Causal inference in neural circuits with open- and closed-loop control

Adam Willats and Matt O'Shaughnessy

April 23, 2021

Abstract

A primary goal of neuroscience is identifying causal relationships between (regions / circuits) of the brain, a task that is challenging when only observational (without intervention?) data is available despite being able to record from an increasing proportion of the brain [cite neuropixels]. Interventions that stimulate parts of the brain, such as lesioning and optogenetics, have been proposed to partially overcome this difficulty. These open-loop interventions, however, do not fully eliminate confounding, limiting their ability to reveal casual structure in some circuits. This is especially problematic when circuits feature reciprocal connections, strong feedback at fast timescales, or when multiple regions are driven by the same upstream cause. In this paper, we show that interventions using closed-loop control can overcome these limitations. We provide a practical framework for applying closed-loop control to circuit identification problems and propose rules that predict when observational data or open-loop interventions are sufficient, and when closed-loop control is needed. Using a wide range of simulated circuits, we then characterize how key process and intervention parameters affect successful identification. This characterization builds towards an understanding of the limits of identification and suggests methods for improving the next generation of circuit identification. Our approach could complement existing technology developments and guide the next generation of (identification in neural circuits).

In this work, we aim to provide a practical framework for applying closed-loop control to circuit identification problems. (The hope is this work) will (open up, further develop) a dialog between those designing and executing systems neuroscience experiments and computational neuroscientists about how to conduct identification experiments moving forward. We aim to start answering questions such as - what additional value does closed-loop control offer? - in which circuits is closed-loop control necessary? in which circuits is open-loop control sufficient? What are the requirements for spatial and temporal degrees of freedom to meet our identification goals? - what can be (said about) the connections identified from such experiments? - How can aspects (parameters?) of an experiment be designed in order to strengthen (and make more data-efficient) our (hypothesis-testing power, the conclusions about circuits)

Overall, this paper summarizes the intersection of causal inference, neuroscience, and control theory, to highlight why and where intervention facilitates circuit identification. We then demonstrate how these principles apply to a simple case study of a 3-node circuit. Then we show a broad characterization of how several key process and intervention parameters affect successful identification. This characterization builds towards and understanding of the (limits of identification) and (begins to suggest) ways to (make the most of the next generation of circuit identification)

This is important because statements of causality are what we mean fundamentally when we talk about how the brain works. Statements of causality are also what we *need* when seeking to develop new therapies to treat disorders of the brain.

make this sound more neuro-legit

challenging → impossible?

more neuro here

add something about 3-node circuit?

mention interventional budget?

sounds condescending

1 Introduction

1.1 Closed-loop control in neuroscience

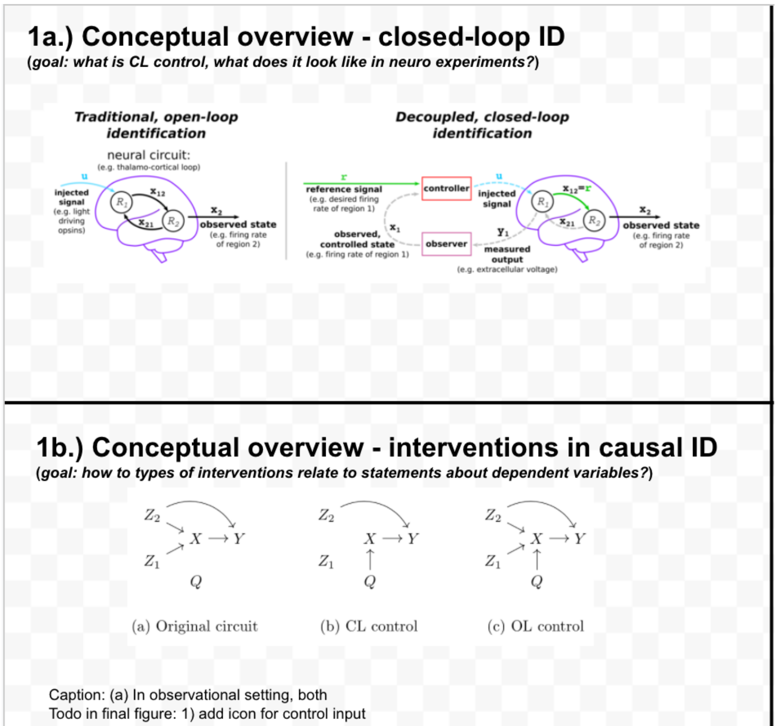

Figure 1: **Text** more text

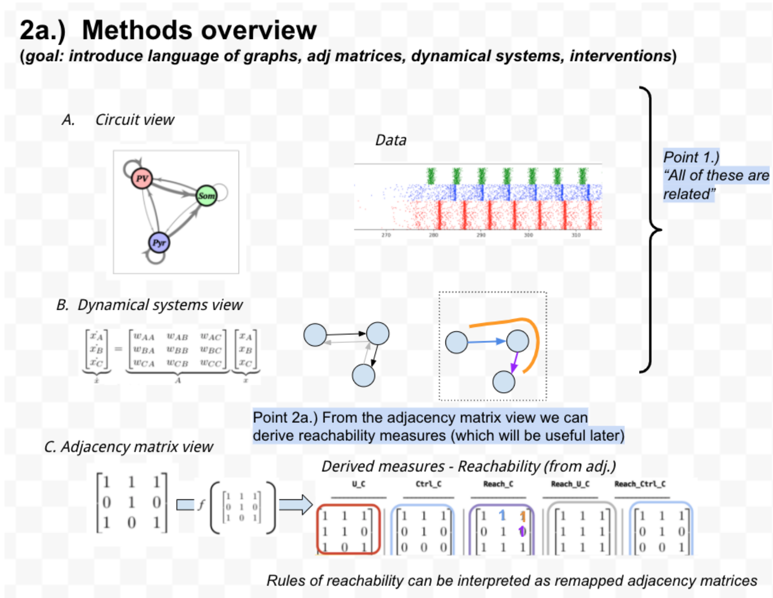

Figure 2: **Text** more text

1.2 Interventions from the perspective of causal inference

2 Methods

2.1 Mathematical models of circuits in the brain

2.1.1 The circuit view

2.1.2 The dynamical system view

2.1.3 The connectivity (adjacency matrix) view

2.2 Interventions in causal identification

2.3 Understanding identification through derived properties of circuits (reachability rules)

3 Results

3.1 Applying (framework) to distinguish a pair of circuits

3.2 Characterizing circuit-pair ambiguity assessed through reachability properties (across circuit size)

3.3 Characterizing identifiability in spiking neural networks

3.3.1 Methods - pipeline for inferring circuits with intervention

relevant properties of spiking circuits within-cell properties

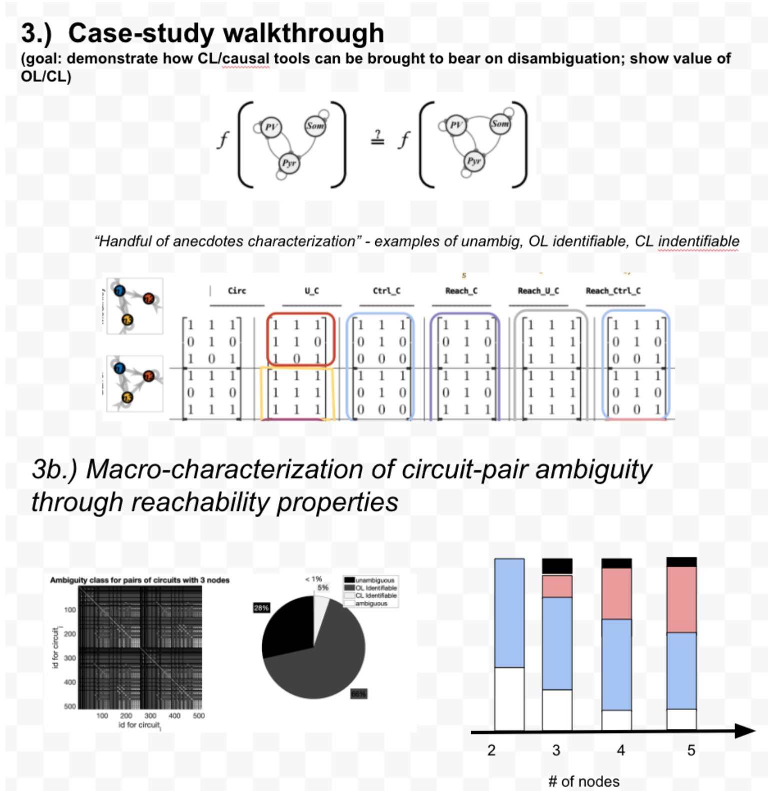

Figure 3: **Text** more text

- membrane time constants
- internal noise

properties of conenctions

- connection strength (weights)
- synaptic delay

properties across a circuit

- heterogeneity of (above properties) across neurons
- proportion of excitatory vs inhibitory connections
- *(spiking vs lfp-like observations)*

intervention parameters

- passive observation v.s. open-loop v.s. closed-loop control

- % of the network observed
- % of the network intervened
- number of samples recorded
- controller bandwidth
- *(intervention location - stimulating driver versus "dead-end" nodes)*
- *(impact of imperfect control)*

Methods for extracting circuit estimates Quantifying successful identification

- false positives/negatives in identified links
- link-identification accuracy / F1 score

Results - open-loop intervention facilitates identification with less data than passive observation

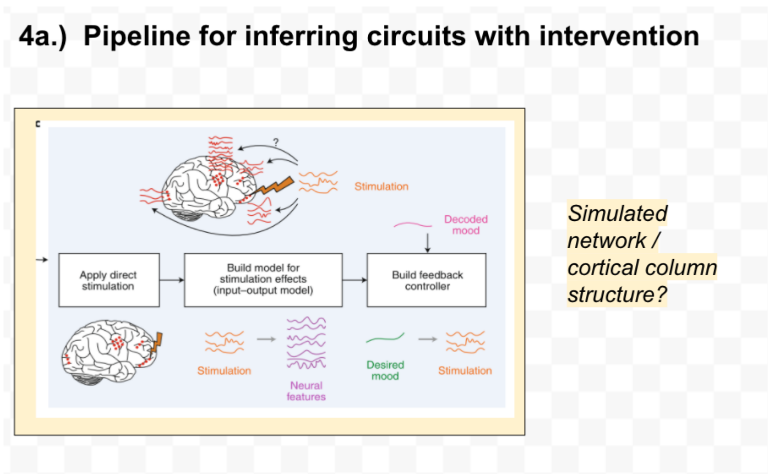

Figure 4: **Text** more text

4 Discussion

4.1 Related work

4.2 Challenges to identification, a spectrum of interventions

4.3 Limitations

4.4 Future Work

References

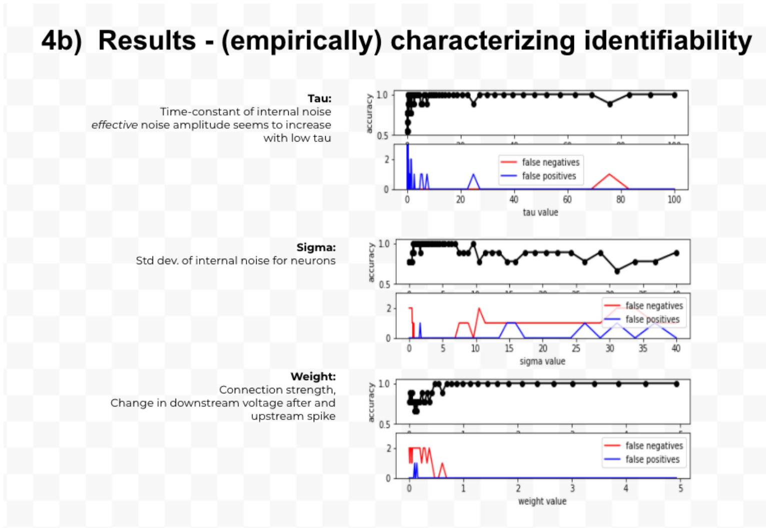

Figure 5: **Text** more text

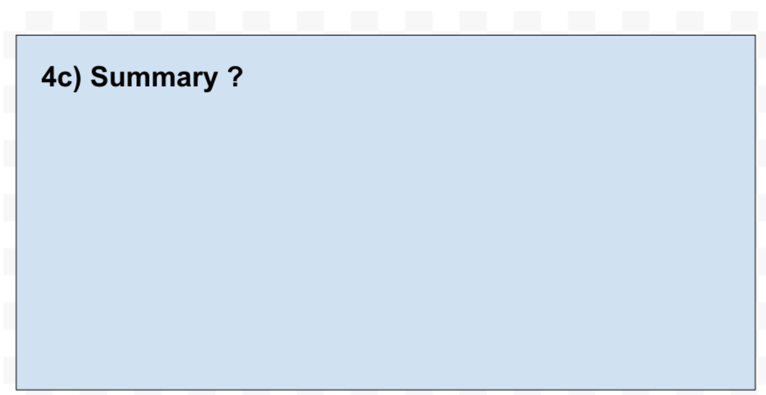

Figure 6: **Text** more text