NEUR 1440: Astrocyte Regulation Confines Representational Drift to Null-space

Yishu Li

November 2025

1 Representational Drift and Null-Hypothesis

1.1 Representational Drift

Representational drift refers to the phenomenon where the neural activity patterns associated with specific sensory stimuli, cognitive states, or motor behaviors change continuously over time, even when the external behavior remains stable. Traditionally, theories of neural coding assumed that the relationship between neural firing and specific variables, such as movement direction, was static. However, recent chronic recording studies have demonstrated that the neural representation is highly dynamic. Individual neurons may alter their tuning curves, drop out of the ensemble, or be replaced by other neurons in the population.

This phenomenon is widespread across cortical areas. In the primary visual cortex (V1), studies have demonstrated that while mice performed a consistent visual task, the cosine tuning of neurons responding to the stimuli of specific directions evolved over weeks, with only a small fraction of neurons maintaining stable tuning [5]. Similarly, in the primary motor cortex (M1), neurons typically exhibit directional cosine tuning, where the firing rate is modulated by the angle of movement related to a specific "preferred direction" [8]. Representational drift manifests here as the instability of these tuning parameters: a neuron's preferred direction may rotate, or its modulation depth may fluctuate significantly over days. Consequently, a neuron that encodes "leftward" movement on one day might shift to encode a different direction—or cease modulation entirely—in subsequent sessions [2]. A simple illustration of the representational drift in cosine tuning is shown below in Figure 1.

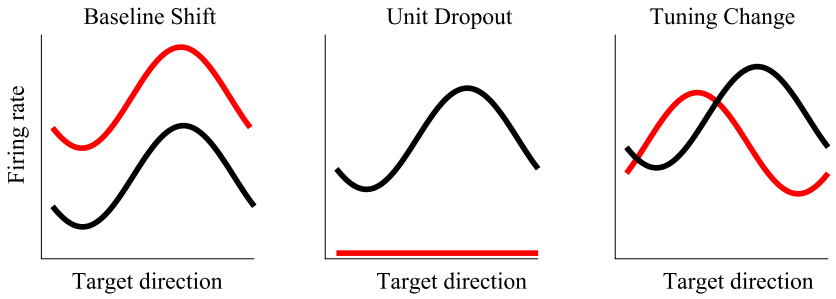


Figure 1: Example of representational drift in cosine tunings.

For intracortical Brain-Computer Interfaces (iBCIs), presentational drift is a primary source of signal non-stationarity, posing a critical hurdle to long-term system robustness. The phenomenon of representational drift presents three key challenges for the effective application of iBCIs:

1. **Decoder Degradation:** Most iBCI decoders, such as Kalman filters [19] or deep learning models [11], are trained on data collected during a specific session. Because the underlying neural distribution shifts over days or weeks due to representational drift, a static decoder inevitably loses accuracy [10]. The mapping learned on "Day 1" becomes obsolete by "Day 10" as the neural representation evolves.

2. **Recalibration Burden:** To counteract drift, traditional iBCI systems often require daily retraining or recalibration. This process is time-consuming and mentally exhausting for users, significantly reducing the clinical viability and "plug-and-play" utility of the device.

3. **Generalization Challenges:** Drift complicates the development of models that can generalize across long time scales. It necessitates advanced signal processing techniques, such as unsupervised domain adaptation [12] and manifold alignment [2], to track the stable latent dynamics from the shifting neural population.

The gap between the inconsistency in neural representations and the stability in behaviors requires more investigation.

1.2 Neural Population and Null-space

Recently, one important explanation on how behavior emerges from unstable neural representations is that instead of being encoded by the individual cells, the behaviors are encoded by the collective dynamics of the neural population.

1.2.1 Neural Population Perspective

Instead of analyzing neurons in isolation (e.g., "Neuron A encodes leftward movement"), the instantaneous activity of N recorded neurons can be viewed as a high-dimensional state vector, $\mathbf{x}(t) \in \mathbb{R}^N$. In this view, the behavior is not encoded by any single neuron firing rate, but by the pattern of co-activation across the entire population [1].

1.2.2 Neural Null-space

The mapping from the high-dimensional neural population in $\mathbb{R}^N$ space to low-dimensional behavior space, which has only very small number of degree of freedom ($\mathbb{R}^K$, where $K \ll N$), induced a massive reduction in dimensionality and information, which implies a high neural redundancy.

Mathematically, if we model a simple additive relationship between two neurons y_1, y_2, and their behavior readout x:

$$x = y_1 + y_2$$

the neural space can be decomposed into two orthogonal spaces, as illustrated in Figure 2. When the neural population activity of y_1 and y_2 moves along the axis labeled as null, the low-dimensional behavior readout x doesn't change, and only when the neural population activity's change has components on the coding axis, the behavior will change. Therefore, the neural population change can be divided into a Coding and a Null-space [17].

1.2.3 The Null-Hypothesis for Drift

The decomposition offered a powerful explanation for the paradox of representational drift. The Null-space Hypothesis suggests that the brain allows representational drift to occur, provided that this drift is largely confined to the Output-Null space [6]. Even though individual neurons could contribute to both coding and null-space, as their tuning curves drift, the plastic mechanism in the network acts to constrain this drift by adjusting other neurons around this drift to maintain the population activity drift in the null-space.

However, as the Null-Hypothesis is getting more widely accepted, the physiological mechanism of this null-space population regulation remains unclear. In this paper, I'm going to discuss astrocyte cells' potential role in regulating the neural population drift in the null-space to close this loop.

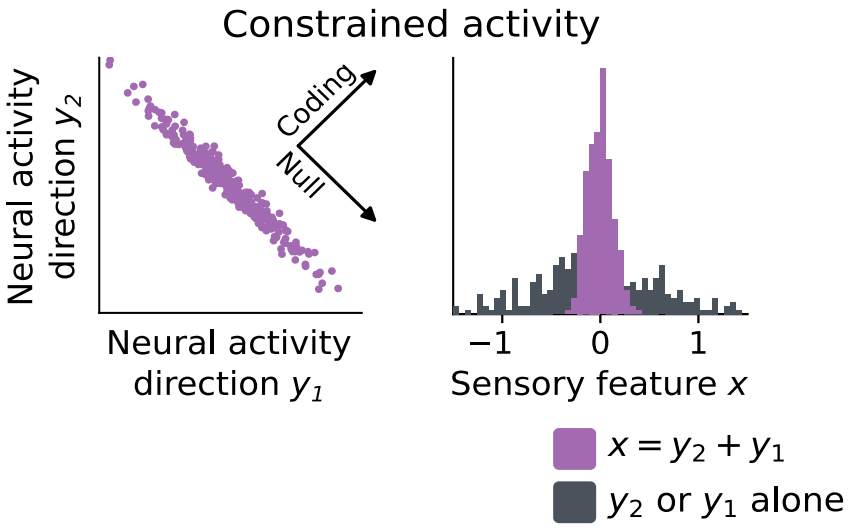


Figure 2: A simple example of neural population activity and low-dimensional readout.

2 Astrocyte Regulation Confines Representational Drift to Null-space

2.1 Astrocyte Properties

Astrocytes are the most abundant glial cells in the mammalian brain. Astrocytes are named for their star-like shape, featuring thousands of fine processes. Astrocytes exhibit "tiling" – they arrange themselves in non-overlapping domains, and each astrocyte has exclusive control over a specific, immense territory, covering up to 2 million synapses. This tiling property ensures that the astrocytic regulation is localized and acts as an ideal "integrator" to detect large-scale network drift that individual neurons cannot perceive.

2.2 Astrocyte Functions

Historically, astrocytes were viewed merely as passive support cells that provided structural scaffolding and metabolic support to neurons as the "neural glue". However, modern neuroscience has shifted to the "Tripartite Synapse" model, which establishes that astrocytes are active participants in information processing [14]. There are key astrocyte functions relevant to neural dynamic regulations:

- **Neurotransmitter Clearance:** Astrocytes could rapidly remove glutamate from the synaptic cleft to prevent neurons from overactivation caused by the accumulation of glutamate in synapses, and maintain temporal precision. Knocking off the glutamate transporters on astrocytes will lead to the overactivation of the nearby neurons and failure in memory formation and learning [4, 7].
- **Gliotransmission** Astrocytes can release chemical signals, such as ATP, D-serine, that directly modulate synaptic plasticity, in response to neuronal activities. d-Serine released by astrocytes could increase the neuron excitability and promote the learning process [9], while the astrocytic ATP/Adenosine enables heterosynaptic suppression to stabilize networks [16].
- **Potassium Spatial Buffering:** Studies have shown that astrocytes can act as a spatial buffer to transport potassium ions, which participate in the formation of the action potential and are important in the regulation of synapse weights, from the high-concentration area to the low-concentration area, thus balance the global electrical potential and excitability of the network [13].

From these functions, it is obvious that astrocytes are able to participate in neural information processing and regulate neural population activities. In the context of representational drift, it's easy to hypothesize

that astrocytes act as a biological feedback control loop that maintains the excitation-inhibition balance of the network.

2.3 Astrocytic Regulation on Representational Drift

From the properties and functions of astrocytes we mentioned above, there might be two constraints on the neural population drift applied to neurons by astrocytes:

1. **Single Neuron Anchoring:** Astrocytes may act as anchors to keep a single neuron from drifting too far from its original representation. While individual neurons may drift, such as changing their firing rate or tuning curves, astrocytes that are directly monitoring their output synapses can sense local hyperactivity through their widespread processes and release gliotransmitters to suppress synaptic gain or upregulate glutamate transporters to dampen excitation. The longer effective scale period of astrocytes [7] may store the previous neuron firing pattern and work as a feedback regulator to anchor the single neuron firing pattern from drifting too much.

2. **Population Drift Direction Constraint:** Astrocytes also constrain the neural population drift direction by activating the Hebbian learning of the entire local neuron ensemble to readjust the tuning and weights, so the population readout of this whole neural ensemble stays constant. The wide spatial range of the astrocytes may also allow the astrocytes to directly read the output of the entire neural ensemble; therefore, astrocytes could make sure the output of the neural ensemble stays unchanged by adjusting the local neurons and monitoring the direct readouts. To physiologically achieve these functions, one possible pathway is that astrocytes could react to the neuron activities by changing the free calcium Ca^{2+} concentration in the corresponding astrocyte process, which induces calcium waves that could spread to the nearby astrocyte processes, causing them to release gliotransmitters and finally promoting the learning process in the nearby neurons [15].

These two processes, facilitated by astrocytes, work together to keep the animal's behavior consistent with the neural representational drifts. Part of the process is illustrated in Figure 3.

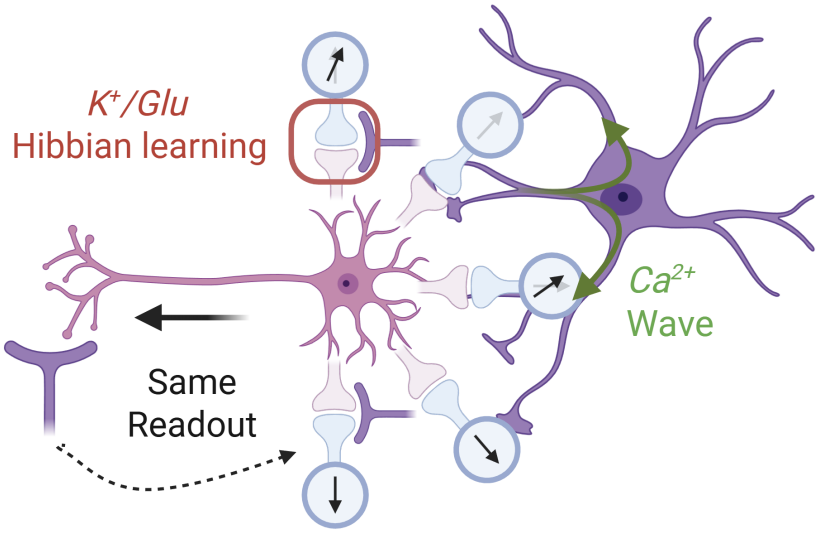


Figure 3: Astrocytes regulate the local neural ensemble in response to the single-neuron representational drift.

3 Astrocyte Regulates Representational Drifts through Neuron Ensembles

3.1 The Excitability-Drift Hypothesis

Beyond their potential role in constraining population dynamics within the Output-Null space, astrocytes actively modulate the intrinsic excitability of individual neurons—a factor increasingly recognized as a primary driver of representational drift [3].

Current hypotheses suggest that representational drift is not a purely stochastic process but is driven by slow, continuous fluctuations in single-neuron excitability. Neurons operate with internal periods of volatility; their responsiveness to synaptic inputs is not static but follows a slow-timescale oscillation. As illustrated in Figure 4, this creates a dynamic "recruitment bias": at any given time, only the subset of neurons currently exhibiting high intrinsic excitability is recruited into the functional encoding ensemble. As the excitability profile of the population shifts over days, the active subset changes, resulting in the observed drift of neural representations.

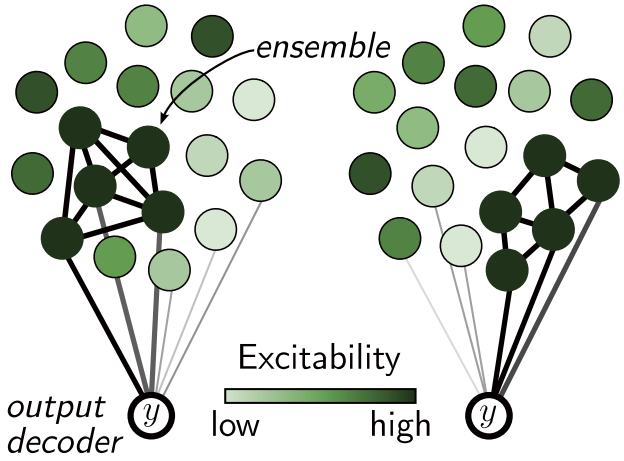


Figure 4: Only neurons with high excitability are recruited for behavior encoding.

3.2 Astrocyte as Excitability Scheduler

Astrocytes are ideally positioned to govern these slow fluctuations in excitability. By controlling the uptake of glutamate via GLT1 transporters, astrocytes effectively tune the "gain" or plasticity of the local circuit. This regulatory mechanism can be analogized to a learning rate scheduler in machine learning optimization algorithms, such as Stochastic Gradient Descent (SGD).

A recent study in Sur's lab provided compelling evidence for this computational role. [4] demonstrated that deactivating the astrocyte glutamate transporter (GLT1) leads to unregulated, persistently high neuronal excitability. As illustrated in Figure 5, the GLT1-deficient mice exhibited significantly lower response times during intermediate training (Sessions 6–10), indicating a faster learning speed compared to healthy controls. However, this rapid acquisition came at a cost: the inter-trial motor trajectory similarity and smoothness were significantly lower.

This phenomenon parallels the behavior of a deep learning model trained with a consistently high learning rate:

- **High Excitability (GLT1-Deficient) ≈ High Learning Rate:** The system updates weights rapidly (fast learning) but "overshoots" or oscillates around the global minimum, resulting in noisy, unstable trajectories.
- **Regulated Excitability (Healthy Control) ≈ Scheduled Decay:** In healthy subjects, astrocytes dynamically dampen excitability as the task is learned. This acts like a scheduler that anneals the

learning rate as the loss decreases, ensuring that the optimization process slows down to allow the neural manifold to converge to a smooth, stable, and robust solution.

Therefore, by modulating the recruitment of neurons through excitability changes, astrocytes essentially balance the trade-off between exploration (neural plasticity) and stability during the learning process.

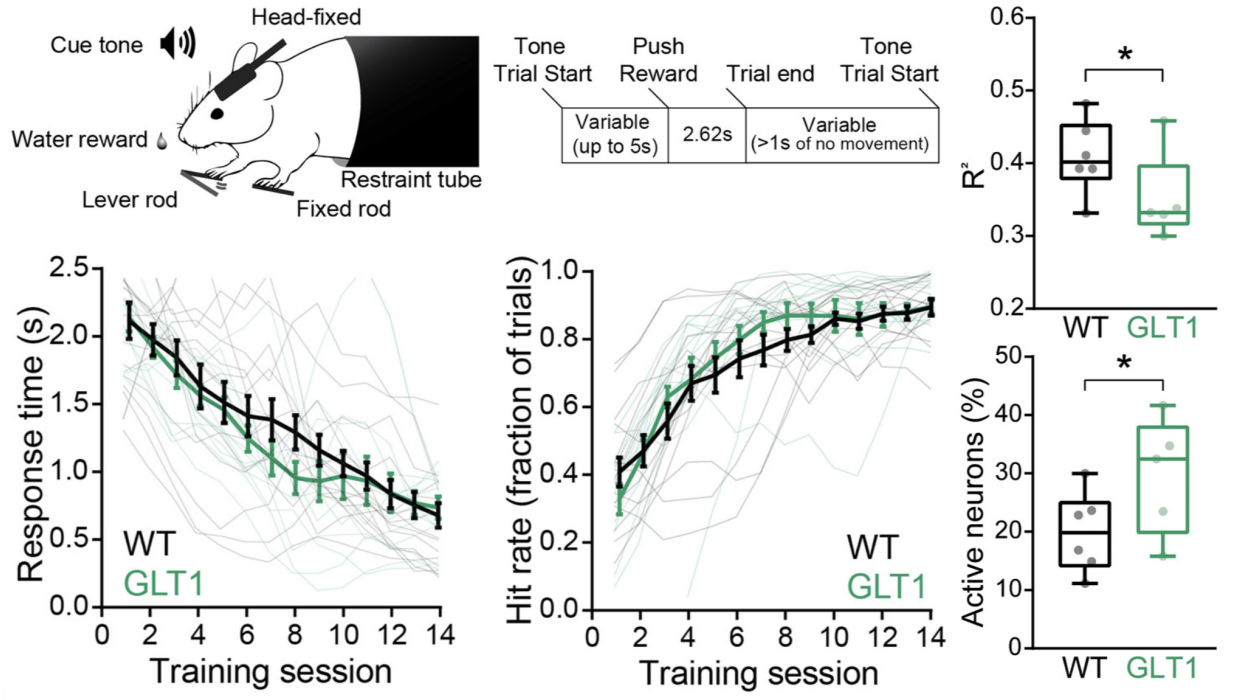


Figure 5: Astrocytes regulate the local neural ensemble in response to the single-neuron representational drift.

3.3 Mechanism for Drift Control

Crucially, combining with the **Excitability-Drift Hypothesis**, this "scheduling" capability can be extended beyond initial learning to the continuous maintenance of the neural code. If drift is indeed caused by stochastic fluctuations in intrinsic excitability, astrocytes likely mitigate this by imposing Homeostatic Gain Control on the population level:

- **Damping Volatility:** By buffering extracellular K^+ and clearing glutamate, astrocytes clamp the physiological range in which neurons can fluctuate. This prevents the intrinsic oscillations of single neurons from becoming large enough to push the population state out of the stable manifold (i.e., crossing from Null-space to coding-space).
- **Compensatory Normalization:** Astrocytes also work as a biological layer normalization [14] to monitor the overall neural excitability in this local area, and maintain the summation of the neurons in this area in a certain range. When a subset of neurons inevitably "drifts out" of the ensemble due to lowered excitability, astrocytes may sense the resulting drop in local network activity and compensatorily increase the excitability of neighboring neurons. This ensures that the total population excitability remains constant.

4 Potential Experimental Validation

To empirically validate the hypothesis that astrocytes actively regulate representational drift and confine it within the output-null space, I propose extending the experimental paradigm established by the Sur lab [4]. If the hypothesis is correct, disrupting astrocytic regulation should uncouple neural drift from behavioral stability. Consequently, we predict that deactivating astrocytic function will cause the neural population to drift out of the null-space, resulting in rapid behavioral regression (catastrophic forgetting).

The proposed experimental design is as follows:

1. **Acquisition Phase:** Train two cohorts of mice (Control vs. Experimental) on an operant conditioning task, such as a lever-press water reward task. Training continues until both groups achieve stable performance, characterized by a plateau in hit rate, reaction time, and kinematic stereotyping.

2. **Manipulation and Retention Interval:** Following mastery, cease all training to prevent further continuous learning. In the **Experimental Group**, inhibit astrocytic function (specifically targeting GLT1 glutamate transporters or using chemogenetic silencing). The **Control Group** receives a sham manipulation. Both groups are then housed normally for a defined retention period to allow natural representational drift to occur.

3. **Cross-Sectional Assessment:** To assess memory retention without the confounding factor of "re-learning" during testing, a cross-sectional approach is employed. At periodic intervals (e.g., Day 3, 7, 14), a distinct subset of mice from each group is subjected to a recall test. Following this single test session, these subjects are excluded from subsequent time-point analyses.

Expected Outcome: I hypothesize that the Experimental Group (GLT1-deficient) will exhibit a significantly faster rate of performance decay compared to the Control Group. In the absence of astrocytic regulation, we postulate that stochastic fluctuations in neuronal excitability will drive the population activity out of the output-null space. This unconstrained drift will degrade the neural representation of the task, manifesting as a failure to retrieve the learned behavior despite the absence of external interference.

5 Potential Applications in iBCIs: Towards a Hybrid Architecture

Recent breakthroughs have challenged the neuron-centric view of cortical coding. Studies demonstrate that astrocytes exhibit tuned responses analogous to neuronal cosine tuning, with evidence showing even sharper directional selectivity in the visual cortex (V1) compared to neurons [18]. With the advent of advanced optical interfaces for calcium imaging, utilizing astrocytic signals for next-generation iBCIs becomes a theoretical possibility. However, relying solely on astrocytes presents significant engineering hurdles:

- **High-Dimensional Signal Multiplexing:** Astrocytes exhibit "mixed selectivity," encoding not only the instantaneous kinematic variables (e.g., preferred direction) but also integrating information regarding behavioral history, reward prediction errors, and metabolic demand [15]. Although information redundancy reveals more acuate neural state, algorithms must separate the simple true motor intent from history effects, making this redundancy double-edged. The encoding of trial success or error information could effectively serve as an intrinsic "ground truth" signal for reinforcement learning-based BCI adaptation.

- **Prohibitive Temporal Latency:** The most critical limitation is the timescale mismatch. Astrocytic calcium transients operate on a timescale of seconds, with response lags up to 4–5 seconds relative to stimulus onset [15]. This latency renders an astrocyte-only system incapable of closing the loop for real-time, high-bandwidth motor control (e.g., cursor movement), which requires millisecond-level precision.

To bridge this gap, I propose a **Neuron-Astrocyte Hybrid BCI (NAH-BCI)** architecture. Rather than treating astrocytes or neurons as a direct source of kinematic control signals, this system utilizes astrocytes as a **"Contextual State Estimator."**

In this framework:

1. **The Fast Stream (Neurons):** A traditional electrode array decodes high-frequency spiking activity for real-time motor execution.

2. **The Slow Stream (Astrocytes):** Simultaneous optical or chemical recording of astrocytes tracks the slowly fluctuating "neural state variables," such as alertness, fatigue, or plasticity levels.

Mechanism of Integration: Astrocytes function as a biological "meta-learner." Since they govern the recruitment of neural ensembles (as discussed in the Excitability-Drift Hypothesis), their activity pattern acts as a **Prior information**. By decoding the current astrocytic state, the BCI can predict *which* subset of neurons is currently highly excitable and likely to be recruited. This information allows the decoder to dynamically re-weight the neural inputs, ignoring the "drifted-out" neurons and prioritizing the "recruited" ones—thereby achieving long-term robustness without frequent recalibration.

6 Conclusion

In conclusion, the paradox of stable behavior emerging from drifting neural representations represents a fundamental challenge for both neuroscience and neuroengineering. This report proposes that this stability is not an intrinsic property of neuronal circuits alone, but a feature of **neuro-glial interaction**. Astrocytes act as the critical "manifold stabilizer" of the cortex by dynamically modulating intrinsic neuronal excitability and integrating local population activity. These mechanisms work in concert to confine stochastic neural drift within the null-space, ensuring that functional readouts remain robust despite the continuous change of individual neurons. The hypothesis could be validated through the proposed GLT1-manipulation experiments to provide the causal link between glial regulation and manifold stability.

Furthermore, translating these biological insights into engineering, through the **Neuron-Astrocyte Hybrid BCI** architecture, offers a promising path toward overcoming the "decoder degradation" challenge. Ultimately, shifting from a neuron-centric view to a holistic neuro-glial perspective is essential for decoding the brain's capacity for long-term resilience, building the next generation of robust brain-computer interfaces.

References

[1] Mark M Churchland, John P Cunningham, Matthew T Kaufman, Justin D Foster, Paul Nuyujukian, Stephen I Ryu, and Krishna V Shenoy. Neural population dynamics during reaching. *Nature*, 487(7405):51–56, 2012.

[2] Alan D Degenhart, William E Bishop, Emily R Oby, Elizabeth C Tyler-Kabara, Steven M Chase, Aaron P Batista, and Byron M Yu. Stabilization of a brain–computer interface via the alignment of low-dimensional spaces of neural activity. *Nature biomedical engineering*, 4(7):672–685, 2020.

[3] Geoffroy Delamare, Yosif Zaki, Denise J Cai, and Claudia Clopath. Drift of neural ensembles driven by slow fluctuations of intrinsic excitability. *eLife*, 12:RP88053.

[4] Chloe Delepine, Jennifer Shih, Keji Li, Pierre Gaudeaux, and Mriganka Sur. Differential Effects of Astrocyte Manipulations on Learned Motor Behavior and Neuronal Ensembles in the Motor Cortex. *Journal of Neuroscience*, 43(15):2696–2713, April 2023. Publisher: Society for Neuroscience Section: Research Articles.

[5] Laura N Driscoll, Noah L Pettit, Matthias Minderer, Selmaan N Chettih, and Christopher D Harvey. Dynamic reorganization of neuronal activity patterns in parietal cortex. *Cell*, 170(5):986–999, 2017.

[6] Shaul Druckmann and Dmitri B Chklovskii. Neuronal circuits underlying persistent representations despite time varying activity. *Current Biology*, 22(22):2095–2103, 2012.

[7] Gabrielle T. Drummond, Arundhati Natesan, Marco Celotto, Jennifer Shih, Prachi Ojha, Yuma Osako, Jiho Park, Grayson O. Sipe, Kyle R. Jenks, Vincent Breton-Provencher, Paul C. Simpson, Stefano Panzeri, and Mriganka Sur. Cortical norepinephrine-astrocyte signaling critically mediates learned behavior, October 2024.

[8] Apostolos P Georgopoulos, John F Kalaska, Roberto Caminiti, and Joe T Massey. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *Journal of Neuroscience*, 2(11):1527–1537, 1982.

[9] Christian Henneberger, Thomas Papouin, Stéphane HR Oliet, and Dmitri A Rusakov. Long-term potentiation depends on release of d-serine from astrocytes. *Nature*, 463(7278):232–236, 2010.

[10] Thomas Hosman, Tsam Kiu Pun, Anastasia Kapitonava, John D Simeral, and Leigh R Hochberg. Months-long high-performance fixed lstm decoder for cursor control in human intracortical brain-computer interfaces. In *2023 11th international IEEE/EMBS conference on neural engineering (NER)*, pages 1–5. IEEE, 2023.

[11] Tommy Hosman, Marco Vilela, Daniel Milstein, Jessica N Kelemen, David M Brandman, Leigh R Hochberg, and John D Simeral. Bci decoder performance comparison of an lstm recurrent neural network and a kalman filter in retrospective simulation. In *2019 9th International IEEE/EMBS conference on neural engineering (NER)*, pages 1066–1071. IEEE, 2019.

[12] Brianna M Karpowicz, Yahia H Ali, Lahiru N Wimalasena, Andrew R Sedler, Mohammad Reza Keshtkaran, Kevin Bodkin, Xuan Ma, Daniel B Rubin, Ziv M Williams, Sydney S Cash, et al. Stabilizing brain-computer interfaces through alignment of latent dynamics. *Nature Communications*, 16(1):4662, 2025.

[13] Paulo Kofuji and EA2322935 Newman. Potassium buffering in the central nervous system. *Neuroscience*, 129(4):1043–1054, 2004.

[14] Leo Kozachkov, Ksenia V. Kastanenka, and Dmitry Krotov. Building transformers from neurons and astrocytes. *Proceedings of the National Academy of Sciences*, 120(34):e2219150120, August 2023. Publisher: Proceedings of the National Academy of Sciences.

[15] Nancy Ann Oberheim, Takahiro Takano, Xiaoning Han, Wei He, Jane HC Lin, Fushun Wang, Qiwu Xu, Jeffrey D Wyatt, Webster Pilcher, Jeffrey G Ojemann, et al. Uniquely hominid features of adult human astrocytes. *Journal of Neuroscience*, 29(10):3276–3287, 2009.

[16] Olivier Pascual, Kristen B Casper, Cathryn Kubera, Jing Zhang, Raquel Revilla-Sanchez, Jai-Yoon Sul, Hajime Takano, Stephen J Moss, Ken McCarthy, and Philip G Haydon. Astrocytic purinergic signaling coordinates synaptic networks. *Science*, 310(5745):113–116, 2005.

[17] Michael E Rule, Timothy O'Leary, and Christopher D Harvey. Causes and consequences of representational drift. *Current Opinion in Neurobiology*, 58:141–147, October 2019.

[18] James Schummers, Hongbo Yu, and Mriganka Sur. Tuned Responses of Astrocytes and Their Influence on Hemodynamic Signals in the Visual Cortex. *Science*, 320(5883):1638–1643, June 2008. Publisher: American Association for the Advancement of Science.

[19] Wei Wu, Yun Gao, Elie Bienenstock, John P Donoghue, and Michael J Black. Bayesian population decoding of motor cortical activity using a kalman filter. *Neural computation*, 18(1):80–118, 2006.