Statistical inference in theoretical models of cognition

Sean, Agostina, Alex, Chunyu, Francesca, Srdjan, Carlos, Ken, and John

1 Abstract

Theoretical neuroscientists often design circuit models of neural activity with the hopes of producing some emergent property observed in data. However, the standard inference toolkit is designed to condition on data points collected in an experiment with a phenomenological observational model, rather than the abstract notion of an emergent property in a biologically realistic model. We introduce a novel machine learning methodology called degenerate solution networks (DSNs), which learn a distribution of neural circuit model parameterizations that produces the emergent property of interest, and is otherwise as random as possible. We use DSNs to advance theoretical understanding of the stomatogastric ganglion (STG), primary visual cortex (V1), superior colliculus (SC), and low rank recurrent neural networks (RNNs). As models continue to increase in complexity and become less tractable in terms of conventional analytic and theoretical approaches, DSNs will be a useful tool for theorists to interrogate their models

2 Introduction

Introduce theoretical neuroscience, the gold standard

- Theoretical modeling is/has been rising in neuroscience [1].

- Define theory of neural computation

- Explain idealized practice of theory – analytic derivation.

- Provide historical examples of the idealized approach: [2], [3], and [4].

- *Transition*: Unfortunately, as biological realism is introduced into neural circuit models, theory through analytic derivation becomes infeasible.

The central challenge of theoretical neuroscience

- In fact, neural circuit models are designed in the context of misaligned, competing incentives.

  - Simplicity (e.g. symmetry, gaussianity), help for analytic derivations.

– Complexity gives biological realism at the cost analytic tractability.

- When biological realism matters and analysis becomes intractable, standard practice is to examine simulated activity [5] (cite a bunch here).

- *Transition* Visualizations and regression are used to provide insights on how parameters govern system activity, however theorists strive for a more formalized understanding of these complex systems.

Careful inference vs careful models

- Bayesian inference is a candidate for formalizing this mapping using probability.

- Introduce modern inference engine. It's use in neuroscience.

- Such careful inference and practical modeling is juxtaposed with the standard practices of the theoretical neuroscience where...

- Can we start doing *careful* inference in *careful* models?

Introduce approach/method and summarize contribution

- Here, we use the state-of-the-art Bayesian inference engines, to learn formal proababilistic relationships between carefully modeled neural circuit parameters and their emergent properties.

- We present a novel machine learning method, DSNs, which ...

- In this study, we use DSNs to provide novel theoretical insights using biological models of list of biological models along with emergent properties studied.

Figure 1 We should have a figure to accompany the introduction section. People seem to like the practice in neural data analysis vs theoretical neuroscience comparison. Maybe put that in Figure 1 along with icons for the four modelsand emergent properties we'll be analyzing?

3   Results

3.1   Degenerate solution networks

- To translate progress in neural data analysis to theoretical neuro, need to key steps.

- 1. Need to learn parameter distributions of biologically realistic (not just phenom.) models.

  - 2. Must be able to condition on emergent properties of interest, not simply computationally convenient sufficient statistics of data sets.

- Bayesian data scientists will say experimental data is all that matters.

- *Transition*: This is untrue when working in a creative, exploratory modeling setting.

Edgy contrarian point about theorists and data

- Common misconception: theoreticians rarely attempt to directly reproduce experimental data.

- Instead, they work with (abstracted?) mathematical definitions of emergent properties.

DSNs

- We introduce DSNs, which bridge methodology in these subfields of comp neuro.

- Combine ideas from MEFNs (cite Gabe) and LFVI (cite Dustin) to learn a deep probability distribution of theoretical model parameterizations z that produce the emergent properties of interest $T(x)$ (see Appendix).

- Explain deep probability distributions.

- DSNs are deep probability distributions of theoretical model parameters, which are optimized to be maximally random (maximum entropy) while producing the specified value of emergent properties:

$$q_\theta^*(z) = \underset{q_\theta \in Q}{\operatorname{argmax}} H(q_\theta(z))$$
$$\text{s.t. } E_{z \sim q_\theta} \left[E_{x \sim p(x|z)} \left[T(x) \right] \right] = \mu \tag{1}$$

Worked example: STG

- For example, consider the STG.

- Explain this STG circuit, emergent property of interest.

- For our choices of STG as model and network syncing as emergent property, we use a DSN to learn a distribution on STG conductance parameters that produces network syncing.

- Emphasize utility of DSN using Hessian.

- An equivalent conceptualization is that DSNs do Bayesian inference (see Appendix).

- Punchline about DSNs and transition to V1.

3.2 Exploratory analysis of a theoretical model

Will focus on this once result finalized. Have a lot of text to pull from.

3.3 Identifying sufficient mechanisms of task learning

Will focus on this once V1 and LRRNN finalized. Have a lot of text to pull from.

3.4 Conditioning on computation with interpretable models of RNNs

Will focus on this once result finalized. Have a lot of text to pull from.

4 Discussion

- Summarize the key methodlogical demonstrations from the results section.

- Talk big picture: If we know we can't analytically derive these things, we need an alternative characterization. Simulate and examine isn't cutting it. We need to be leveraging the modern inference engine to gain this understanding. Bayesian probability is the framework we should use for this formalism.

- Expand on idea of posterior predictive checks / hypothesis testing / exploratory analyses of models themselves. Give the whole, we don't even understand the models we're developing pitch.

- Elaborate on idea of conditioning on flexibly defined statistics i.e. emergent properties. Emphasize how this is practical. Link to sufficient statistics, esp. commonly used in phenom models like spike counts etc.

- Summarize the respective strengths SNPE and DSN.

- Link conditioning on task execution with work done today with RNNs. Basically, we're training overparameterized models with regression, and get a distribution (we have no prob treatment of). Emphasize utility of low-dim interpretable parameterizations.

- A paragraph on bridging large scale recordings with theory.

References

[1] Larry F Abbott. Theoretical neuroscience rising. *Neuron*, 60(3):489–495, 2008.

[2] John J Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the national academy of sciences*, 81(10):3088–3092, 1984.

[3] Haim Sompolinsky, Andrea Crisanti, and Hans-Jurgen Sommers. Chaos in random neural networks. *Physical review letters*, 61(3):259, 1988.

[4] Misha V Tsodyks, William E Skaggs, Terrence J Sejnowski, and Bruce L McNaughton. Paradoxical effects of external modulation of inhibitory interneurons. *Journal of neuroscience*, 17(11):4382–4388, 1997.

[5] Gabrielle J Gutierrez, Timothy OLeary, and Eve Marder. Multiple mechanisms switch an electrically coupled, synaptically inhibited neuron between competing rhythmic oscillators. *Neuron*, 77(5):845–858, 2013.

[6] Brendan K Murphy and Kenneth D Miller. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron*, 61(4):635–648, 2009.

[7] Hirofumi Ozeki, Ian M Finn, Evan S Schaffer, Kenneth D Miller, and David Ferster. Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron*, 62(4):578–592, 2009.

[8] Daniel B Rubin, Stephen D Van Hooser, and Kenneth D Miller. The stabilized supralinear network: a unifying circuit motif underlying multi-input integration in sensory cortex. *Neuron*, 85(2):402–417, 2015.

[9] Henry Markram, Maria Toledo-Rodriguez, Yun Wang, Anirudh Gupta, Gilad Silberberg, and Caizhi Wu. Interneurons of the neocortical inhibitory system. *Nature reviews neuroscience*, 5(10):793, 2004.

[10] Bernardo Rudy, Gordon Fishell, SooHyun Lee, and Jens Hjerling-Leffler. Three groups of interneurons account for nearly 100% of neocortical gabaergic neurons. *Developmental neurobiology*, 71(1):45–61, 2011.

[11] Ashok Litwin-Kumar, Robert Rosenbaum, and Brent Doiron. Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes. *Journal of neurophysiology*, 115(3):1399–1409, 2016.

[12] Carsten K Pfeffer, Mingshan Xue, Miao He, Z Josh Huang, and Massimo Scanziani. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nature neuroscience*, 16(8):1068, 2013.

[13] Mario Dipoppa, Adam Ranson, Michael Krumin, Marius Pachitariu, Matteo Carandini, and Kenneth D Harris. Vision and locomotion shape the interactions between neuron types in mouse visual cortex. *Neuron*, 98(3):602–615, 2018.

[14] Chunyu A Duan, Marino Pagan, Alex T Piet, Charles D Kopec, Athena Akrami, Alexander J Riordan, Jeffrey C Erlich, and Carlos D Brody. Collicular circuits for flexible sensorimotor routing. *bioRxiv*, page 245613, 2018.