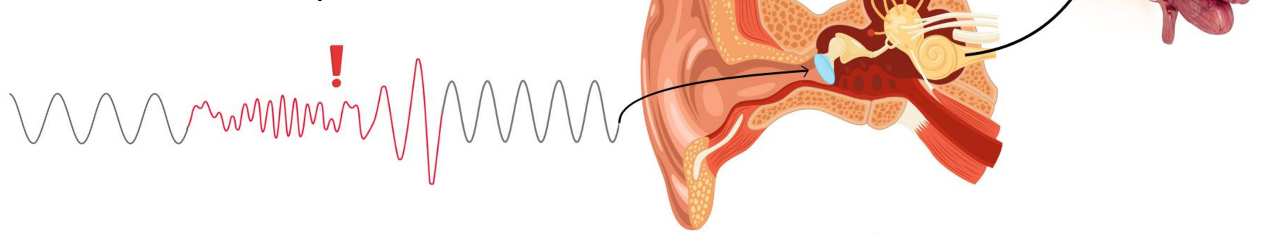
Emergence of Auditory Receptive Fields Based on Surprise at Multiple Timescales

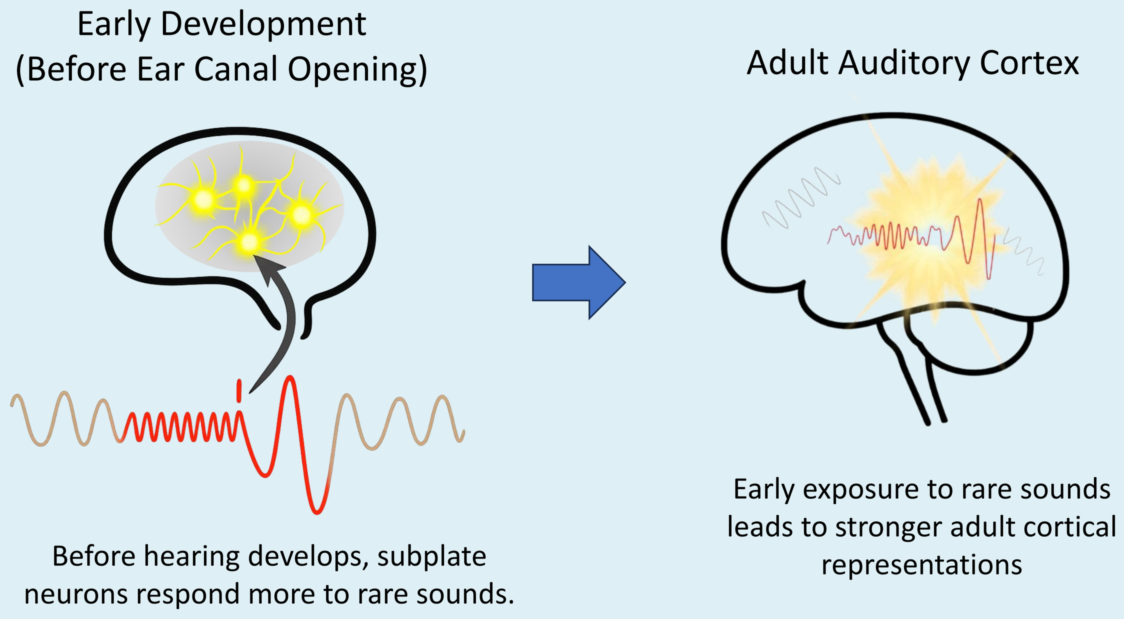
Yashaswini[1]; Sneha Dash[1]; Sharba Bandyopadhyay[1]
[1]Information Processing Laboratory, Dept of E&ECE, Indian Institute of Technology Kharagpur

Surprise! Did you hear that coming?

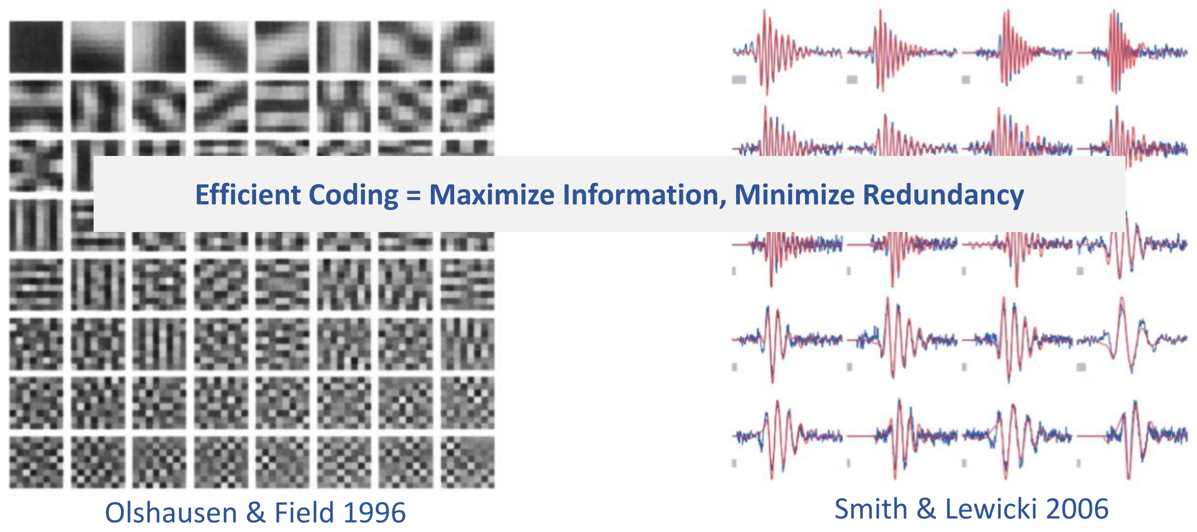
Auditory cortical neurons respond strongly to rare, surprising sounds in a sequence.

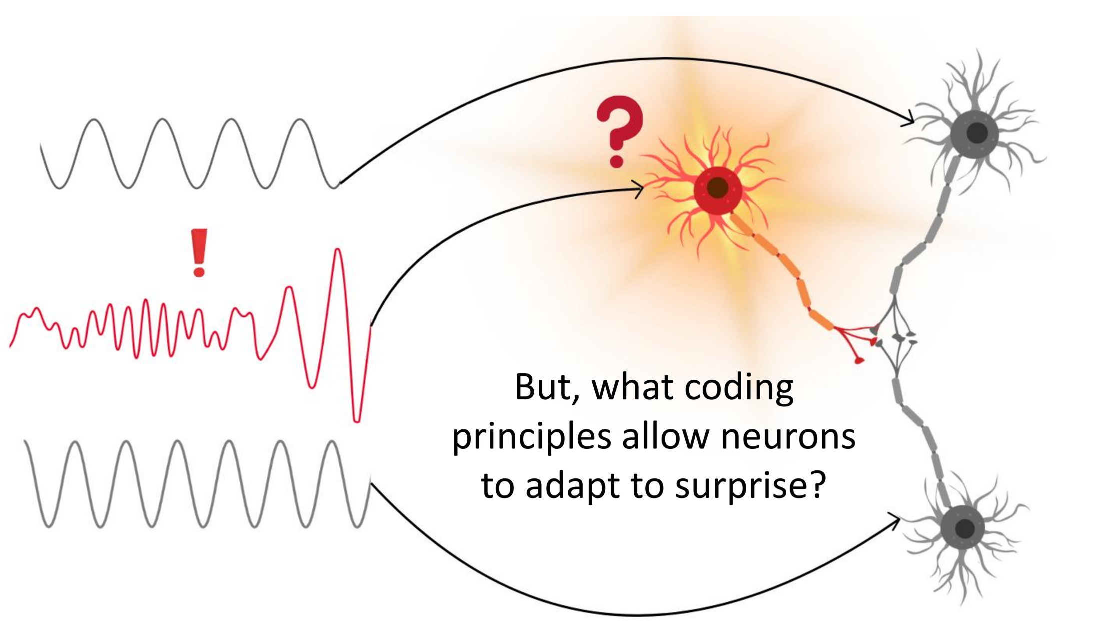


Insights from Mehra et. al., *J Neurosci.* 2022: Early Experience of Rare Sounds Causes Long-Term Changes in the Adult Auditory Cortex

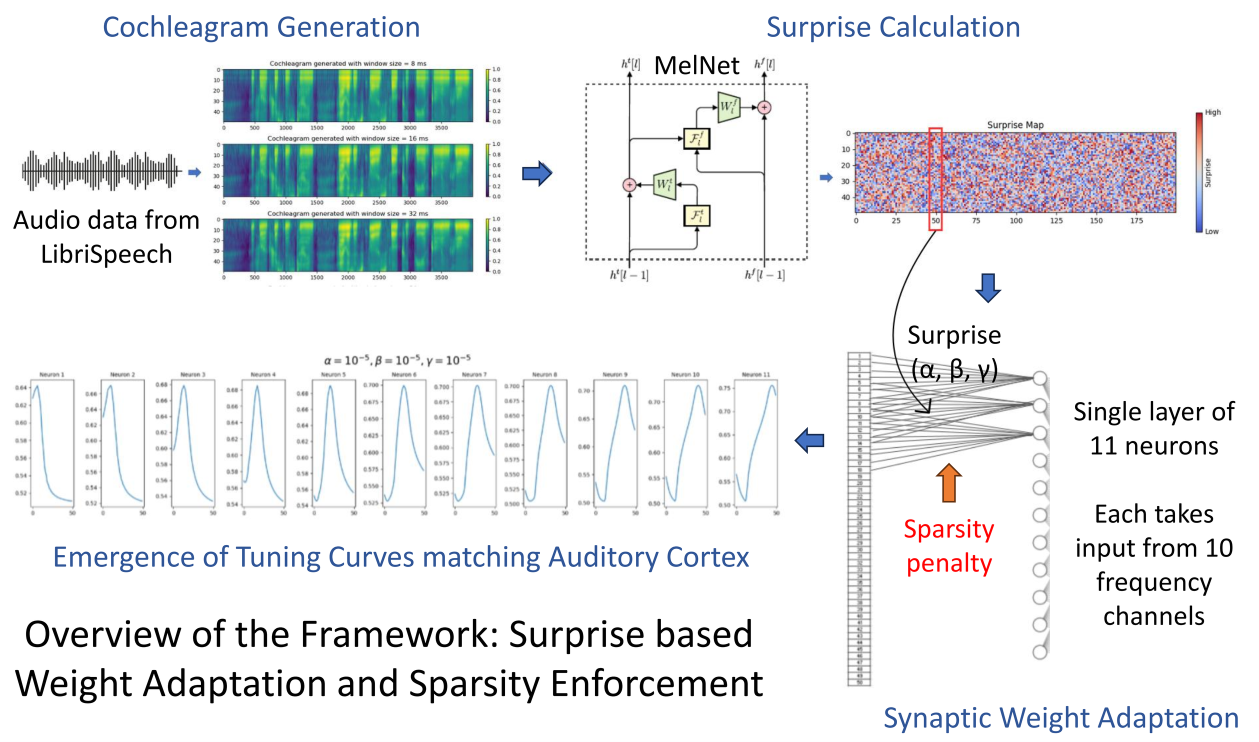


Classical models show how neurons encode natural stimuli efficiently and sparsely, using as few active neurons as possible.



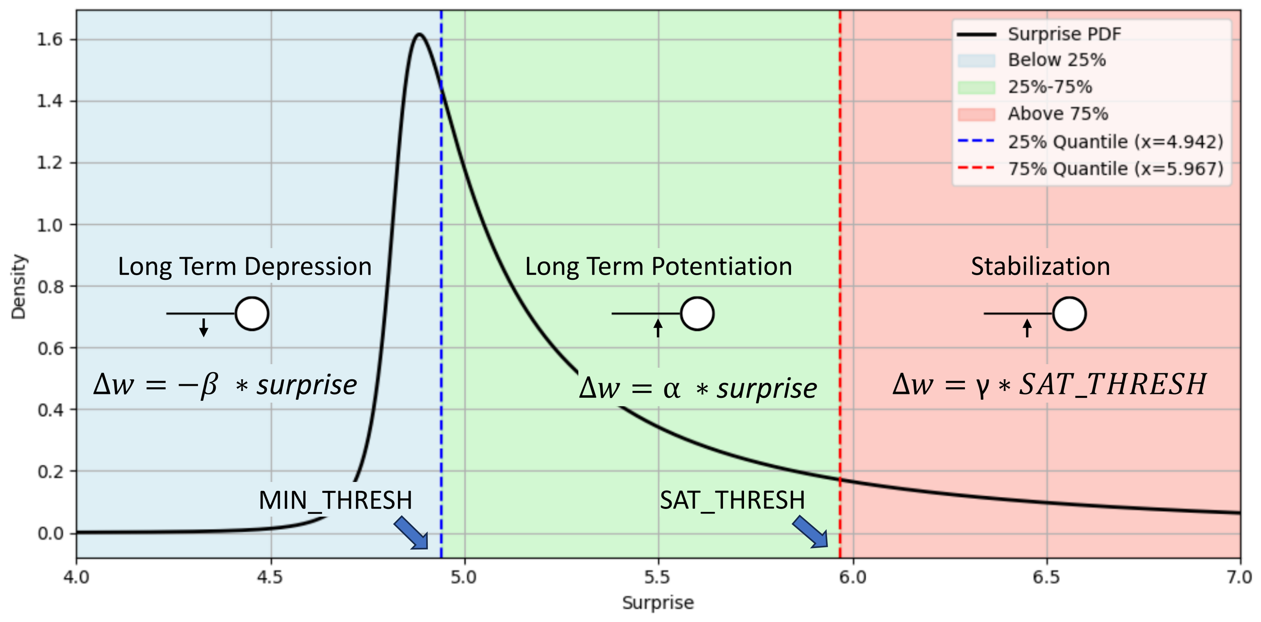


Surprise-Driven Adaptation



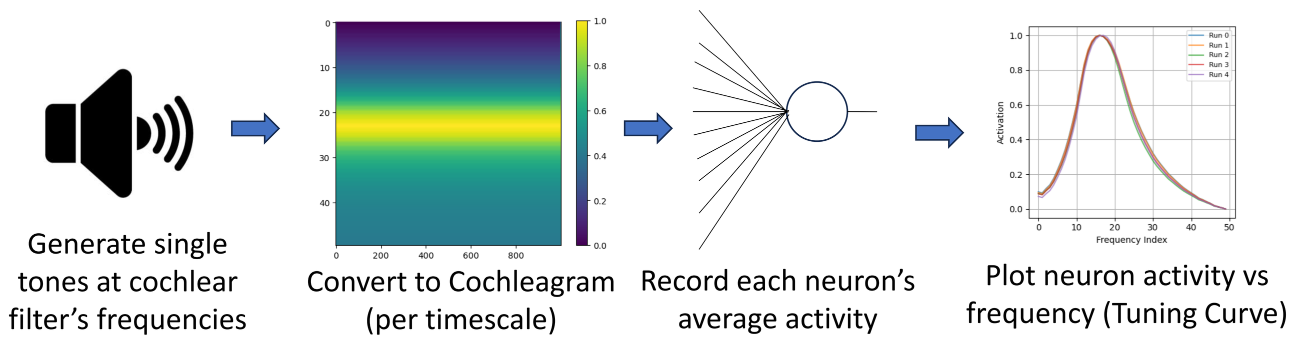
Overview of the Framework: Surprise based Weight Adaptation and Sparsity Enforcement

The Learning Rule

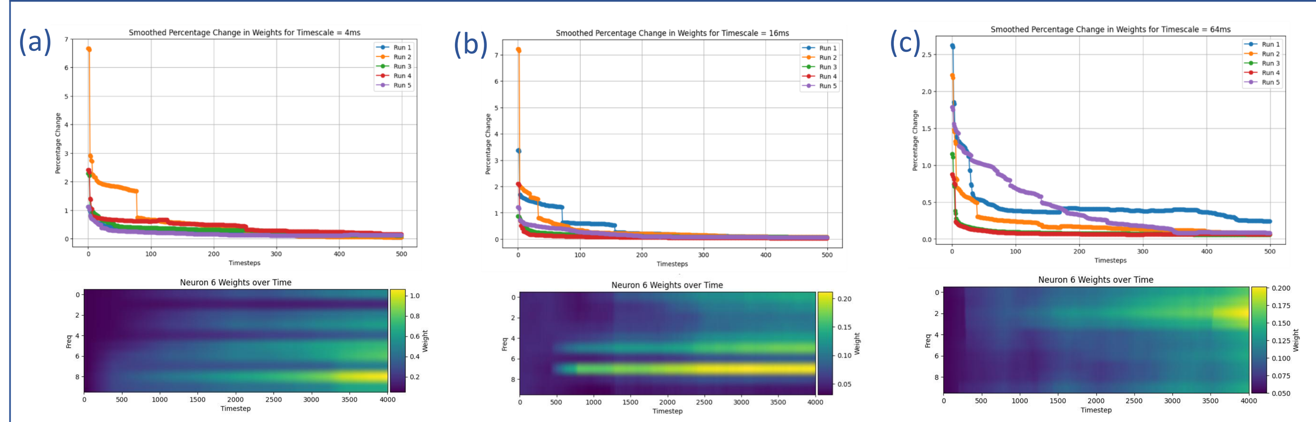


Surprise distribution estimated from MelNet is used to set thresholds at quartiles (MIN_THRESH, SAT_THRESH), driving a three-factor learning rule: depression (β) for low surprise, potentiation (α) for moderate surprise, and stabilization (γ) for high surprise. Post-update, sparsity is enforced via L1-norm gradient descent on activations.

Generating Tuning Curves

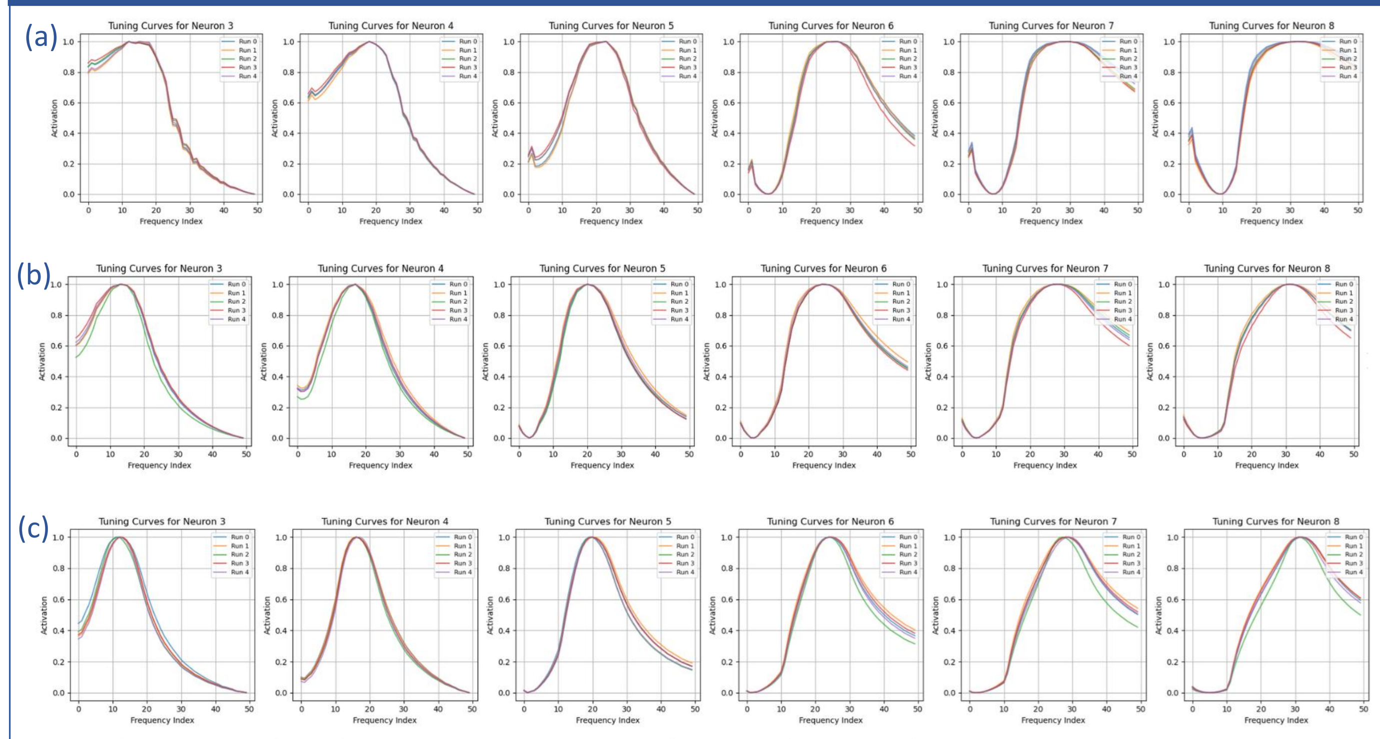


Fast Convergence of Weights



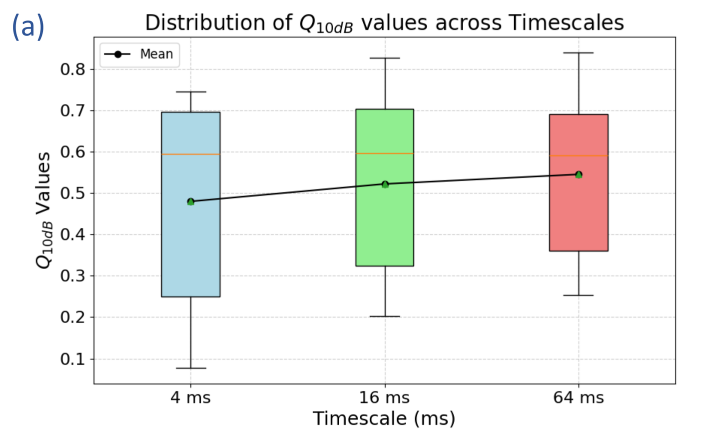
(a-c) Smoothed percentage change in synaptic weights across five runs shows fast and robust convergence at (a) 4 ms, (b) 16 ms, and (c) 64 ms timescales. **Lower panels**: Weight matrices of a neuron over time illustrate stable adaptation and frequency-specific structure at each timescale.

Tuning Curves at Different Timescales



(a–c) Normalized tuning curves of a subset of neurons across 5 runs demonstrate robustness, frequency selectivity and sideband inhibition at (a) 4 ms, (b) 16 ms, and (c) 64 ms timescales.

Alignment with Neurophysiological Data



(b) Table 1: Q_{10dB} values (median, 10th, and 90th percentiles) from our model across timescales. Highlighted cells show close correspondence with A1 cortex measurements (Table 2).

Parameter	4 ms	16 ms	64 ms
Q_{10dB} (median)	0.594	0.596	0.590
10th percentile	*0.146*	*0.234*	*0.330*
90th percentile	*0.738*	*0.778*	*0.760*

(c) Table 2: Q_{10dB} values (median and 10th–90th percentile) for different ferret auditory cortical areas, adapted from Bizley et al. (2005). A1 cortex values are highlighted.

Parameter	All areas	A1	AAF	PSF	PPF	ADF	AVF
Q_{10dB} (median)	1.37	0.59	0.27	0.45	0.67	0.35	0.25
10th–90th percentile	*0.84–3.1*	*0.41–0.74*	*0.19–0.88*	*0–2.31*	*0–0*	*0.18–1.96*	*0.21–1.5*

(a) Q10dB boxplots show a **mean increase across timescales**, indicating **broader tuning at slower timescales** and **sharper tuning at faster timescales**, consistent with neurophysiological findings (Rodriguez et al. (2010)). **(b-c)** Model Tuning Curves Q10dB values show close correspondence with A1 cortex reference data from Bizley et al. (2005), particularly in **median values** (**0.594–0.596** vs. **0.59**) and **90th percentiles** (**0.738–0.778** vs **0.74**) as highlighted in Table 1 and Table 2.

Key Findings and Future Work

- Our model demonstrates that **surprise-driven synaptic adaptation**, combined with **efficient coding**, can generate **biologically plausible auditory receptive fields**.
- The emergence of **broader tuning at fast timescales** and **sharper tuning at slow timescales** closely matches neurophysiological data obtained from the auditory cortex.
- **Q10dB values** of our model's tuning curves closely match those measured in the **ferret A1 cortex**, supporting the biological plausibility of our results.
- Our framework predicts **intensity-dependent tuning**, narrow Frequency Response Areas at low SPLs and broad at high SPLs. Quantitative validation against neurophysiological data will be addressed in future work.
- Future work would include **replacing MelNet** with an **adaptive network** that updates its **prior distribution** through **experience** and calculates surprise **dynamically**, creating a more biologically plausible surprise-driven learning system. The current approach can also be extended to **multi-layer** architectures to better capture **hierarchical** auditory processing.

Contact

Yashaswini
Indian Institute of Technology Kharagpur
Email: yashaswini27aug@gmail.com
Phone: +91 9538539797

References

1. Vasquez, S. and Lewis, M. (2020) 'MelNet: A generative model for audio', arXiv preprint arXiv:1910.06711.
2. Panayotov, V., Chen, G., Povey, D. and Khudanpur, S. (2015) 'LibriSpeech: An ASR corpus', Interspeech 2015.
3. Olshausen, B.A. and Field, D.J. (1996) 'Emergence of simple-cell receptive field properties', Nature, 381(6583), pp. 607–609.
4. Lewicki, M.S. (2002) 'Efficient coding of natural sounds', Nature Neuroscience, 5(4), pp. 356–363.
5. Mehra, M., Bandyopadhyay, S., Kumar, A., et al. (2022) 'Rare sound experience causes changes in adult auditory cortex', Journal of Neuroscience, 42(8), pp. 1454–1476.
6. Frémaux, N. and Gerstner, W. (2016) 'Neuromodulated spike-timing-dependent plasticity', Frontiers in Neural Circuits, 9, p. 85.
7. Rodriguez, F.A., Chen, G.D., Read, H.L., et al. (2010) 'Spectral and temporal modulation tradeoff', Journal of Neurophysiology, 103(2), pp. 887–903.
8. Bizley, J.K., Nodal, F.R., Nelken, I. and King, A.J. (2005) 'Functional organization of ferret auditory cortex', Cerebral Cortex, 15(10), pp. 1637–1653.