bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Unraveling Network Dynamics via Riemannian Multi-Scale Decomposition

Yonatan Kleerekoper
Technion, Israel Institute of Technology
`yonatan.kle@campus.technion.ac.il`

Mohammad Kurtam
Technion, Israel Institute of Technology
`kurtam@campus.technion.ac.il`

Yonatan Keselman
Technion, Israel Institute of Technology
`ykeselman@campus.technion.ac.il`

Amir Ghanayim
Technion, Israel Institute of Technology
`amir.gh122@campus.technion.ac.il`

Jackie Sciller
Technion, Israel Institute of Technology
`jackie@technion.ac.il`

Simon Musall
Juelich, Research Center
`s.musall@fz-juelich.de`

Yitzhak Schiller
Technion, Israel Institute of Technology
`syitzhak@technion.ac.il`

Hadas Benisty
Technion, Israel Institute of Technology
`hadasbe@technion.ac.il`

Abstract

Understanding the temporal dynamics of high-dimensional, time-varying systems is a fundamental challenge across scientific disciplines. Existing methods primarily model the temporal evolution of sensor readouts while often assuming that the correlations between different components of the observed signal are static or neglecting them entirely. However, in many real-world systems, these interactions are dynamic, adding substantial complexity to the extraction of meaningful representations and the design of downstream models. In this work, we introduce a novel multi-resolution framework for analyzing temporal sequences of Symmetric Positive Semi-Definite (SPSD) matrices, such as correlation matrices. Our method constructs a hierarchical representation by recursively applying similarity and difference operators, respecting the underlying Riemannian geometry of the SPSD manifold. We further employ spectral analysis to identify subgroups that drive correlational dynamics. By extracting these key sub-networks, we denoise the signal and, importantly, promote the interpretability of time-varying systems. We validate our approach using both synthetic data and real neural recordings across modalities, including Calcium imaging and EEG, and at different scales, from local dendritic networks to cortical-level networks. Our results show that the proposed method extracts meaningful structures from complex temporal datasets and offers new insights into evolving network connectivity in high-dimensional systems.

1 Introduction

In many real-world experiments, we collect large-scale recordings of dynamic signals in an effort to uncover the underlying structure of the system that generated them. The challenge becomes even more profound when the intrinsic properties of the system, such as connectivity patterns governing

Preprint. Under review.

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

dynamics, or functional organization, evolve over time. In such cases, modeling both the observed signals and the dynamics of the underlying structure is essential for gaining meaningful insights.

In neuroscience, for example, learning induces plastic changes in the connectivity between neurons and brain regions to optimize behavioral performance. Synaptic plasticity cannot be directly measured in behaving animals. Instead, researchers use correlations between neural recordings, estimated using short segments of the data, as proxies for functional connectivity [4, 19, 14, 2, 39, 1]. This motivates the study of correlation dynamics as a means to identify which sub-networks are actively driving change and are thus most involved in learning-related reorganization.

A fundamental challenge is that correlation matrices are Symmetric Positive Semi-Definite (SPSD) and lie on a non-Euclidean manifold, making standard linear operations inapplicable in this space. Recent work has leveraged Riemannian geometry to perform tasks such as temporal analysis [35, 3, 16, 17], feature selection [9], and deep learning [20, 40, 25, 32]. Notably, [35] proposed a Riemannian wavelet-like framework based on the recursive application of the Riemannian mean as a low-pass filter. However, their method does not support recursive high-pass filtering, thus dramatically limiting the frequency resolution of the decomposed signal.

We address this gap by introducing a multi-resolution, time-frequency analysis for a temporal sequence of SPSD matrices. Our key innovation is a difference operator that preserves symmetry and positive semi-definiteness, allowing both low-pass and high-pass operations to be recursively applied on the manifold. This yields a Haar-like, multi-scale decomposition on the Riemannian manifold. Using this representation, we further propose a spectral method for identifying which components of the original observed signal drive the correlational dynamics. Our contributions are:

- A novel difference operator for SPSD matrices that preserves manifold structure.
- A theoretical characterization of the spectral behavior and stability of the difference operator.
- A multi-scale decomposition of SPSD matrices based on recursive application of Haar-like filtering on the SPSD manifold.
- A spectral algorithm for identifying dynamic drivers of correlational change.

By extending signal processing tools to the SPSD manifold, we provide a principled and scalable framework for multi-resolution temporal analysis of structured, symmetric data. We validate our method on both synthetic and real neural data, including Calcium imaging (both dendritic and somatic), widefield cortical recordings, and EEG. Our method uncovers key elements involved in correlational reorganization, offering interpretable and biologically relevant insights.

2 Related works

2.1 Time-varying graphs

Time varying graphs have emerged as powerful tools for capturing such dynamic relationships and have been successfully applied in diverse domains, including social networks [30, 41, 7, 23], traffic forecasting [42, 27], recommender systems [31, 12], and neuroscience [36]. Recent developments have leveraged neural network architectures to model temporal graphs, achieving impressive empirical results. However, many of these methods overlook a key geometric structure: when interactions are symmetric, i.e., graph edges are bidirectional, the resulting connectivity matrices are SPSD.

2.2 Dynamics of functional connectivity for neuroscience

Understanding connectivity mechanisms in the brain has been a growing focus in neuroscience [18, 8, 29, 26]. Recent studies have explored functional connectivity between cells and brain regions and their relationship to behavioral and cognitive states [34, 33]. These works refer to functional connectivity as a static entity by averaging over large segments of data, corresponding to a specific state [22]. Graph theory methods are widely used by the fMRI community for the study of functional connectivity [21, 38]. Much of the existing work tracks specific features of the network, informing the individual role of each node to the rest of the network [6]. However, this type of analysis does not capture the structure of the overall network, as it collapses functional connectivity into a set of centrality scores.

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Other works combine graph theory with Riemannian geometry to explore the relation between the dynamics of functional connectivity and behavior. For example, [4] shows that spontaneous behavior can be meaningfully represented through the dynamics of the correlation between cortical brain regions. [19] investigates the role of dopamine signaling in a cellular network organization during motor learning. [24] utilizes the Riemannian mean to classify different human activities during task performance based on EEG recordings. While gaining meaningful insights and improving model robustness, these works are limited in the sense that they are unable to identify the key elements driving changes in the correlational structure of the networks.

2.3 Multi-resolution Riemannian analysis

Wavelet transforms have long been a fundamental tool for multiscale analysis [13], yet extending them to Riemannian manifolds poses significant challenges. The work by [35] is a notable advance in this direction, defining two key operators between SPSD matrices: a similarity operator that preserves shared structural components and a difference operator that captures differences between them. This method has a key limitation: the difference operator maps to the tangent space, and its output is not guaranteed to remain within the manifold. Consequently, it cannot be applied recursively, and the decomposition must rely solely on repeated low-pass operations. This precludes the construction of a full wavelet packet transform and limits the resolution at which dynamic features can be identified. Our approach overcomes this limitation by introducing a novel difference operator that yields an SPSD matrix as output. This key property enables both low-pass and high-pass filtering to be applied recursively without leaving the Riemannian manifold.

3 Methods

3.1 Problem formulation

Let $\mathbf{x}[t] \in \mathbb{R}^d$ be observations of a time-varying system. We assume the intrinsic properties of the system vary slowly so that pairwise interactions between components of $\mathbf{x}[t]$ are approximately constant within a short time window of T_w samples. We slide this window along the observed time series and extract a series of correlation matrices $\{\mathbf{C}_n\}_{n=1}^N$, where n denotes the segment index. This sequence of SPSD matrices represents the temporal evolution of the correlational structure of the observed signal. Such data is common in neuroscience, where neural activity is recorded as subjects learn to perform a task. The temporal index t corresponds to neural sampling rate, and the segment index n corresponds to behavioral epochs such as trials, for example.

Two central questions arise: 1) Can we identify a low-dimensional representation for the sequence of matrices to infer specific dynamics? 2) Which components in the original high-dimensional space drive these dynamics? The first question is typically answered using Riemannian Geometry [5] and manifold learning techniques such as diffusion map [10]. The second question is more challenging. Most nonlinear and Riemannian embedding methods are not interpretable, as they typically do not offer a well-defined mapping between points in the original high-dimensional space and their representations in the embedded space. To address this, we introduce a principled multi-resolution framework, inspired by Wavelet Packet Decomposition (WPD), that enables interpretable analysis of SPSD matrix sequences while preserving Riemannian structure.

3.2 Low-pass and high-pass filtering on the SPSD manifold

A key aspect of WPD is the choice of low-pass and high-pass filters, where Haar filters are a common choice[13]:

$$h_{\text{LP}}[n] = \frac{\sqrt{2}}{2}[1,1], \quad h_{\text{HP}}[n] = \frac{\sqrt{2}}{2}[-1,1] \tag{1}$$

Inspired by eqn. (1), we construct our decomposition based on a similarity and a difference operator, both preserving SPSD properties. Let $\mathbf{W}_1$ and $\mathbf{W}_2$ be two SPSD matrices. Let $\gamma(p)$ be the geodesic connecting them, meaning that $\mathbf{W}_1 = \gamma(0)$ and $\mathbf{W}_2 = \gamma(1)$. We follow [35] and use the similarity operator, defined by the midpoint on the geodesic:

$$\mathbf{W}_1 \# \mathbf{W}_2 \triangleq \gamma(0.5) = \mathbf{W}_1^{0.5}\left(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5}\right)^{0.5}\mathbf{W}_1^{0.5} \tag{2}$$

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

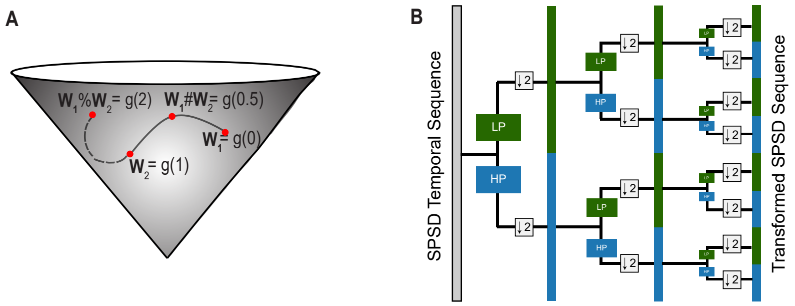

Figure 1: (A) A schematic illustration of the geodesic connecting two matrices where the similarity operator is positioned at the midpoint, while the difference operator extends the geodesic such that $\mathbf{W}_2$ becomes the midpoint between $\mathbf{W}_1\%\mathbf{W}_2$ and $\mathbf{W}_1$. (B) An illustration of the proposed analysis: recursive application of low-pass and high-pass filters, followed by down-sampling by a factor of two.

Applying this operator to a sequence of SPSD matrices results in the Riemannian mean of every consecutive pair:

$$\mathbf{C}_n^{\mathrm{LP}} = \mathbf{C}_n \# \mathbf{C}_{n+1}, \quad \forall n = 1, ..., N-1 \tag{3}$$

We note that the filtered sequence, $\mathbf{C}_n^{\mathrm{LP}}$, is also SPSD, since every matrix is the mid-point on the geodesic between two other SPSD matrices. This property is fundamental, since WPD is based on the recursive application of these filters. Analogous to the Haar high-pass filter, we propose the following difference operator:

$$\mathbf{W}_1\%\mathbf{W}_2 \triangleq \gamma(2) = \mathbf{W}_1^{0.5}\left(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5}\right)^2\mathbf{W}_1^{0.5} \tag{4}$$

Proposition 1. For any pair of SPSD matrices $\mathbf{W}_1$ and $\mathbf{W}_2$, the geodesic can be extended from $\gamma(-\infty)$ to $\gamma(\infty)$. Moreover, the SPSD manifold is geodesically complete. [5]

Proposition 1 ensures that the resulting matrix is indeed SPSD, allowing us to apply it recursively. Figure 1A illustrates the geometric interpretation: the similarity operator lies at the midpoint of the geodesic. In contrast, the output of the difference operator is the extension of the geodesic such that $\mathbf{W}_2$ is the midpoint, thus making the similarity and difference operators mutually inverse. (See a formal proof of Proposition 2 in Appendix 7.2.) We note that this operator is not symmetric, meaning that in general, $\mathbf{W}_1\%\mathbf{W}_2 \neq \mathbf{W}_2\%\mathbf{W}_1$. Importantly, our spectral analysis shows that this operator enhances differences between $\mathbf{W}_1$ and $\mathbf{W}_2$ (Sec. 4, Theorems 1-4). We, therefore, use it to define the following high-pass operation:

$$\mathbf{C}_n^{\mathrm{HP}} = \mathbf{C}_n \% \mathbf{C}_{n+1}, \quad \forall n = 1, ..., N-1 \tag{5}$$

We further explore the effect of the similarity and difference operators as low- and high-pass filters and their resemblance to the classical (Euclidean 1D) Haar wavelet functions in Appendix 7.4.

3.3 Riemannian multi-resolution decomposition

Inspired by standard WPD, we extract a multi-resolution decomposition of a sequence of SPSD matrices. Given a temporal sequence of N SPSD matrices, we apply the low- and high-pass operators defined in eqns. (3) and (5), followed by down-sampling by a factor of two. This splits the sequence of matrices into low- and high-frequency sequences, each of length $N/2$. Recursively repeating this process for $\log_2(N)$ steps results in a hierarchical decomposition where the last level consists of a series of SPSD matrices corresponding to N frequency bins denoted by $\left\{\hat{C}_f\right\}_{f=1}^{N}$, as illustrated in Fig. 1B. In the next section, we utilize this decomposition and propose an algorithm for detecting elements driving the dynamics.

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3.4 Detection of dynamics drivers

Our proposed decomposition yields a sequence of matrices, each corresponding to a frequency bin. In this section, we propose a spectral analysis algorithm to detect informative bins and identify which components of $\mathbf{x}[t]$ drive this dynamics. First, we extract the dominant eigenvector ψ_f corresponding to the largest eigenvalue:

$$\hat{C}_f \psi_f = \lambda_f \psi_f \tag{6}$$

For each eigenvector, we quantify how informative it is using the normalized entropy of its absolute values:

$$h_f \triangleq -\frac{1}{\log d} \sum_{i=1}^{d} p_f(i) \log p_f(i), \quad p_f(i) = \frac{|\psi_f(i)|}{\sum_{i=1}^{d} |\psi_f(i)|} \tag{7}$$

We then select frequency bins by retaining eigenvectors with low entropy and lower frequency index, as these typically represent structured slow dynamics. Next, we apply k-means (k=2) to each eigenvector to classify its components into high- or low-absolute value. Finally, we compute the voting score for each component across the high-value cluster and identify elements that are consistently voted for as dynamic drivers. This process is summarized in Algorithm 1.

Algorithm 1 Riemannian Multi-Scale Decomposition for Dynamics Detection

```
Require: SPSD matrix sequence $C_n$ of length N, cutoff frequency $f_{cutoff}$, cutoff entropy $h_{cutoff}$
    and a voting threshold $M$.
Construct wavelet packet decomposition:
for $k = 1$ to $\log_2(N)$ do
    for each node in level $k-1$ do
        Apply low-pass and high-pass filtering using eqns. (3) and (5)
        Downsample the resulting signals by 2
    end for
end for
for each frequency bin $f$ do
    Extract the eigenvector corresponding to the highest eigenvalue. eqn. (6).
    Compute the normalized entropy $h_f$ using eqn. (7).
end for
Create a subset of eigenvectors $\{\psi_f^*\} = \{\psi_f | f < f_{cutoff}, h_f \leq h_{cutoff}\}$
Apply k-means to each retained eigenvector to yield an indicator such that $\mathbb{1}_f[i] = 1$ if the
    element $i$ in frequency bin $f$ was clustered with the high absolute value and $\mathbb{1}_f[i] = 0$ otherwise.
Identify the dynamic drivers subset. $D = \{i | \sum_f \mathbb{1}_f[i] > M\}$
```

We demonstrate the complete pipeline on a controlled synthetic dataset in Section 5.1, highlighting how the proposed decomposition and dynamic driver detection recover the underlying structure.

4 Spectral analysis

In this section, we provide a theoretical analysis of the spectral behavior of the low-pass ($\#$) and high-pass ($\%$) operators. Specifically, we demonstrate how the shared and distinct structures of the input SPSD matrices are reflected in the spectra of the filtered matrices, and we provide further analysis on the stability of the operators under small perturbations.

4.1 Strictly common components

The first two theorems discuss the case where $\mathbf{W}_1$ and $\mathbf{W}_2$ share a common eigenvector ψ, but with different eigenvalues.

Theorem 1. Let ψ be an eigenvector of both $\mathbf{W}_1$ and $\mathbf{W}_2$ with eigenvalues λ_1 and λ_2, respectively: $\mathbf{W}_1\psi = \lambda_1\psi, \quad \mathbf{W}_2\psi = \lambda_2\psi$. Then ψ is also an eigenvector of the similarity operator $\mathbf{S} = \mathbf{W}_1 \# \mathbf{W}_2$, with eigenvalue: $\lambda_S = \sqrt{\lambda_1\lambda_2}$.

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

This result, first shown in [35], implies that the low-pass operator preserves the spectral structure of shared components. If ψ is dominant in both matrices, it remains dominant in their geometric mean.

Theorem 2. Under the same assumptions as above, ψ is also an eigenvector of the high-pass operator $\mathbf{Q} = \mathbf{W}_1 \% \mathbf{W}_2$, with eigenvalue: $\lambda_Q = \lambda_1^{-1}\lambda_2^2$.

Proof. See Appendix 7.1.

This shows that the high-pass operator amplifies differences in eigenvalues between the two matrices. That is, if two matrices share a component ψ but with differing strength ($\lambda_1 \neq \lambda_2$), then ψ is highlighted in $\mathbf{Q}$.

Together, these theorems reveal that low-pass filtering emphasizes commonly expressed structure, while high-pass filtering emphasizes variation, supporting their use in multi-scale correlational analysis.

4.2 Stability under weak perturbations

In practice, shared structure across matrices may only be approximate. The following analysis demonstrates that our operators are stable under small perturbations in shared eigenvectors, indicating that they behave well in the presence of noise. We denote the ϵ-pseudo-spectrum, which captures how eigenvalues shift under perturbations as defined by [37]:

Definition 1. Let $\mathbf{W} \in \mathbb{R}^{N\times N}$, the following definitions of the ϵ-psuedo-spectrum are equivalent for a small $\epsilon > 0$:

1. The set of $\lambda \in \mathbb{R}$ such that $\|(\lambda\mathbf{I} - \mathbf{W})^{-1}\| \geq \epsilon^{-1}$
2. The set of $\lambda \in \mathbb{R}$ such that $\lambda \in \Lambda(\mathbf{W} + \mathbf{E})$ for some $\mathbf{E} \in \mathbb{R}^{N\times N}$ with $\|\mathbf{E}\| \leq \epsilon$
3. The set of $\lambda \in \mathbb{R}$ such that $\|(\lambda\mathbf{I} - \mathbf{W})\mathbf{v}\| \leq \epsilon$ for some $\mathbf{v} \in \mathbb{R}^N$ with $\|\mathbf{v}\| = 1$

The following two theorems extend our spectral results to the case of approximately common components, showing that pseudo-eigenstructure is preserved under our operators.

Theorem 3. [35] Suppose there exists an eigen-pair (λ_k, ψ_k) of $\mathbf{W}_k$ for $k = 1, 2$ so that $\psi_1 = \psi_2 + \psi_\epsilon$, where $\|\psi_\epsilon\|_2 \leq \frac{\sqrt{\lambda_2}}{\tilde{\lambda}_{2\,\max}\sqrt{\lambda_1}}\epsilon$ for a small $\epsilon > 0$, where $\tilde{\lambda}_{2\,\max} = \|\mathbf{W}_2 - \lambda_2\mathbf{I}\|$. Then we have: $\sqrt{\lambda_1\lambda_2} \in \Lambda_\epsilon(\mathbf{S})$. Specifically, we have: $\|(\mathbf{S} - \sqrt{\lambda_1\lambda_2}\mathbf{I})\psi_1\|_2 \leq \epsilon$, and ψ_1 is a corresponding ϵ–pseudo-eigenvector of $\mathbf{S}$.

Theorem 4. Suppose there exists an eigenpair λ_k and ψ_k of $\mathbf{W}_k$ for $k = 1, 2$ so that $\psi_1 = \psi_2 + \psi_\varepsilon$, where $\|\psi_\varepsilon\|_2 \leq \frac{1}{(\sigma_{\max}(\mathbf{W}_2{\mathbf{W}_1}^{-1}) + \lambda_2\lambda_1^{-1})\tilde{\lambda}_2^{\max}}\varepsilon$, where $\sigma_{\max}$ denotes the maximum eigenvalue, for a small $\varepsilon > 0$, and where $\tilde{\lambda}_2^{\max} = \|\mathbf{W}_2 - \lambda_2\mathbf{I}\|$. Then we have $\lambda_2^2\lambda_1^{-1} \in \Lambda_\varepsilon(\mathbf{Q})$. Specifically, we have $\|(\mathbf{Q} - \lambda_2^2\lambda_1^{-1}\mathbf{I})\psi_1\|_2 \leq \varepsilon$, and ψ_1 is a corresponding ε-pseudo-eigenvector of $\mathbf{Q}$.

Proof. See Appendix 7.1.

These results demonstrate that our operators are robust to perturbations in shared structure—a desirable property when working with noisy real-world data, such as neural recordings.

4.3 Implications for our framework

These theoretical results justify the core design of our method. The similarity operator ($\#$) preserves shared structure and suppresses noise, while the difference operator ($\%$) amplifies subtle variations. Their recursive application produces a multi-resolution decomposition for a sequence of SPSD matrices. Moreover, spectral analysis provides interpretability to the structure of each frequency bin in the decomposed sequence: dominant eigenvectors reflect coherent activity, and their entropy can be used to identify components with consistent dominance.

5 Results

We evaluate the performance of our framework on a simulated toy dataset and several real-world neural datasets. Due to the algorithm's short runtime, all experiments were run on a CPU. However,

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

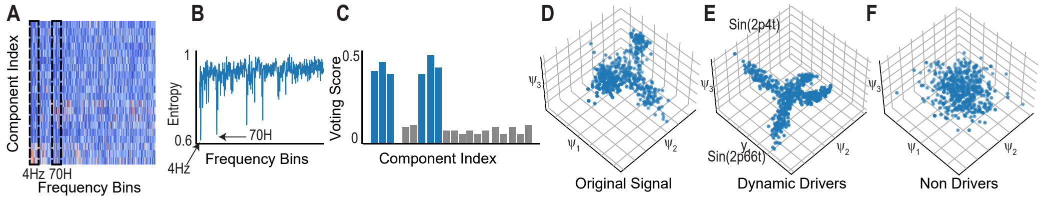

Figure 2: Toy example: a simulated sequence of SPSD matrices. (A) The first eigenvector of the decomposed sequence expressing the distribution across components and frequency bins. (B) Entropy as a function of frequency bins. (C) Voting scores across components, blue - true dynamic drivers, gray - non-drivers. (D-F) Diffusion embedding of the sequence of matrices considering: D, all elements (original data), E, dynamics drivers, F, non-drivers only.

our geometric low- and high-pass operators are inherently parallelizable and can be efficiently implemented on GPUs for further acceleration.

5.1 Toy example

We construct a sequence of $20X20$ SPSD matrices. Two disjoint subsets of entries (indices 1–3 and 7–9) were assigned low- and high-frequency sinusoidal dynamics, while all other off-diagonal entries were sampled from a uniform distribution:

$$\begin{aligned} \mathbf{C}_n\left(k,l\right) &= 0.5\sin\left(2\pi f_1 n\right) + U\left(-0.25, 0.25\right), \quad k,l \in \{1,2,3\}, \quad n \leq 256 \\ \mathbf{C}_n\left(k,l\right) &= 0.5\sin\left(2\pi f_2 n\right) + U\left(-0.25, 0.25\right), \quad k,l \in \{7,8,9\}, \quad n > 256 \\ \mathbf{C}_n\left(k,l\right) &\sim U\left(-0.25, 0.25\right), \quad otherwise; \quad \mathbf{C}_n\left(k,k\right) = 1 \end{aligned} \tag{8}$$

with $f_1 = 4Hz$ and $f_2 = 66Hz$. Applying our algorithm, we identify a distinct pattern within the leading eigenvectors indicating elements 1-3 at 4Hz and elements 7-9 at approximately 70Hz, as illustrated in Figure 2A. This pattern is reflected by low-entropy values in these frequencies 2B. Moreover, our framework successfully identifies the elements driving the sinusoidal dynamics, as illustrated in Fig. 2C. We compared the original SPSD sequence to a sequence that included components identified as dynamic drivers, and a third sequence that retained the rest of the network (i.e., components not identified as dynamic drivers). We extracted a low-dimensional representation for each sequence using Riemannian distances [5] and diffusion map [10] presented in Figure 2D-F. The dynamic drivers preserved the intrinsic geometry and recovered a clean two-segment manifold reflecting the intrinsic properties of the signal 2E. In contrast, excluding them resulted in unstructured noise 2F. Overall, our proposed analysis successfully captured the intrinsic structure of this system, identifying its dynamic drivers while effectively removing noise.

5.2 Dendritic networks in the primary motor cortex (M1) during motor learning

We applied our method to Calcium imaging data from 8 mice as they were trained to perform a lever-pull task over 15 days. Regions of interest (ROIs) were manually identified from Layer 5 dendritic trees from the primary motor cortex (M1). Each trial lasted 12 seconds, with 30 trials per session (for a more detailed description of the experimental setup, see Appendix 7.5). We computed a correlation matrix per trial across all ROIs, resulting in hundreds of SPSD correlation matrices reflecting the correlational dynamics during learning across all training sessions. We applied the proposed analysis and identified the dynamic drivers, i.e., a sub-group of dendrites per animal. We then re-computed the correlation matrices using either dynamic drivers or non-drivers (the rest) and extracted the first diffusion map component for both sequences. We evaluated the variance of the first diffusion map component across time, as depicted in Fig. 3B: the dynamic drivers (blue) exhibit higher temporal variability than the rest of the network (gray), suggesting their greater involvement in functional reorganization during learning (p=0.002, Wilcoxon rank-sum). To assess the behavioral relevance of the dendrites identified as dynamic drivers, we trained a linear model to predict lever

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

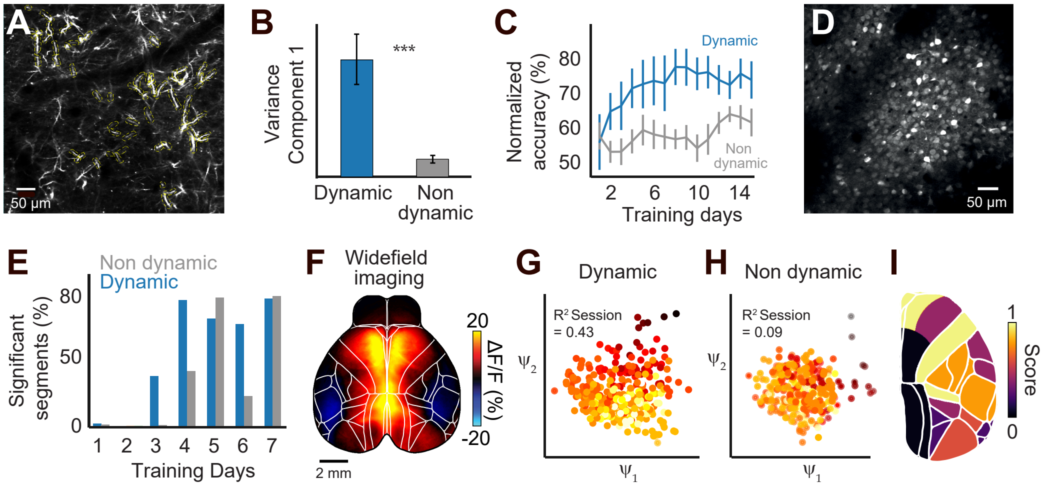

Figure 3: **Dendritic networks during motor learning**: (A) Example field of view from the Calcium imaging data, where white streaks indicate dendritic activity. (B) Variance across sessions of the first diffusion map component, average across 8 animals, showing the significantly greater temporal variability captured by the dynamic drivers compared to the rest. (C) Model accuracy R^2 for predicting lever position from neuronal activity across all training days, considering dynamics driving ROIs (blue), non-driving ROIs (gray). Normalized by the R^2 score achieved by all ROIs. For B and C, error bars indicate mean ± SEM, n=8 animals. **Large population networks during motor learning**: (D) Example field of view from Calcium imaging data, where white circles indicate neuronal activity of cell bodies. (E) Percentage of time windows in which outcome separation was detected (p-value < 0.01, permutation testing). **Widefield Calcium imaging during audio discrimination**:(F) Example data (one frame) for widefield Calcium data overlayed with the Allen Mouse Brain CCF. (G+H) The first two components of diffusion map, taking dynamic drivers, G, or non-drivers,H. Colors stand for the session index. (I) Importance score of each component on the mouse brain.

position based on the temporal activity of the dynamic drivers or based on non-drivers. As shown in Fig. 3C, dynamic drivers achieved higher R^2 values across sessions and animals compared to the non-drivers. This result indicates that the sub-population of dynamic drivers during learning was also involved in the emergence of behavioral encoding of the task.

5.3 Large population networks in M1 during motor learning

We next analyzed a high-density Calcium imaging dataset tracking the activity of 178 neurons in layer 2/3 of M1 during 7 sessions of training for a hand-reach task (for a detailed description of the experimental setup, see Appendix 7.6). Previous work showed that learning induces global network reorganization and emergence of outcome encoding, but did not isolate the responsible subpopulations [19]. Using our method, we identified 91 neurons (out of 178) as dynamic drivers in the 30 lowest-frequency bins. We quantified outcome separation using either dynamic drivers or the rest of the network using a PCA-based sensitivity index, as performed by [19]:

$$d_{prime}\left(t\right) = \sum_{i} \left(\frac{\mu_S^i\left(t\right) - \mu_F^i\left(t\right)}{\sigma_S^i\left(t\right)\sigma_F^i\left(t\right)} \right)^2$$

where μ_S and μ_F are the means of the i-th PC, average across successful and failed trials, respectively. Similarly, σ_S^i and σ_F^i are the corresponding standard deviations. Significant separation was determined using randomized permutation of outcome labels. We then aggregated our results over t and counted the fraction of time (within the trial) each population had separable dynamics by outcome, per day. Our findings show that the subpopulation identified by our algorithm as driving the correlational dynamics also exhibited an earlier and more consistent emergence of outcome separation, as illustrated in Fig. 3E. This supports the hypothesis that motor learning unfolds through

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1: Accuracy for left-right arm classification

Subject	Baseline accuracy	With feature selection	Number of features selected (out of 22)
S1	84	**90**	10
S2	50	**51**	4
S3	71	**85**	5
S4	50	**65**	8
S5	50	**54**	10
S6	55	**58**	5
S7	50	**52**	4
S8	**97**	94	6
S9	**92**	89	9
Average	66.5	**71.9**	-

gradual, low-frequency changes in network dynamics and highlights the utility of our method in identifying functional subpopulations involved in encoding behaviorally relevant variables.

5.4 Widefield imaging during learning of an audio discrimination task

We utilized our method for studying large-scale changes in correlation structure across cortical areas. We analyzed widefield Calcium imaging data from [28], where mice learned to lick left/right in response to auditory cues. We focused on right-hemisphere components extracted using the Allen Mouse Brain Common Coordinate Framework (CCF) and analyzed trials where animals correctly chose the right side. While the original paper by Musall et al. focused on fully trained expert animals, here we analyzed the sessions preceding this stage, as the animal was trained to perform the task (see Appendix 7.7 for a detailed description of the experimental setup). Our method identified 11 dynamic drivers (out of 18 components). To quantify relevance to the learning process, we computed the sequence of correlations, including either dynamic drivers or the non-drivers. As before, we extracted a low-dimensional representation for both sequences, as depicted in Fig. 3G and H. The gradient of colors in G, as opposed to the unordered structure in H, indicates that the sub-network of dynamic drivers better reflects long-term changes in correlation structure throughout task learning. To quantify this we trained a linear model for prediction of session index based on the first leading diffusion components. Using dynamic drivers, we achieved a fair accuracy value of $R^2 = 0.43$ while using the non-drivers yielded a much lower value of $R^2 = 0.09$. Using the overall network led to accuracy of $R^2 = 0.25$. The drivers sub-network alone captured the dynamics of the correlational embedding, demonstrating that it identified the key cortical areas that are driving learning-related changes in functional connectivity. Finally, we superimposed the importance scores of each component in the network (indicating whether it is a dynamic driver or not) to its spatial location in the brain mask, Fig. 3I. The parietal and frontal cortices emerge as key drivers of the observed correlational dynamics, consistent with their well-established roles in learning-related processes.

5.5 EEG classification in a BCI task

To further demonstrate the versatility of our approach to different modalities and spatio-temporal scales, we applied our framework to EEG dataset IIa from BCI Competition IV [15], focusing on motor imagery of left vs. right hand. Each trial was band-pass filtered (8–30Hz) and correlation matrices computed for a 3-second period starting from 0.5 secs before the go-cue. Dynamic drivers were identified separately for each movement type. We retained electrodes identified as drivers for either side and excluded those that were ambiguous or irrelevant. Using the selected electrodes, we computed diffusion embeddings and trained an RBF-SVM classifier with regularization parameter $C = 0.1$ to distinguish left from right. As shown in Table 1, classification accuracy on the test set improved in 7 out of 9 subjects. This demonstrates the method's ability to isolate task-relevant EEG channels for improved decoding.

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

6 Conclusion

We presented a novel multi-resolution framework for analyzing temporal sequences of SPSD matrices that fully respects their underlying Riemannian geometry. By defining recursive low- and high-pass filtering operations directly on the manifold, our method enables Haar-like wavelet packet decompositions of structured correlation data. Through spectral analysis of the decomposed components, we identified dynamic drivers as sub-networks most responsible for temporal changes in correlation structure. The successful application of our method to a different types of experimental settings and data, demonstrate the general utility of the approach to neuroscience. Beyond neuroscience, our framework opens new avenues for analyzing evolving structured data in domains such as finance, climate science, and spatiotemporal graph learning. Future directions include generalizing the filter design on the SPSD manifold and addressing the method's reliance on consistent components across time, a challenge in settings like neuroscience, where tracking identical cells across sessions is often infeasible. Integrating our approach with deep learning may enable analysis of dynamic correlation structures with partial or mismatched observations.

References

[1] Xingwei An, Yutao Zhou, Yang Di, Ying Han, and Dong Ming. A novel method to identify mild cognitive impairment using dynamic spatio-temporal graph neural network. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2024.

[2] Jessica R Andrews-Hanna, Abraham Z Snyder, Justin L Vincent, Cindy Lustig, Denise Head, Marcus E Raichle, and Randy L Buckner. Disruption of large-scale brain systems in advanced aging. *Neuron*, 56(5):924–935, 2007.

[3] Ralf Banisch and Péter Koltai. Understanding the geometry of transport: diffusion maps for lagrangian trajectory data unravel coherent sets. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(3), 2017.

[4] Hadas Benisty, Daniel Barson, Andrew H Moberly, Sweyta Lohani, Lan Tang, Ronald R Coifman, Michael C Crair, Gal Mishne, Jessica A Cardin, and Michael J Higley. Rapid fluctuations in functional connectivity of cortical networks encode spontaneous behavior. *Nature Neuroscience*, 27(1):148–158, 2024.

[5] Silvère Bonnabel and Rodolphe Sepulchre. Riemannian metric and geometric mean for positive semidefinite matrices of fixed rank. *SIAM Journal on Matrix Analysis and Applications*, 31(3):1055–1070, 2010.

[6] Weiwei Chang, Jinping Liu, Liluo Nie, Xiaomin Pang, Zongxia Lv, and Jinou Zheng. The degree centrality and functional connectivity in patients with temporal lobe epilepsy presenting as ictal panic: a resting state fmri study. *Frontiers in Neurology*, 13:822253, 2022.

[7] Yizhou Chen, Anxiang Zeng, Qingtao Yu, Kerui Zhang, Cao Yuanpeng, Kangle Wu, Guangda Huzhang, Han Yu, and Zhiming Zhou. Recurrent temporal revision graph networks. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, volume 36, pages 69348–69360. Curran Associates, Inc., 2023.

[8] Joseph Cichon and Wen-Biao Gan. Branch-specific dendritic ca2+ spikes cause persistent synaptic plasticity. *Nature*, 520(7546):180–185, 2015.

[9] David Cohen, Tal Shnitzer, Yuval Kluger, and Ronen Talmon. Few-sample feature selection via feature manifold learning. In *International Conference on Machine Learning*, pages 6296–6319. PMLR, 2023.

[10] Ronald R. Coifman and Stéphane Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006. Special Issue: Diffusion Maps and Wavelets.

[11] Joao Couto, Simon Musall, Xiaonan R Sun, Anup Khanal, Steven Gluf, Shreya Saxena, Ian Kinsella, Taiga Abe, John P Cunningham, Liam Paninski, et al. Chronic, cortex-wide imaging of specific cell populations during behavior. *Nature protocols*, 16(7):3241–3263, 2021.

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[12] Hanjun Dai, Yichen Wang, Rakshit Trivedi, and Le Song. Deep coevolutionary network: Embedding user and item features for recommendation. *arXiv preprint arXiv:1609.03675*, 2016.

[13] Ingrid Daubechies. Ten lectures on wavelets. *Society for industrial and applied mathematics*, 1992.

[14] Emily L Dennis and Paul M Thompson. Functional brain connectivity using fmri in aging and alzheimer's disease. *Neuropsychology Review*, 24(1):49–62, 2014.

[15] Mehrdad Fatourechi, Ali Bashashati, Rabab K Ward, and Gary E Birch. Emg and eog artifacts in brain computer interface systems: A survey. *Clinical Neurophysiology*, 118(3):480–494, 2007.

[16] Gary Froyland. Dynamic isoperimetry and the geometry of lagrangian coherent structures. *Nonlinearity*, 28(10):3587, 2015.

[17] Gary Froyland and Eric Kwok. A dynamic laplacian for identifying lagrangian coherent structures on weighted riemannian manifolds. *Journal of Nonlinear Science*, 30:1889–1971, 2020.

[18] Min Fu, Xinzhu Yu, Ju Lu, and Yi Zuo. Repetitive motor learning induces coordinated formation of clustered dendritic spines in vivo. *Nature*, 483(7387):92–95, 2012.

[19] Amir Ghanayim, Hadas Benisty, Avigail Cohen Rimon, Sivan Schwartz, Sally Dabdoob, Shira Lifshitz, Ronen Talmon, and Jackie Schiller. Vta projections to m1 are essential for reorganization of layer 2-3 network dynamics underlying motor learning. *Nature Communications*, 16(1):200, 2025.

[20] Zhiwu Huang and Luc Van Gool. A riemannian network for spd matrix learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 31, 2017.

[21] Kaihua Jiang, Yang Yi, Lin Li, Hongxin Li, Huijuan Shen, Fangqiao Zhao, Yunpin Xu, and Aibin Zheng. Functional network connectivity changes in children with attention-deficit hyperactivity disorder: A resting-state fmri study. *International Journal of Developmental Neuroscience*, 78:1–6, 2019.

[22] Rongtao Jiang, Nianming Zuo, Judith M Ford, Shile Qi, Dongmei Zhi, Chuanjun Zhuo, Yong Xu, Zening Fu, Juan Bustillo, Jessica A Turner, et al. Task-induced brain connectivity promotes the detection of individual differences in brain-behavior relationships. *NeuroImage*, 207:116370, 2020.

[23] Ming Jin, Yuan-Fang Li, and Shirui Pan. Neural temporal walks: Motif-aware representation learning on continuous-time dynamic graphs. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 19874–19886. Curran Associates, Inc., 2022.

[24] Fotis P Kalaganis, Nikos A Laskaris, Elisavet Chatzilari, Spiros Nikolopoulos, and Ioannis Kompatsiaris. A riemannian geometry approach to reduced and discriminative covariance estimation in brain computer interfaces. *IEEE Transactions on Biomedical Engineering*, 67:245–255, 2019.

[25] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

[26] Manoj Kumar, Gregory Handy, Stylianos Kouvaros, Yanjun Zhao, Lovisa Ljungqvist Brinson, Eric Wei, Brandon Bizup, Brent Doiron, and Thanos Tzounopoulos. Cell-type-specific plasticity of inhibitory interneurons in the rehabilitation of auditory cortex after peripheral damage. *Nature communications*, 14(1):4170, 2023.

[27] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv preprint arXiv:1707.01926*, 2017.

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[28] Simon Musall, Xiaonan R Sun, Hemanth Mohan, Xu An, Steven Gluf, Shu-Jing Li, Rhonda Drewes, Emma Cravo, Irene Lenzi, Chaoqun Yin, et al. Pyramidal cell types drive functionally distinct cortical activity patterns during decision-making. *Nature neuroscience*, 26(3):495–505, 2023.

[29] Andrew J Peters, Jun Lee, Nathan G Hedrick, Keelin O'Neil, and Takaki Komiyama. Reorganization of corticospinal output during motor learning. *Nature neuroscience*, 20(8):1133–1141, 2017.

[30] Emanuele Rossi, Ben Chamberlain, Fabrizio Frasca, Davide Eynard, Federico Monti, and Michael Bronstein. Temporal graph networks for deep learning on dynamic graphs. *arXiv preprint arXiv:2006.10637*, 2020.

[31] Aravind Sankar, Yanhong Wu, Liang Gou, Wei Zhang, and Hao Yang. Dysat: Deep neural representation learning on dynamic graphs via self-attention networks. In *Proceedings of the 13th international conference on web search and data mining*, pages 519–527, 2020.

[32] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE transactions on neural networks*, 20(1):61–80, 2008.

[33] Yan-Liang Shi, Nicholas A Steinmetz, Tirin Moore, Kwabena Boahen, and Tatiana A Engel. Cortical state dynamics and selective attention define the spatial pattern of correlated variability in neocortex. *Nature communications*, 13(1):44, 2022.

[34] Yan-Liang Shi, Roxana Zeraati, Anna Levina, and Tatiana A Engel. Spatial and temporal correlations in neural networks with structured connectivity. *Physical review research*, 5(1):013005, 2023.

[35] Tal Shnitzer, Hau-Tieng Wu, and Ronen Talmon. Spatiotemporal analysis using riemannian composition of diffusion operators. *Applied and Computational Harmonic Analysis*, 68:101583, 2024.

[36] Le Song, Mladen Kolar, and Eric Xing. Time-varying dynamic bayesian networks. In Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta, editors, *Advances in Neural Information Processing Systems*, volume 22. Curran Associates, Inc., 2009.

[37] Lloyd N Trefethen and Mark Embree. Spectra and pseudospectra: the behavior of nonnormal matrices and operators. *Princeton university press*, 2020.

[38] Rong Wang, Mianxin Liu, Xinhong Cheng, Ying Wu, Andrea Hildebrandt, and Changsong Zhou. Segregation, integration, and balance of large-scale resting brain networks configure different cognitive abilities. *Proceedings of the National Academy of Sciences*, 118(23):e2022288118, 2021.

[39] S. Wein, A. Schüller, A. M. Tomé, W. M. Malloni, M. W. Greenlee, and E. W. Lang. Forecasting brain activity based on models of spatiotemporal brain dynamics: A comparison of graph neural network architectures. *Network Neuroscience*, 6(3):665–701, 07 2022.

[40] Or Yair, Mirela Ben-Chen, and Ronen Talmon. Parallel transport on the cone manifold of spd matrices for domain adaptation. *IEEE Transactions on Signal Processing*, 67(7):1797–1811, 2019.

[41] Rex Ying, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L Hamilton, and Jure Leskovec. Graph convolutional neural networks for web-scale recommender systems. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 974–983, 2018.

[42] Jiani Zhang, Xingjian Shi, Junyuan Xie, Hao Ma, Irwin King, and Dit-Yan Yeung. Gaan: Gated attention networks for learning on large and spatiotemporal graphs. *arXiv preprint arXiv:1803.07294*, 2018.

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

7 Appendix

7.1 Proofs for spectral analysis

Theorem 2. Let ψ be an eigenvector of both $\mathbf{W}_1$ and $\mathbf{W}_2$ with eigenvalues λ_1 and λ_2, respectively: $\mathbf{W}_1\psi = \lambda_1\psi, \quad \mathbf{W}_2\psi = \lambda_2\psi$. Then ψ is also an eigenvector of the difference operator $\mathbf{W}_1\%\mathbf{W}_2$, with eigenvalue: $\lambda_1^{-1}\lambda_2^2$.

Proof.

$$\begin{aligned}(\mathbf{W}_1\%\mathbf{W}_2)\,\psi &= \mathbf{W}_1^{0.5}\left(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5}\right)^2\mathbf{W}_1^{0.5}\psi\\ &= \mathbf{W}_1^{0.5}\left(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5}\right)^2\sqrt{\lambda_1}\psi\\ &= \mathbf{W}_1^{0.5}\left(\lambda_1^{-1}\lambda_2\right)^2\sqrt{\lambda_1}\psi\\ &= \lambda_1^{-1}\lambda_2^2\psi.\end{aligned}$$

□

Theorem 4. Suppose there exists an eigenpair λ_k and ψ_k of $\mathbf{W}_k$ for $k = 1, 2$ so that $\psi_1 = \psi_2 + \psi_\varepsilon$, where $\|\psi_\varepsilon\|_2 \le \frac{1}{(\sigma_{\max}(\mathbf{W}_2\mathbf{W}_1{}^{-1})+\lambda_2\lambda_1^{-1})\tilde{\lambda}_2^{\max}}\varepsilon$, where $\sigma_{\max}$ denotes the maximum eigenvalue, for a small $\varepsilon > 0$, and where $\tilde{\lambda}_2^{\max} = \|\mathbf{W}_2 - \lambda_2\mathbf{I}\|$. Denote $\mathbf{Q} = \mathbf{W}_1\%\mathbf{W}_2$. Then we have $\lambda_2^2\lambda_1^{-1} \in \Lambda_\varepsilon(\mathbf{Q})$. Specifically, we have $\|(\mathbf{Q} - \lambda_2^2\lambda_1^{-1}\mathbf{I})\psi_1\|_2 \le \varepsilon$, and ψ_1 is a corresponding ε-pseudo-eigenvector of $\mathbf{Q}$.

Proof. By using (10 from appendix 7.3) , we have:

$$\mathbf{Q} = \mathbf{W}_1{}^{0.5}\left(\mathbf{W}_1{}^{-0.5}\mathbf{W}_2\mathbf{W}_1{}^{-0.5}\right)^2\mathbf{W}_1{}^{0.5} = (\mathbf{W}_2\mathbf{W}_1{}^{-1})^2\mathbf{W}_1.$$

Since ψ_1 is an eigenvector of $\mathbf{W}_1$ with eigenvalue λ_1, we obtain:

$$\mathbf{Q}\psi_1 = (\mathbf{W}_2\mathbf{W}_1{}^{-1})^2\mathbf{W}_1\psi_1 = \lambda_1(\mathbf{W}_2\mathbf{W}_1{}^{-1})^2\psi_1.$$

Thus, it suffices to show that ψ_1 is an ε-pseudo-eigenvector of $(\mathbf{W}_2\mathbf{W}_1{}^{-1})^2$. By direct expansion, we have:

$$\mathbf{W}_2\mathbf{W}_1{}^{-1}\psi_1 = \mathbf{W}_2\mathbf{W}_1{}^{-1}\psi_1 = \frac{1}{\lambda_1}\mathbf{W}_2\psi_1 = \frac{1}{\lambda_1}\mathbf{W}_2(\psi_2 + \psi_\varepsilon).$$

$$= \frac{\lambda_2}{\lambda_1}\psi_1 + \frac{1}{\lambda_1}(\mathbf{W}_2 - \lambda_2\mathbf{I})\psi_\varepsilon.$$

This implies:

$$\left((\mathbf{W}_2\mathbf{W}_1{}^{-1})^2 - \lambda_2^2\lambda_1^{-2}\mathbf{I}\right)\psi_1 = (\mathbf{W}_2\mathbf{W}_1{}^{-1} + \lambda_2\lambda_1^{-1}\mathbf{I})(\mathbf{W}_2\mathbf{W}_1{}^{-1} - \lambda_2\lambda_1^{-1}\mathbf{I})\psi_1.$$

Multiplying both sides by λ_1, we get:

$$(\mathbf{Q} - \lambda_2^2\lambda_1^{-1}\mathbf{I})\psi_1 = (\mathbf{W}_2\mathbf{W}_1{}^{-1} + \lambda_2\lambda_1^{-1}\mathbf{I})(\mathbf{W}_2 - \lambda_2\mathbf{I})\psi_\varepsilon.$$

Taking the ℓ_2 norm, we obtain:

$$\|(\mathbf{Q} - \lambda_2^2\lambda_1^{-1}\mathbf{I})\psi_1\|_2 \le \|\mathbf{W}_2\mathbf{W}_1{}^{-1} + \lambda_2\lambda_1^{-1}\mathbf{I}\|_2\|\mathbf{W}_2 - \lambda_2\mathbf{I}\|_2\|\psi_\varepsilon\|_2.$$

Applying the given bound on $\|\psi_\varepsilon\|_2$, we conclude:

$$\|(\mathbf{Q} - \lambda_2^2\lambda_1^{-1}\mathbf{I})\psi_1\|_2 \le \varepsilon.$$

Thus, $\lambda_2^2\lambda_1^{-1}$ is an ε pseudo-eigenvalue of $\mathbf{Q}$, where ψ_1 is a corresponding ε pseudo-eigenvector. □

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

7.2 Proof mutually inverse

Proposition 2. Let $\mathbf{W}_1$ and $\mathbf{W}_2$ be SPSD matrices, the operators defined in (2)(4) are mutually inverse, meaning:

$$\begin{aligned}\mathbf{W}_2\%(\mathbf{W}_1\#\mathbf{W}_2) &= \mathbf{W}_1\\ \mathbf{W}_2\#(\mathbf{W}_2\%\mathbf{W}_1) &= \mathbf{W}_1\end{aligned} \tag{9}$$

Proof. We will first prove that $\mathbf{W}_2\%(\mathbf{W}_1\#\mathbf{W}_2) = \mathbf{W}_1$:

$$\mathbf{W}_1\#\mathbf{W}_2 = \mathbf{W}_1^{0.5}(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5})^{0.5}\mathbf{W}_1^{0.5} = \mathbf{W}_3$$

$$\begin{aligned}\mathbf{W}_2\%\mathbf{W}_3 &= \mathbf{W}_2^{0.5}(\mathbf{W}_2^{-0.5}\mathbf{W}_3\mathbf{W}_2^{-0.5})^2\mathbf{W}_2^{0.5}\\ &= \mathbf{W}_2^{0.5}(\mathbf{W}_2^{-0.5}\mathbf{W}_1^{0.5}(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5})^{0.5}\mathbf{W}_1^{0.5}\mathbf{W}_2^{-0.5})^2\mathbf{W}_2^{0.5}\\ &= \mathbf{W}_1^{0.5}(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5})^{0.5}\mathbf{W}_1^{0.5}\mathbf{W}_2^{-1}\mathbf{W}_1^{0.5}(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5})^{0.5}\mathbf{W}_1^{0.5}\end{aligned}$$

Assume this expression equals $\mathbf{W}_1$:

$$\begin{gathered}\mathbf{W}_1 = \mathbf{W}_1^{0.5}(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5})^{0.5}\mathbf{W}_1^{0.5}\mathbf{W}_2^{-1}\mathbf{W}_1^{0.5}(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5})^{0.5}\mathbf{W}_1^{0.5}\\ \mathbf{W}_1^{-0.5}(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5})^{-0.5}\mathbf{W}_1^{-0.5}\mathbf{W}_1\mathbf{W}_1^{-0.5}(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5})^{-0.5}\mathbf{W}_1^{-0.5} = \mathbf{W}_2^{-1}\\ \mathbf{W}_1^{-0.5}(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5})^{-0.5}(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5})^{-0.5}\mathbf{W}_1^{-0.5} = \mathbf{W}_2^{-1}\\ \mathbf{W}_1^{-0.5}(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5})^{-1}\mathbf{W}_1^{-0.5} = \mathbf{W}_2^{-1}\\ \mathbf{W}_1^{-0.5}\mathbf{W}_1^{0.5}\mathbf{W}_2^{-1}\mathbf{W}_1^{0.5}\mathbf{W}_1^{-0.5} = \mathbf{W}_2^{-1}\\ \mathbf{W}_2^{-1} = \mathbf{W}_2^{-1}\\ \Rightarrow \mathbf{W}_1 = \mathbf{W}_2\%(\mathbf{W}_1\#\mathbf{W}_2)\end{gathered}$$

Now, we prove that $\mathbf{W}_2\#(\mathbf{W}_2\%\mathbf{W}_1) = \mathbf{W}_1$:

$$\mathbf{W}_2\%\mathbf{W}_1 = \mathbf{W}_2^{0.5}(\mathbf{W}_2^{-0.5}\mathbf{W}_1\mathbf{W}_2^{-0.5})^2\mathbf{W}_2^{0.5} = \mathbf{W}_4$$

$$\begin{aligned}\mathbf{W}_2\#\mathbf{W}_4 &= \mathbf{W}_2^{0.5}(\mathbf{W}_2^{-0.5}\mathbf{W}_4\mathbf{W}_2^{-0.5})^{0.5}\mathbf{W}_2^{0.5}\\ &= \mathbf{W}_2^{0.5}(\mathbf{W}_2^{-0.5}\mathbf{W}_2^{0.5}(\mathbf{W}_2^{-0.5}\mathbf{W}_1\mathbf{W}_2^{-0.5})^2\mathbf{W}_2^{0.5}\mathbf{W}_2^{-0.5})^{0.5}\mathbf{W}_2^{0.5}\\ &= \mathbf{W}_2^{0.5}((\mathbf{W}_2^{-0.5}\mathbf{W}_1\mathbf{W}_2^{-0.5})^2)^{0.5}\mathbf{W}_2^{0.5}\\ &= \mathbf{W}_2^{0.5}\mathbf{W}_2^{-0.5}\mathbf{W}_1\mathbf{W}_2^{-0.5}\mathbf{W}_2^{0.5}\\ &= \mathbf{W}_1\end{aligned}$$

□

7.3 Equivalent form

Proposition 3. Equivalent form of the filters. We have

$$\mathbf{W}_1^{0.5}\left(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5}\right)^p\mathbf{W}_1^{0.5} = \left(\mathbf{W}_2\mathbf{W}_1^{-1}\right)^p\mathbf{W}_1$$

Proof. For some $p \geq 0$, consider the expression:

$$\mathbf{W}_1^{0.5}\left(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5}\right)^p\mathbf{W}_1^{0.5}.$$

Define the matrices:

$$\mathbf{M} = \mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5}, \quad \tilde{\mathbf{M}} = \mathbf{W}_2\mathbf{W}_1^{-1}.$$

Since $\mathbf{W}_1$ and $\mathbf{W}_2$ are positive definite, so is $\mathbf{M}$, which implies that $\mathbf{M}$ and $\tilde{\mathbf{M}}$ are similar via:

$$\mathbf{M} = \mathbf{W}_1^{-0.5}\tilde{\mathbf{M}}\mathbf{W}_1^{-0.5}.$$

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

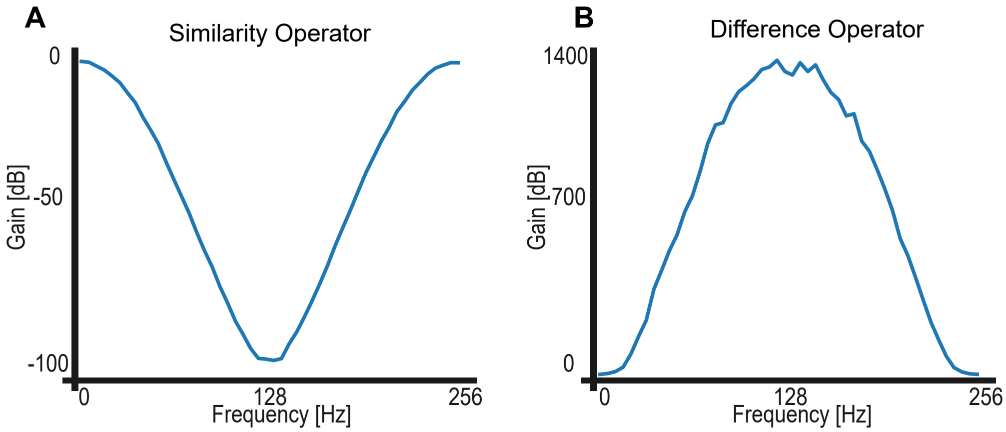

Figure 4: (A) Gain of the similarity operator as a function of input frequency. (B) Gain of the difference operator as a function of input frequency.

Let Λ^M and $\mathbf{V}^M$ be the eigenvalue and eigenvector matrices of $\mathbf{M}$, respectively. Then the eigenvalue matrix of $\tilde{\mathbf{M}}$ satisfies:

$$\Lambda^{\tilde{M}} = \Lambda^M,$$

with right and left eigenvectors given by:

$$\mathbf{V}_R^{\tilde{M}} = \mathbf{W}_1^{0.5}\mathbf{V}^M, \quad \mathbf{V}_L^{\tilde{M}} = \mathbf{W}_1^{-0.5}\mathbf{V}^M.$$

Thus, we obtain the matrix powers:

$$\mathbf{M}^p = \mathbf{V}^M \left(\Lambda^M\right)^p \left(\mathbf{V}^M\right)^T,$$

$$\tilde{\mathbf{M}}^p = \mathbf{W}_1^{0.5}\mathbf{V}^M \left(\Lambda^M\right)^p \left(\mathbf{V}^M\right)^T \mathbf{W}_1^{-0.5}.$$

From this, we conclude:

$$\tilde{\mathbf{M}}^p = \mathbf{W}_1^{0.5}\mathbf{M}^p\mathbf{W}_1^{-0.5}.$$

Therefore, we derive:

$$\mathbf{W}_1^{0.5}\mathbf{M}^p\mathbf{W}_1^{0.5} = \mathbf{W}_1^{0.5}\mathbf{M}^p\mathbf{W}_1^{-0.5}\mathbf{W}_1 = \tilde{\mathbf{M}}^p\mathbf{W}_1.$$

Substituting back for $\mathbf{M}$ and $\tilde{\mathbf{M}}$, we obtain:

$$\mathbf{W}_1^{0.5}\left(\mathbf{W}_1^{-0.5}\mathbf{W}_2\mathbf{W}_1^{-0.5}\right)^p\mathbf{W}_1^{0.5} = \left(\mathbf{W}_2\mathbf{W}_1^{-1}\right)^p\mathbf{W}_1. \tag{10}$$

□

7.4 Similarity and difference operators as low- and high-pass filters

In classical signal processing, low- and high-pass filters are characterized by their frequency response, typically determined by a 3dB drop in the gain of the output signal. In this section, we extend this concept to the SPSD setting to further demonstrate the properties of the similarity and difference operators (eqns. (3) and (5). We simulate a set of 256 SPSD matrices of size 10×10, where the first three components are sinusoidal signals, and the remaining entries are a small random noise:

$$\begin{aligned} \mathbf{W}_n(k,l) &= 0.5\sin\left(2\pi f_1 n\right) + U(-0.01, 0.01), \quad k,l \in \{1,2,3\} \\ \mathbf{W}_n(k,l) &\sim U(-0.01, 0.01), \quad \text{otherwise} \\ \mathbf{W}_n(k,k) &= 1, \quad \text{for all } k \end{aligned}$$

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

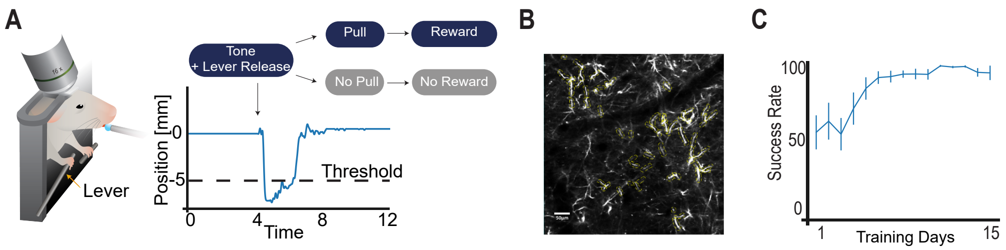

Figure 5: (A) Schematic of the experimental setup: a head-fixed mouse performs a lever pull task during 2-photon Calcium imaging. Each trial lasts 12 seconds, where a go-cue is given after 4 seconds. Lever position crossing the 5mm threshold is rewarded with a sweetened drop of water delivered by a water spout. (B) Example field of view from the calcium imaging data, where yellow lines indicate ROI labeling of dendritic activity. (C) Success rate (mean±SEM) of 8 mice performing the lever-pull task across training days, showing an increase in performance with training.

where $U(a, b)$ denotes a uniform distribution in the range $[a, b]$, and f_1 is the sinusoidal frequency.

Analogous to classical 1D filters, we measure the spectral response of the filters by comparing the gain of the sequence before filtering, $\{\mathbf{W}_n\}_{n=1}^{N}$, to the gain of the sequence after filtering, $\{\mathbf{U}_n\}_{n=1}^{N}$, as a function of the frequency f_1. We note that all matrices in the input signal have one dominant component. If filtering rejected the sinusoidal components, the filtered output would consist of noise alone, meaning there would be no dominant component. Following this logic, we propose taking the ratio between the dominance of the largest eigenvalue before and after filtering as a measure of the gain of the filter. Formally, the dominance of the first eigen-component of a matrix, $\mathbf{A}$, is:

$$D_{\mathbf{A}} = 1 - \frac{\lambda_1}{\mathrm{Tr}(\mathbf{A})} \tag{11}$$

where, λ_1 is the largest eigenvalue of $\mathbf{A}$ and $\mathrm{Tr}(\mathbf{A})$ is the trace of $\mathbf{A}$. Given the two sequences of matrices we thus take the average ratio of dominance values before and after filtering:

$$G = \sum_{n=1}^{N} 20 \log_{10} \frac{D_{\mathbf{U}_n}}{D_{\mathbf{W}_n}} \tag{12}$$

We computed the gain of both similarity and difference operators as a function of frequency by varying the f_1, as depicted in Figure 4. The gain of the similarity operator aligns with the behavior of the Haar low-pass filter, preserving low frequencies and rejecting high frequencies. In contrast, the proposed difference operator matches the behavior of the Haar high-pass filter, emphasizing high frequencies.

7.5 M1 dendritic structures data set

In this longitudinal experiment, eight mice were trained to perform a motor task during 15 sessions. Each trial lasted 12 seconds, during which an auditory go cue was given four seconds into the trial, prompting the animals to pull the lever. Pulls that surpassed a preset displacement threshold were rewarded by a drop of sweetened water delivered to the animal's mouth through a water spout. Lever position was sampled at 150 Hz, yielding millisecond-resolution kinematic traces. Throughout every session, simultaneous two-photon Calcium imaging (30 Hz) was conducted to monitor the apical tuft dendrites of 5–10 layer-5 Pyramidal Tract (L5 PT) neurons in M1 forelimb region. For stable tracking of dendrites, a sparse labeling technique of L5 PT neurons was applied. In addition, the pons was injected at three depths with a viral mixture comprising (i) a diluted retrograde AAV carrying Cre recombinase, (ii) a concentrated retroAAV-CAG-FLEX-jGCaMP7s, and (iii) a concentrated retroAAV-FLEX-tdTomato. A chronic cranial window implanted over forelimb M1 enabled repeated imaging of the same field of view (FOV) across days. After each session, regions of interest (ROIs) were manually annotated, preserving the identity of individual dendritic segments across all recordings.

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

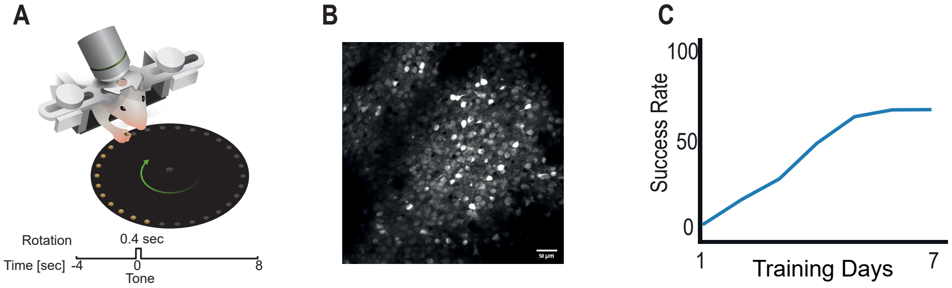

Figure 6: (A) Experimental setup: a head-fixed mouse performing a motor learning task during 2-photon calcium imaging. Following a go-cue, the rotating plate prompts the mouse to reach and grasp a food pellet. Each trial lasts 12 seconds, including a 4-second pre-cue phase. (B) Example field of view from the Calcium imaging data. (C) Success rate (percentage of successful trials) of the analyzed mouse across training sessions. Motor proficiency improves over days and plateaus by session 5.

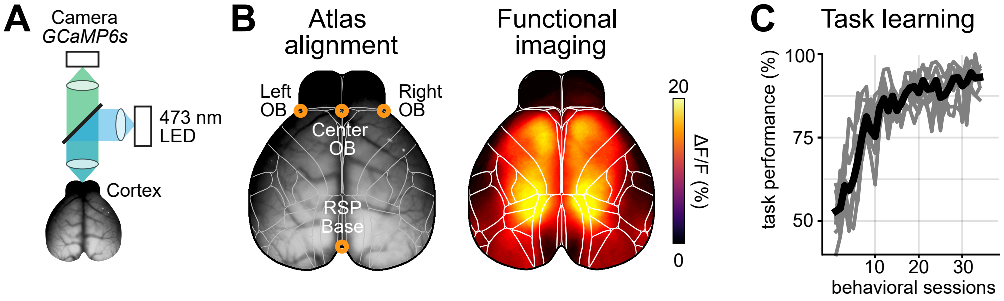

Figure 7: (A) Schematic of the widefield imaging setup for cortex-wide functional imaging. Imaging was performed through the intact, cleared skull. (B) Imaging data was aligned to the Allen CCF using 4 anatomical landmarks. Cortical activity, measured as relative changes in measured fluorescence, could then be related to the activity of cortical areas (white lines). (C) Learning curves of mice, trained in the auditory discrimination task. Animals achieved expert task performance within 10-20 sessions of behavioral training.

7.6 Large population in M1 data set

A brief description of the experimental setup is provided below. For comprehensive details, see [19]. In this experiment, head-fixed mice were trained to perform a hand-reach and grasp motor task, Fig. 6. Training took place over seven sessions, during which the neuronal activity of pyramidal neurons (PNs) in layer 2-3 of M1 was recorded using two-photon Calcium imaging. Chronic Calcium imaging was performed with the genetically encoded Calcium indicator GCaMP6s, which was injected via a viral vector injection into the M1 forelimb area. Layer 2-3 Pyramidal Neurons were selectively labeled using a specific promoter (CaMKII). A chronic window was then implanted over the M1 forelimb area, allowing for imaging of activity over long periods. At each imaging session, Regions Of Interest (ROIs) with approximately the same Field Of View (FOV) were imaged to track the activity of the same cell population across days. Recordings were conducted at a sampling rate of 30 Hz over 12-second trials, with an average of 40-50 trials per session. Registration of FOVs across sessions and ROIs extraction were performed manually. The dataset used in our analysis included seven sessions from one publicly available control animal, comprising the activity of 178 neurons over seven training sessions.

bioRxiv preprint doi: https://doi.org/10.1101/2025.03.26.645425; this version posted May 15, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

7.7 Widefield dataset

Comprehensive details of the widefield imaging are provided in [28] and [11]. Widefield imaging was done with an inverted tandem-lens macroscope and a sCMOS camera (Edge 5.5, PCO) running at 30 frames per second and a field of view of 12.5 x 10.5 mm2 (Fig. 7A). Imaging resolution was 640×540 pixels after $4\times$ spatial binning, resulting in a spatial resolution of about 20 μm per pixel. Mice were Ai93D;Emx-Cre;LSL-tTA animals, expressing the Calcium indicator GCaMP6f in all excitatory cortical neurons. All imaging data were then rigidly aligned to the Allen Mouse Brain Common Coordinate Framework (CCF), using anatomical landmarks, to allow for inferring the activity of different cortical areas over the course of task learning (Fig. 7B). Accurate atlas alignment was further confirmed using retinotopic mapping, showing high agreement of functionally inferred locations of visual areas.

During widefield imaging, head-fixed mice were trained in an auditory discrimination task, where two sequences of Poisson-distributed click sounds were presented to either the left or right side of the animal for 1 to 1.5 seconds. After the stimulus and a subsequent 0.5-long delay period, mice had to perform a licking response on the side where more clicks were presented to obtain a water reward. The auditory discrimination task, therefore, required the accumulation of the auditory information over time and working memory to obtain high task performance (Fig. 7C). Animals were trained between 30 and 40 sessions, and we analyzed changes in cortical correlation structure over the course of task learning. Results here include cortex-wide population activity recorded from all training sessions of one animal out of four included in the full dataset.