Emergent oscillations in a cortical column model of predictive coding with multiple interneuron types

Kwangjun Lee (k.lee@uva.nl)

Cyriel M.A. Pennartz (c.m.a.pennartz@uva.nl)

Jorge F. Mejias (j.f.mejias@uva.nl)

Cognitive and Systems Neuroscience Group, Swammerdam Institute for Life Sciences
University of Amsterdam, Amsterdam, the Netherlands

Abstract

We propose a biologically grounded computational model of predictive coding (PC) that integrates a neuroanatomically informed hierarchy of cortical areas with laminar organization and cell-type-specific connectivity. The model performs PC on naturalistic images through Hebbian learning and prediction error minimization. The model assumes that stereotypical pyramidal-PV-SST-VIP circuits with the same structure but different bottom-up and top-down inputs compute positive and negative prediction errors. Sensory inference in the model generates neural oscillations, with simulations of optogenetic inactivation revealing distinct roles for PV, SST, and VIP cells in these dynamics. Furthermore, the model exhibits mismatch negativity-like responses to deviant stimuli. This work offers a biologically plausible framework for understanding the neural circuits underlying PC in the cortex.

Keywords: predictive coding; interneuron; oscillation; cortical circuit; sensory inference; perception

Introduction

The brain faces the challenge of inferring the properties of objects from inherently noisy sensory signals. Predictive coding (PC) (Rao & Ballard, 1999; Mumford, 1992; Srinivasan, Laughlin, & Dubs, 1982) views the brain as an inference machine that constantly generates an internal model of the world and minimizes the discrepancies between predicted and actual sensory inputs (i.e. prediction error). Although PC principles have been extensively explored at a theoretical level, their biological plausibility within cortical circuits is still largely lacking. Recent experimental findings (Attinger, Wang, & Keller, 2017; Hertäg & Clopath, 2022; Keller & Mrsic-Flogel, 2018) have shown the crucial role of different inhibitory interneuron subtypes, such as PV, SST, and VIP cells, in the calculation and propagation of prediction errors. To address this gap, our study presents a biologically grounded computational model of visual perception based on PC. This model uniquely integrates neuroanatomically informed projections between cortical areas with precise laminar organization and the diversity of cell types characteristic of sensory cortex neural circuits, aiming to provide a more mechanistic understanding of predictive processing in the brain.

Methods

Neurons followed a linear firing rate model (Wilson & Cowan, 1972): $\tau\, dr_i/dt = -r_i + f\left(\sum_j W_{ij} r_j\right)$. The model consisted of two cortical areas (Fig. 1A): Area 1 received bottom-up sensory input in layer 4 and top-down predictions from Area 2 to compute prediction errors in layer 2/3; Area 2 encoded internal representations using pyramidal and PV cells. A key element was the identification of a canonical microcircuit motif (composed of pyramidal, PV, SST, and VIP cells) for prediction error computation through an exhaustive combinatorial search that explored numerous potential within-circuit connectivity (among the four neuron types) and synaptic input patterns (bottom-up sensory inputs and top-down predictions). This search was constrained by established neurobiological principles (Bastos et al., 2012; Pfeffer, Xue, He, Huang, & Scanziani, 2013; Pennartz, Dora, Muckli, & Lorteije, 2019), such as pyramidal cells receiving bottom-up input and VIP cells receiving top-down input. For our study, we randomly selected one of the 173 resulting combinations of within-circuit connectivity and synaptic input patterns to illustrate the general principles of prediction error computation within our model. This circuit exclusively computes either positive or negative prediction errors (Lee, Dora, Mejias, Bohte, & Pennartz, 2024) based on the specific pattern of bottom-up and top-down input received by its constituent cells. The synaptic weights between Area 1 and Area 2 were plastic and learned using the Hebbian learning rule with grayscale CIFAR-10 images, while the within-microcircuit weights, determined by combinatorial search, remained fixed.

Results

The model effectively predicted novel naturalistic images after training, demonstrating the successful learning of an internal model. Notably, oscillatory dynamics emerged in both representation and prediction error microcircuits (Fig. 1B), reflecting the temporal dynamics of error minimization, consistent with experimental findings (Alamia & VanRullen, 2019). These oscillatory dynamics persisted even during an oddball paradigm (Fig. 1C), where a deviant stimulus (D) within a sequence of repeating standard stimuli (S) elicited increased neural activity (mismatch negativity-like response) in both representation and prediction error microcircuits. Note that these oscillations dampened over time and converged to zero in prediction error circuits and a stable focus in representation circuits, indicating error minimization and the formation of latent representations, respectively. Selective inactivation of each cell type in the network, mimicking optogenetic experiments, revealed that interneurons have differential roles in PC (Fig. 1D): while PV provides a blanket of inhibition, dampening network activity, SST and VIP serve to control the amplitude of neural oscillations and the number of cycles between representation and error neurons in opposing directions.

Conclusion

A novel cortical network model with a laminar and cell type-specific architecture is proposed, demonstrating the capacity to perform perceptual inference and learning of sensory input. In particular, rhythmic activity spontaneously emerged during the propagation of predictions and errors. We dissected the contribution of different types of cortical neurons to the generation of prediction error (Lee, Pennartz, & Mejias, 2024).

Acknowledgements

This work was done with the support of the EBRAINS and HBP computing services. We thank Maria Nefeli Panagiotou for the early iterations of this work.

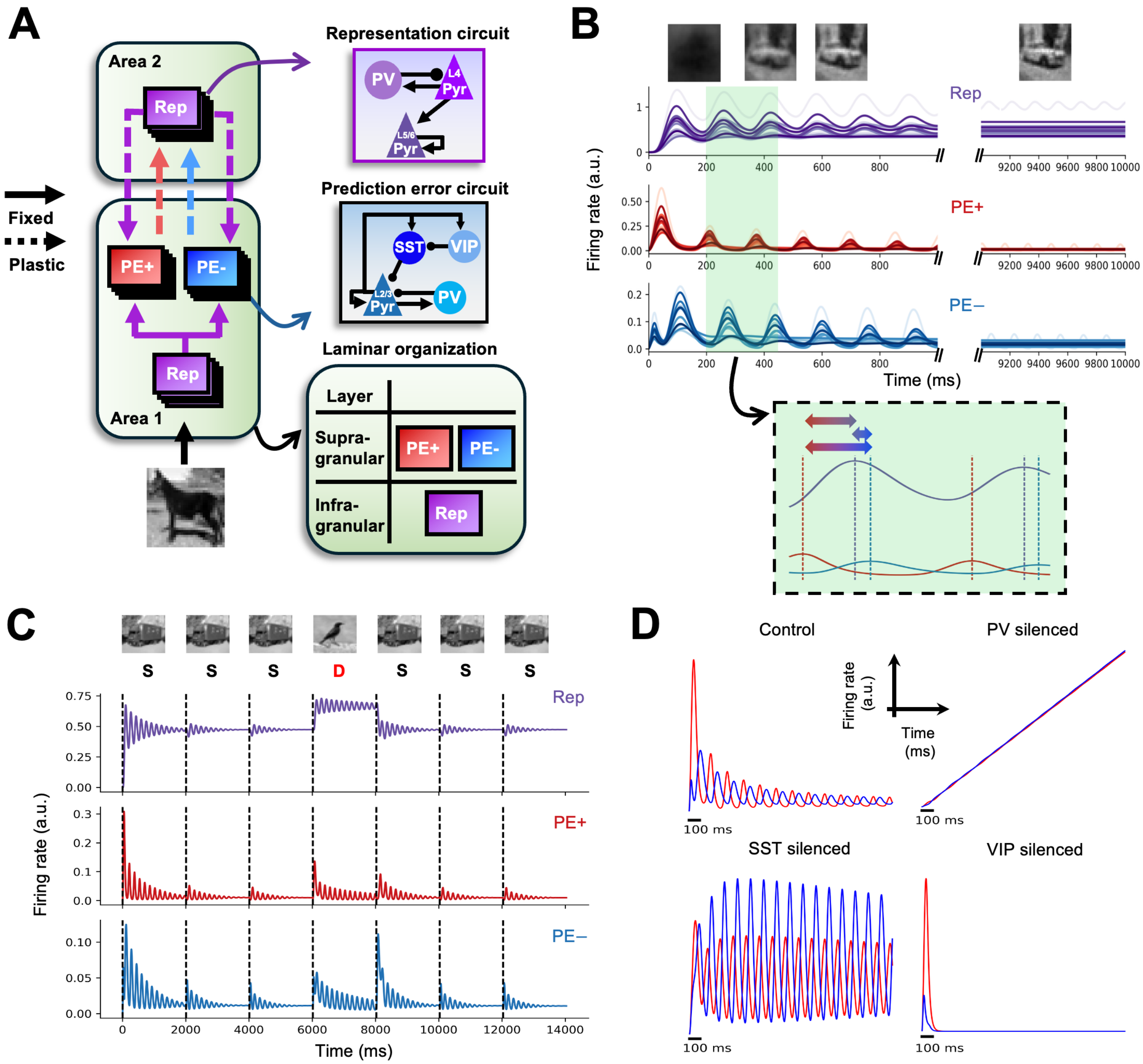

Figure 1: Cortical column model of predictive coding. (A) Schematic of the cortical column model showing laminar organization and the division into prediction error and representation circuits. (B) Population activity of pyramidal cells in representation (Rep; purple), positive (PE+; red), and negative (PE-; blue) prediction error microcircuits exhibited rhythmic responses to naturalistic images. As prediction errors declined, Rep activity stabilized, supporting improved reconstructions (top images). The highlighted area in green shows stable phase relationships between PE+ and PE- (bottom arrow), PE- and Rep (middle), and PE+ and Rep (top). (C) Activity of L5/6 pyramidal cells in Rep (purple) and L2/3 pyramidal cells in PE+ and PE- (red and blue) during an oddball sequence (e.g., truck and bird images). Prediction errors decreased over time through inference. Repetition of the standard stimulus (S) led to lower initial errors and faster convergence, while repeated exposure also accelerated Rep stabilization. Deviant stimuli (D) evoked heightened PE activity and distinct steady-state representations. (D) Optogenetic silencing of PV, SST, or VIP interneurons revealed their distinct roles. Panels show firing rates of L2/3 pyramidal cells in PE+ and PE- (red and blue) under control and silenced conditions. In control, prediction errors declined over time. Silencing PV cells resulted in a continuous increase of pyramidal cell firing rates in both PE+ and PE- circuits and rid of oscillatory behavior. Silencing SST cells disrupted prediction error minimization, while silencing VIP cells shut down both PE circuits.

References

Alamia, A., & VanRullen, R. (2019). Alpha oscillations and traveling waves: Signatures of predictive coding? *PLoS Biology*, *17*(10), e3000487.

Attinger, A., Wang, B., & Keller, G. B. (2017). Visuomotor coupling shapes the functional development of mouse visual cortex. *Cell*, *169*(7), 1291–1302.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695–711.

Hertäg, L., & Clopath, C. (2022). Prediction-error neurons in circuits with multiple neuron types: Formation, refinement, and functional implications. *Proceedings of the National Academy of Sciences*, *119*(13), e2115699119.

Keller, G. B., & Mrsic-Flogel, T. D. (2018). Predictive processing: a canonical cortical computation. *Neuron*, *100*(2), 424–435.

Lee, K., Dora, S., Mejias, J. F., Bohte, S. M., & Pennartz, C. M. (2024). Predictive coding with spiking neurons and feedforward gist signaling. *Frontiers in Computational Neuroscience*, *18*, 1338280.

Lee, K., Pennartz, C. M., & Mejias, J. F. (2024). Cortical networks with multiple interneuron types generate oscillatory patterns during predictive coding. *bioRxiv*, 2024–10.

Mumford, D. (1992). On the computational architecture of the neocortex: Ii the role of cortico-cortical loops. *Biological cybernetics*, *66*(3), 241–251.

Pennartz, C. M., Dora, S., Muckli, L., & Lorteije, J. A. (2019). Towards a unified view on pathways and functions of neural recurrent processing. *Trends in neurosciences*, *42*(9), 589–603.

Pfeffer, C. K., Xue, M., He, M., Huang, Z. J., & Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nature neuroscience*, *16*(8), 1068–1076.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, *2*(1), 79–87.

Srinivasan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *216*(1205), 427–459.

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