Model of Hallucination using Bayesian Inference and Dopamine-Glutamate Interactions

Nikhil Mukraj

Abstract

Various computational models of schizophrenia and related disorders utilize auto-associative networks to model the phenomenon of memory and hallucination. These models using attractor states have proven accurate for modeling the effect of varying ion channel conductances on memory recall and hallucination. This paper demonstrates the efficacy of these models in a Bayesian inference architecture while varying dopaminergic and glutamatergic interactions. It also shows the potential of these models to highlight potential drug targets such as glutamate and dopamine transporter modulation.

1 Introduction

Schizophrenia and related psychotic disorders are characterized by positive symptoms such as hallucinations and delusions. These are generally accepted to be modulated by both neurochemical and environmental conditions. While the dopamine hypothesis of schizophrenia does explain the action of antipsychotics, it does not explain the phenomena of hallucinations completely. Computational psychiatry elaborates that hallucinations can be explained through a combination of attractor states and Bayesian inference.

Assuming the brain can operate as an attractor, as demonstrated by previous research, hallucinations can be explained as a failure to maintain a low firing resting state [1, 2, 3]. Failure to maintain an attractor state can be explained by an imbalance in excitation and inhibition. If a neural network has too much excitatory activity or too little inhibition, an attractor can erroneously jump to a firing state recalling a memory erroneously. This imbalance can be generated through atypicalities in AMPA, NMDA, and GABAa conductances through simulation but is also supported by experimental evidence.

Bayesian inference can explain hallucination as prior cues having too much influence over current cognition. While having some influence over perception of a current cue is necessary for cognition, too much influence can cause erroneous calculation [4]. Various studies have shown that Bayesian inference can explain clinical and experimental data but there appears to be a gap in terms of studying Bayesian inference by computationally simulating neurochemical changes and observing the results at a systems level.

Simulation of Bayesian inference can be done using auto-associative attractor models. Simulations of auto-associative attractor neural networks can utilize biophysical neuronal models to demonstrate systems level effects from biophysical changes, namely changes present due to ion channel activity as well as connectome changes. The aim of this paper is to use those simulations to study, computationally, the effect of modifying different potential drug targets in a Bayesian inference setting on hallucinations.

2 Methodology

2.1 Models

2.1.1 Izhikevich Equations

There are various biophysical models that can be used to simulate systems level changes in neurobiology, including but not limited to Hodgkin-Huxley models, Morris-Lecar models, and leaky integrate-and-fire models. An Izhikevich neuron model was used to due its accuracy in describing membrane potential while being computationally efficient in terms of speed.

The following equations describe a modified

Izhikevich neuron that can model dopaminergic pathways that modify AMPA and NMDA action (Potentially will be modified such that $c_{\text{glutamate}}$ will be modified by D_1 or D_2) (Currently s_{D_1} and s_{D_2} are 1 and 0.05 respectively):

$$C_m \frac{dv}{dt} = 0.04v^2 + 5v + 140 - u + I_{\text{syn}}$$

$$\tau_m \frac{du}{dt} = a(bv - u)$$

2.1.2 Receptor Models

The receptor gating value of each receptor, (AMPA, NMDA, D1, and D2) was described by the following equation:

$$r_x = -c_y t_y dt + H(v_{\text{th}}) t_{y_{\max}}$$

Where $H(x)$ refers to the Heaviside function, $-c_y$ refers to the clearance constant of the associated neurotransmitter acting on the receptor, t_y refers to the concentration of the associated neurotransmitter, and $t_{y_{\max}}$ is the maximal concentration of the neurotransmitter. v_{th} refers to voltage threshold of the Izhikevich model, or when the neuron spikes. r_x should remain within the domain $1 \geq r_x \geq 0$. Clearance constants were chosen to reflect physiological conditions [5, 6].

Dopamine is able to modulate the efficacy of other receptors, particularly AMPA and NMDA [7, 8, 9]. This dopaminergic gain modulation is modeled by two receptors, D_1 and D_2, as shown by the following equations:

$$M_{D_1} = 1 - r_{D_1} s_{D_1}$$
$$M_{D_2} = 1 - r_{D_2} s_{D_2}$$

These variables are used to modify the conductances of the AMPA and NMDA channels in an excitatory manner for D_1 and an inhibitory manner for D_2.

It should be noted that the effect of D_1 tapers off as the concentration of dopamine increases whereas the effect of D_2 is linear @tseng2004dopamineglutamate. AMPA current is modeled as follows:

$$I_{\text{AMPA}} = g_{\text{AMPA}} M_{D_2} r_{\text{AMPA}} (v - E_{\text{AMPA}})$$

The M_{D_2} modulation factor decreases the conductance of AMPA as the concentration of dopamine rises. NMDA is modeled as follows:

$$B(v) = \frac{1}{1 + \frac{e^{-0.062v}[\text{Mg}^{2+}]}{3.75}}$$

$$I_{\text{NMDA}} = g_{\text{NMDA}} M_{D_2} r_{\text{NMDA}}^{M_{D_1}} B(v)(v - E_{\text{NMDA}})$$

Where M_{D_2} decreases the efficacy of NMDA while M_{D_1} increases the efficacy. The $[Mg^{2+}]$ was chosen to be 0.3 mM to reflect physiological conditions. Finally GABAa was modeled as follows:

$$I_{\text{GABAa}} = g_{\text{GABAa}} r_{\text{GABAa}} (v - E_{\text{GABAa}})$$

No dopaminergic effect was modeled on GABAa specifically.

The input to the neuron model was modeled as the sum of the ligand gated channel receptor currents:

$$I_{\text{syn}} = I_{\text{AMPA}} + I_{\text{NMDA}} + I_{\text{GABAa}}$$

2.2 Model Architecture

The auto-associative point attractor was split into two groups: the excitatory group (E_1) and the inhibitory group (I_1). The excitatory group expresses the recalled pattern while the inhibitory group ensures stability. The each excitatory group had weights structured to store a set of random binary patterns. In each simulation the excitatory group received input from two separate cues, either through a glutamate input or a dopamine input. The network was setup such that the secondary glutamate or dopamine input was from two coupled groups, one excitatory (E_2) and one inhibitory (I_2). The input itself was taken from E_2. There were three types of potential neurotransmitter action: glutamate, D_1, and D_2.

When glutamate was used as a secondary cue, E_1 was directly connected to the E_2. One specific binary pattern stored in E_2 was chosen at random to be connected to another specific binary pattern stored in E_1. Every on bit in the E_2 pattern was mapped in a one-to-one manner to an on bit in E_1.

When dopamine was used to influence E_1, there was an intermediate dopamine group (D) between E_2 and E_1 so there were no internal dopaminergic synapses in E_2, but dopamine could be transmitted to E_2. If D_1 was simulated, E_1 was modified to have D_1 receptors ($s_{D_1} > 0$), and if D_2 was simulated, E_1 was modified to have D_2 receptors ($s_{D_2} > 0$). However, if s_{D_2} was simulated, the pattern in E_2 mapped on bits to off bits in E_1 since D_2 action was inhibitory. The model first applied C_2 and then C_1 to ensure that C_2 was a prior cue. The simulations tested the preference for a certain cue under different distortion conditions; a healthy control should see that C_1 takes priority when distortion is

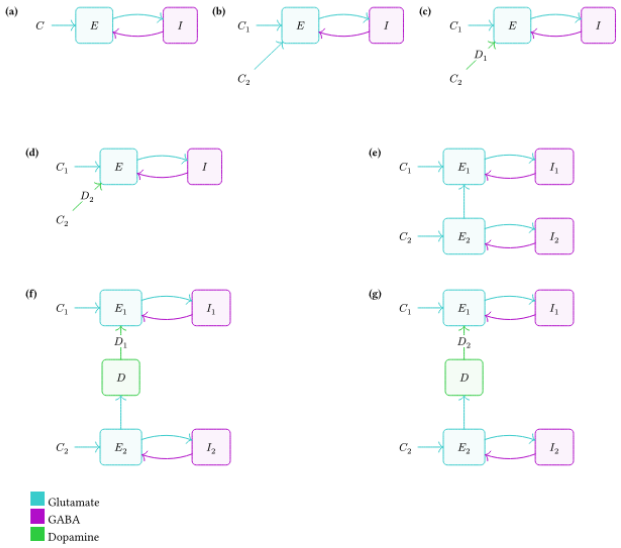

Figure 1: Model structure

low but C_2 takes priority when distortion is high. Schizophrenic behavior should see a higher tendency to tend towards representing C_2 as Bayesian inference is deregulated in a manner that results in prior cues having a larger effect.

As a control, E_1 coupled to E_2 was simulated with no Bayesian inference to demonstrate how hallucinations could form independently of inference.

quick notes: there should be four simulations control with no bayesian inference where cue is applied immediately to E_1 control with no bayesian inference where no cue is applied (if when no cue is applied, memory never holds, try using a small noisy cue) one where a cue is applied immediately directly to E_1 and then removed one where no cue is applied to E_1 but E_2 inputs a cue pattern switching simulations too

D2 SHOULD BE FIT TO DOPAMINE-GLUTAMATE PAPER should have svg figure somewhere

maybe analyze manifolds too

3 Results and Discussion

3.1 Control with No Bayesian Inference

3.1.1 Cue Applied

3.1.2 No Cue Applied

3.2 Cues Applied to Same Group

3.2.1 Glutamate Based Inference

3.2.2 Dopamine Based Inference

3.3 Inter-group Cues

3.3.1 Glutamate Based Inference

3.3.2 Dopamine Based Inference

Interaction	$\Pr(>F)$
C (distortion):C ($\mathbf{AMPA_g}$)	2.36×10^{-7}
C (distortion):C ($\mathbf{NMDA_g}$)	6.06×10^{-1}
C (distortion):C ($\mathbf{GABAa_g}$)	2.26×10^{-7}

The D1 stimulation demonstrates that the AMPA, NMDA, and GABAa conductances, along with distortion, have a statistically significant effect on Bayesian accuracy. This is likely due to C_2 attractor basins deepening with higher AMPA and lower GABAa conductances by stronger dopamine synapses either directly or indirectly. NMDA does not appear to deepen the basins but instead seems to generally increase the bias towards C_2 as the prior

cue. For 20% distortion when varying AMPA, increasing AMPA conductance from 0.3 to 0.9 shows a 21.4% increase in hallucinating the C_2 to be the correct cue. Conversely, for 80% distortion, increasing AMPA conductance from 0.3 to 0.9 shows a 13.5% decrease in accuracy. Increasing AMPA conductance seems to generally show an increase in hallucination for more intact cues but a decrease for more distorted cues. This increase in hallucination is likely due to the strengthening of dopamine modulatory synapses between attractors. However, the decrease in hallucination for more distorted cues is likely due to the model settling on an attractor state that is neither C_1 or C_2.

GABAa conductance appears to have the opposite effect of varying AMPA conductance, where lower GABAa conductances seem to show a higher tendency of hallucination at higher distortion levels. GABAa conductance is likely decreasing the attractor basin depth for C_1 and C_2, since the accuracy of each cue becomes more similar as GABAa conductance increases. This is seen clearly given the GABAa results with a distortion of 20% which show a 35.2% decrease in hallucinating C_2, while a distortion of 80% shows an increase of hallucination by 30.6%. The inhibition that GABAa provides appears to be necessary to prevent runaway excitation from dopamine modulation that causes hallucinations. Given stronger perceptual cues, it is likely that AMPA and GABAa conductances will have a lessened effect on the non-Bayesian C_1 cue.

NMDA conductance seems to have a different pattern of affect on hallucination compared to AMPA and GABAa conductances. Higher NMDA conductance biases all distortions towards hallucinating C_2 as the correct cue. For a distortion of 20%, there is a 23.8% increase in hallucinating C_2 when NMDA conductance is increased from 0.5 to 1.5. For a distortion of 80%, there is a 27.0% increase in hallucinating C_2 when NMDA conductance is given the same increase. This is likely due to the nature of NMDA channels which only activate during depolarization and may have a harder time deepening attractor basins.

4 Conclusion

5 quick notes

- Generate the 4×4 grids for each parameter
- Demonstrate that c2 takes over with high enough distortion
- Should have noisy cue/no cue on c1 and c2 input, to trigger hallucinations

References

[1] E. T. Rolls, "Glutamate, obsessive–compulsive disorder, schizophrenia, and the stability of cortical attractor neuronal networks," *Pharmacology Biochemistry and Behavior*, vol. 100, no. 4, pp. 736–751, 2012.

[2] E. T. Rolls, "Cortical neurodynamics, schizophrenia, depression, and obsessive-compulsive disorder," in *Phenomenological Neuropsychiatry*, pp. 119–143, 2024.

[3] C. Rennó-Costa, J. E. Lisman, and P. F. M. J. Verschure, "A signature of attractor dynamics in the ca3 region of the hippocampus," *PLOS Computational Biology*, vol. 10, no. 5, 2014.

[4] J. Siemerkus, S. Tomiello, and K. E. Stephan, "Bayesian inference and hallucinations in schizophrenia," *Brain*, vol. 142, no. 8, pp. 2178–2181, 2019.

[5] R. N. Lippert, A. L. Cremer, S. Edwin Thanarajah, C. Korn, T. Jahans-Price, L. M. Burgeno, M. Tittgemeyer, J. C. Brüning, M. E. Walton, and H. Backes, "Time-dependent assessment of stimulus-evoked regional dopamine release," *Nature Communications*, vol. 10, no. 1, p. 336, 2019.

[6] K. Moussawi, A. Riegel, S. Nair, and P. W. Kalivas, "Extracellular glutamate: functional compartments operate in different concentration ranges," *Frontiers in systems neuroscience*, vol. 5, p. 94, 2011.

[7] E. Hernández-Echeagaray, A. J. Starling, C. Cepeda, and M. S. Levine, "Modulation of ampa currents by d2 dopamine receptors in striatal medium-sized spiny neurons: are dendrites necessary?," *European Journal of Neuroscience*, vol. 19, no. 9, pp. 2455–2463, 2004.

[8] M. J. Higley and B. L. Sabatini, "Competitive regulation of synaptic ca2+ influx by d2

dopamine and a2a adenosine receptors," *Nature neuroscience*, vol. 13, no. 8, pp. 958–966, 2010.

[9] K.-Y. Tseng and P. O'Donnell, "Dopamine–glutamate interactions controlling prefrontal cortical pyramidal cell excitability involve multiple signaling mechanisms," *The Journal of Neuroscience*, vol. 24, no. 22, pp. 5131–5139, 2004.