423, October 1948.
4. W. Bialek, I. Nemenman, and N. Tishby. Predictability, complexity, and learning. Neural computation, 13(11):2409–2463, 2001.
5. Eric DeGiuli. Random language model. Physical Review Letters, 122, March 2019

ACKNOWLEDGEMENTS

EDG acknowledges support of the Natural Sciences and Engineering Research Council of Canada (NSERC), Discovery Grant RGPIN-2020-04762.