Excitatory-Inhibitory Dynamics in Adaptive Decision-Making

Veronica Chelu
Mila, McGill University
Montreal, QC, Canada
veronica.chelu@mail.mcgill.ca

Doina Precup
Mila, McGill University, Google DeepMind
Montreal, QC, Canada
doina.precup@mcgill.ca

Abstract

Neural circuits rely on excitatory–inhibitory (E/I) interactions to support adaptive learning and decision-making. Here, we investigate how these dynamics contribute to flexible behaviour across three modelling levels. First, using a simplified mean-field model of two-choice decision-making, we examine the computational role of selective excitation and inhibition in stabilizing or amplifying competition between alternative choices. Building on these insights, we embed a similar E/I mechanism into the preference function of a reinforcement learning (RL) agent, showing how inhibitory feedback modulates behavioural adaptation in reversal learning. Finally, we assess the scalability of these principles by training RL agents with E/I-constrained recurrent neural networks (RNNs) in dynamic tasks. While a general E/I architecture allows broader forms of inhibitory influence, our results indicate it hinders learning in these settings. In contrast, a structured architecture enforcing local inhibition preserves biological plausibility while maintaining robust performance. Together, these findings suggest that E/I dynamics may provide a feasible computational mechanism for adaptive learning and decision-making.

Keywords: excitation-inhibition; neural dynamics; E/I-constrained RNNs; adaptive decision-making; reinforcement learning

Introduction

Adaptive learning and decision-making are critical in dynamic environments. A fundamental organizing principle of neural circuits, Dale's law (Dale, 1935), states neurons release either E or I neurotransmitters (Eccles et al., 1954). Recurrent E/I circuits support adaptive decision-making (Lam et al., 2022; Najafi et al., 2020; Roach et al., 2023).

We investigate the computational role of selective E/I dynamics in adaptive decision-making, focusing on two canonical motifs: excitatory recurrence and inhibitory feedback, across three scales of abstraction[1]: (i) a two-choice E/I mean-field model, (ii) an RL agent with E/I-structured action-preferences, and (iii) two E/I-constrained RNNs.

Methods and Results

A Mean-Field E/I Model of Decision-Making

We use an E/I mean-field model (Wilson & Cowan, 1972) with two decision-selective E and I populations cf. Fig.1-a in a fixed contextual decision-making task with ambiguous input.

Results Increasing recurrent excitation (w_{EE}) shifts dynamics from a single to bi-stable attractors (Fig. 1-b). Inhibition shapes dynamics based on specificity: **non-specific** (no effect), **ipsi-specific** (compresses attractors, increases stability; Fig. 1-c), **contra-specific** (enhances choice competition, may destabilize dynamics; Fig. 1-d).

E/I in Value-Based Control

We integrate E/I dynamics into the action-preferences of an RL agent, with E preferences encoding expected rewards and I integrating reward prediction errors (RPE): $Q_{\mathrm{E}}(A) \leftarrow (1-\alpha)Q_{\mathrm{E}}(A)+\alpha R - w_{\mathrm{I}}Q_{\mathrm{I}}(A)$, and $Q_{\mathrm{I}}(A) \leftarrow (1-\alpha)Q_{\mathrm{I}}(A)+\alpha\delta$, with $\delta = R - Q_{\mathrm{E}}(A)$. Action selection follows entropy-regularized policy updates (Bhandari & Russo, 2021): $\pi'(a) \propto \pi(a)^{1-\tau}\exp\big(w_{\mathrm{E}}\, Q_{\mathrm{E}}(a)\big)$, where τ regulates entropy, whereas the precision w_{E} controls the integration of preferences into behaviour and encapsulates the computational role of excitation without modelling the underlying dynamics explicitly. The specificity of inhibition is controlled through an inhibitory strength w_{I}. We examine performance in a reversal learning task with deterministic reward contingency shifts.

Results Under **high precision** (large w_{E}, Fig. 2-a, Left), ipsi-specific inhibition stabilizes trajectories (Fig. 2-a, Center) and learning (Fig. 2-a, Left); contra-specific inhibition accelerates adaptation but risks instability (Fig. 2-a/b, Right). Under **low precision**, ipsi-specific inhibition hinders learning through over-regularization; contra-specific inhibition facilitates faster adaptation (not shown).

E/I in Recurrent Neural Networks

We compare two E/I-constrained RNN architectures (Fig. 3-a) against a Vanilla RNN (no constraints): **(Left) Column E/I RNNs (ColEI)** (Song et al., 2016), which have separate E/I units with column-wise sign constraints, supporting broad inhibitory motifs (including global inhibition), and **(Right) Dale's ANNs (DANNs)** (Cornford et al., 2021; Li et al., 2023), featuring strictly local inhibition via a reparameterization that reflects the two cortical motifs ($E \rightarrow E$ and $E \rightarrow I \rightarrow E$): $\mathbf{W} = \mathbf{W}_{\mathrm{EE}} - \mathbf{W}_{\mathrm{EI}}\mathbf{W}_{\mathrm{IE}}$, with $\mathbf{W}_{\{\mathrm{EE,EI,IE}\}}$ sign-constrained.

Setting (i): Fixed Contextual Task RNNs trained via supervised learning and backpropagation through time must sustain low pre-stimulus activity, then maintain elevation post-stimulus for the choice corresponding to the largest input stimulus (Fig. 3-b).

Results All models achieve high accuracy; E/I RNNs learn slower, with greater variability between runs. ColEI networks struggle more with ambiguous stimuli (not shown). DANNs exhibit structured attractor dynamics aligned with de-

[1] Code is available at https://github.com/veronicachelu/EI_RLDM.

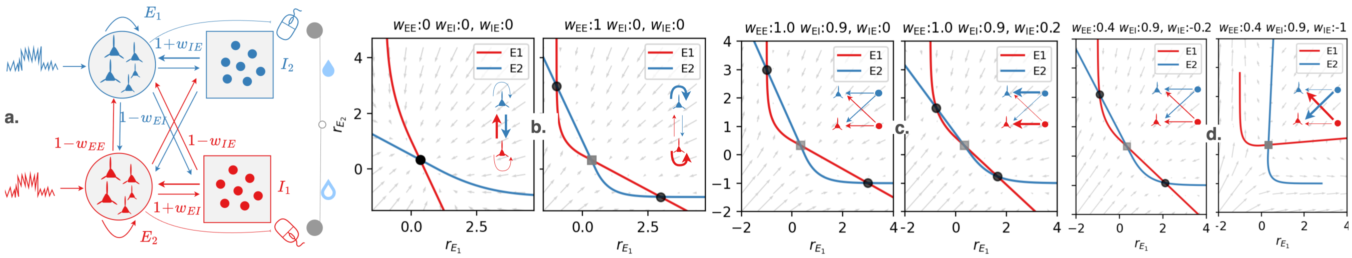

Fig. 1: (a) E/I mean-field model. **(b) Recurrent excitation**. **Selective inhibition**: **(c)** ipsi-specific **(d)** contra-specific.

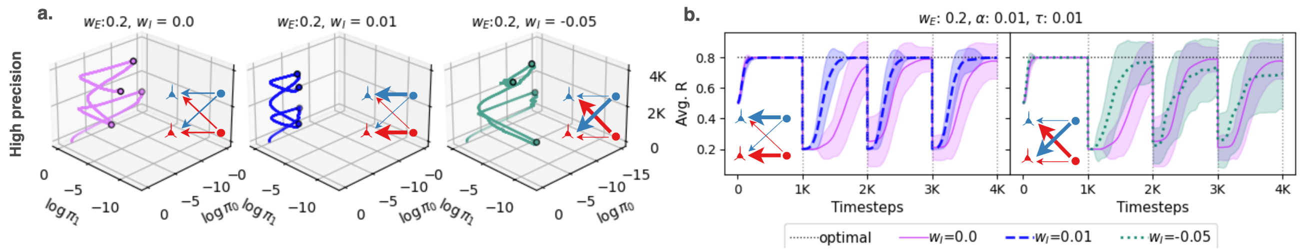

Fig. 2: (a) Trajectories & **(b) performance** in high precision (large w_{E}); ipsi-specific: blue, $w_{\mathrm{I}} > 0$, contra-specific inhibition: green, $w_{\mathrm{I}} < 0$.

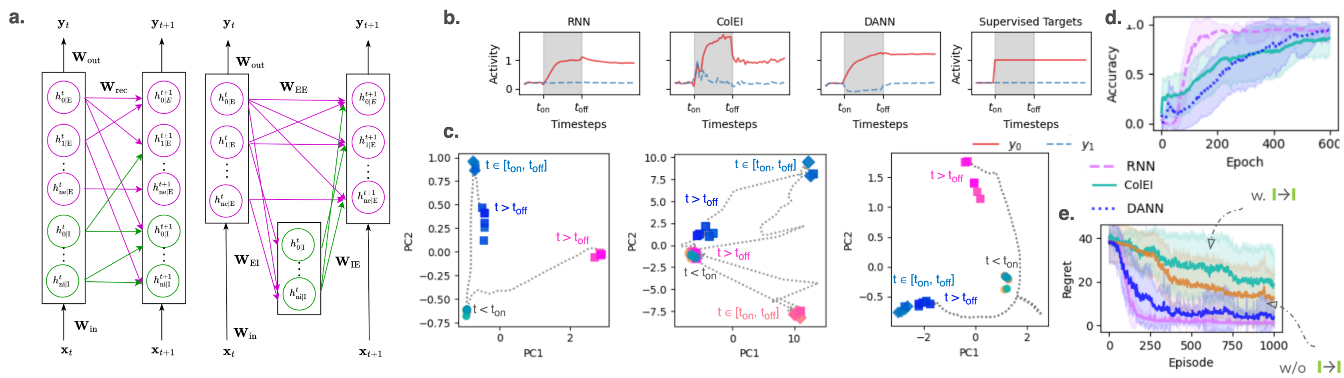

Fig. 3: (a) RNN architectures. (a-Left) *ColEI*: E and I neurons are partitioned by column within each layer. **(a-Right)** *DANN*: inhibition is local between E layers via a separate I layer. **(b) Output activity traces** in setting (i): Example trials from trained RNNs ($c = 0.2$). **(c) Phase-plane analysis:** Trajectory and fixed points found, differentiating pre-, during-, and post-stimulus phases for two values of the evidence coherence c, favouring the two choices. **(d) Performance** in setting (i). **(e) Performance** in setting (ii).

cision phases (Fig. 3-c, Right), similar to vanilla RNNs (Fig. 3-c, Left), whereas ColEI RNNs have poor attractor separation (Fig. 3-c, Center).

Setting (ii): Dynamic Task RNN-based agents are trained to adaptively solve a sequence of tasks drawn from a structured distribution of Bernoulli bandits via meta-learning—a plasticity-based outer-loop RL algorithm (REINFORCE cf. Williams (1992)) slowly establishes an inner-loop algorithm that performs trial-by-trial RL using recurrent activity dynamics.

Results ColEI networks learn very slowly and have high variability over runs; constraining inhibition to be strictly local helps optimization. DANNs demonstrate robust performance, comparable to RNNs (Fig. 3-e).

Discussion

We investigated how two canonical E/I motifs shape adaptive learning and decision-making across three modelling levels:

- **Mean-field model:** Recurrent excitation amplifies signals; inhibition either stabilizes or enhances competition.
- **RL agent:** Interactions between E/I-structured preferences modulate behavioural adaptation.
- **RNNs:** Column-based E/I partitioning (ColEI) impairs learning in dynamic tasks, and inhibitory connectivity plays a role in this. Local inhibition and unconstrained recurrent weights (DANNs) facilitates optimization and mitigates these problems, ensuring robust, biologically plausible performance, comparable to vanilla RNNs.

Overall, these findings highlight the computational role of E/I interactions in modulating behavioural flexibility.

Limitations & Future Directions Future research should focus on integrating biological details such as interneuron diversity, homeostasis, spike-rate adaptation, and disentangle E/I dynamics across inner- and outer-loop learning in dynamic tasks.

Acknowledgements

The authors thank Blake Richards, Roman Pogodin, Colin Bredenberg, and all anonymous reviewers for feedback and insights. Veronica Chelu is grateful for support from IVADO, Fonds d'excellence en recherche Apogée Canada, Bourse d'excellence au doctorat.

References

Bhandari, J., & Russo, D. (2021, 13–15 Apr). On the linear convergence of policy gradient methods for finite mdps. In A. Banerjee & K. Fukumizu (Eds.), *Proceedings of the 24th international conference on artificial intelligence and statistics* (Vol. 130, pp. 2386–2394). PMLR.

Cornford, J., Kalajdzievski, D., Leite, M., Lamarquette, A., Kullmann, D. M., & Richards, B. A. (2021). Learning to live with dale's principle: Anns with separate excitatory and inhibitory units. In *International conference on learning representations.*

Dale, H. (1935). *Pharmacology and nerve-endings.*

Eccles, J. C., Fatt, P., & Koketsu, K. (1954). Cholinergic and inhibitory synapses in a pathway from motor-axon collaterals to motoneurones. *The Journal of physiology*, *126*(3), 524.

Lam, N. H., Borduqui, T., Hallak, J., Roque, A., Anticevic, A., Krystal, J. H., ... Murray, J. D. (2022, 02). Effects of altered excitation-inhibition balance on decision making in a cortical circuit model. *The Journal of Neuroscience*, *42*(6), 1035.

Li, P., Cornford, J., Ghosh, A., & Richards, B. (2023). Learning better with dale's law: a spectral perspective. In *Proceedings of the 37th international conference on neural information processing systems.*

Najafi, F., Elsayed, G., Cao, R., Pnevmatikakis, E., Latham, P., Cunningham, J., & Churchland, A. (2020, Jan). Excitatory and inhibitory subnetworks are equally selective during decision-making and emerge simultaneously during learning. *Neuron*, *105*(1), 165–179.

Roach, J. P., Churchland, A. K., & Engel, T. A. (2023). Choice selective inhibition drives stability and competition in decision circuits. *Nature Communications*, *14*(1), 147.

Song, H. F., Yang, G. R., & Wang, X.-J. (2016, 02). Training excitatory-inhibitory recurrent neural networks for cognitive tasks: A simple and flexible framework. *PLOS Computational Biology*, *12*(2), e1004792–.

Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.*, *8*, 229–256. Retrieved from `https://doi.org/10.1007/BF00992696` doi: 10.1007/BF00992696

Wilson, H. R., & Cowan, J. D. (1972, Jan). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys J*, *12*(1), 1–24.