Proceedings Track

Analytical Formulation of LFP Manifold

Editors: List of editors' names

Abstract

Manifold geometry has become a powerful lens for understanding how the brain organizes high-dimensional activity into compact, interpretable structure. Most of these insights, however, come from spiking activity, leaving open the question of whether the same geometrical principles extend to local field potentials (LFPs)—the collective signals that reflect the coordination of neural populations. Here we take advantage of the spectral structure of LFPs to derive a closed-form analytical expression for their intrinsic dimensionality. We prove the core geometry of this manifold is a K-torus, where the dimension K is determined by the number of distinct oscillatory rhythms. We validate the computational feasibility and theoretical correctness of our framework using a synthetic benchmark with known chaotic dynamics. This work provides a new, analytically grounded language for interpreting population signals, laying the theoretical foundation for transforming noisy LFPs into reusable geometric abstractions.

Keywords: Neural manifolds, Local Field Potential, Autoregressive Models, Dimensionality, Torus, Systems Neuroscience

1. Introduction

The capacity to record from large neural populations has fundamentally reshaped neuroscience, confronting the field with high-dimensional datasets that defy traditional single-neuron analysis. A powerful paradigm for understanding this complexity is the manifold hypothesis, which posits that despite the immense number of neurons, their collective activity is constrained to a lower-dimensional geometric structure, or neural manifold Cunningham and Yu (2014). This perspective reframes neural computation: instead of arising from the independent firing of single cells, it is seen as the result of coordinated dynamics unfolding along these intrinsic manifolds Vyas et al. (2020). Seminal work in this area has provided both the theoretical motivation and practical tools for using dimensionality reduction to extract these latent structures from population activity Gao and Ganguli (2017). Methods such as Gaussian Process Factor Analysis have been particularly influential, enabling the analysis of neural trajectories on a moment-by-moment basis without losing the dynamics in trial-averaging Yu et al. (2009). By moving the focus from individual neurons to the shared, low-dimensional space they collectively inhabit, the manifold framework offers a principled way to uncover the geometric foundations of neural information processing.

While much of this progress has centered on spiking activity Levy et al. (2023); Chung et al. (2016); Cohen et al. (2020); Chung and Abbott (2021), another crucial signal, the local field potential (LFP), offers a complementary window into collective neural dynamics. The LFP is understood to primarily reflect the summed synaptic transmembrane currents from a local population of neurons, offering a measure of aggregate input and subthreshold

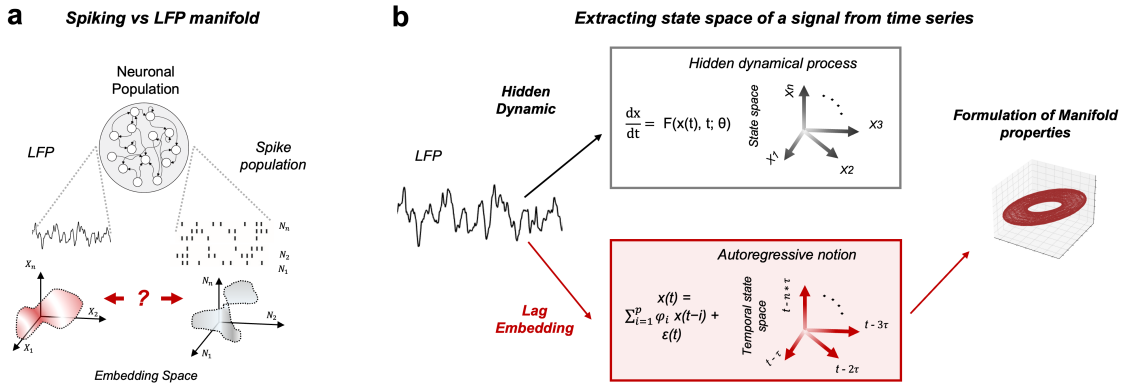

Figure 1: **Conceptual framework for deriving LFP manifold geometry. (a)** While neural manifold studies have successfully characterized the geometry of spiking activity, the corresponding geometry of the aggregate LFP signal remains largely unknown. **(b)** Our approach bridges this gap. Instead of assuming a hidden, unobserved dynamical process, we model the LFP time series directly using an autoregressive (AR) model. This allows us to use lag embedding to reconstruct the system's state space and analytically derive the properties of its underlying manifold, which we prove to be a K-torus.

processing rather than spiking output Buzsáki et al. (2012); Einevoll et al. (2013); Choubdar et al. (2023); Fazlali et al. (2020). The complex biophysical origins of the LFP have been extensively studied, establishing it as a rich, albeit intricate, signal reflecting the cooperative behavior of neural ensembles. Furthermore, numerous studies have powerfully linked LFP oscillations to high-level cognitive functions, such as cognitive control, working memory, and attention, demonstrating that these field potentials are not mere epiphenomena but carry vital computational information Miller et al. (2018); Siegel et al. (2012); Zabeh et al. (2023). The LFP therefore provides a unique vantage point on the shared inputs and processing within a neural circuit, which should, in principle, also possess a geometric structure.

Given the success of manifold analysis for spiking data and the clear importance of LFPs in neural computation, it is surprising that the intrinsic dimensionality and geometric properties of the LFP signal itself have not been clearly investigated. This represents a significant gap in our understanding. Because LFPs integrate subthreshold activity, their underlying geometric structure could reveal the organization of shared inputs to a circuit, providing a crucial counterpart to the output-related manifolds derived from spikes. The challenge of interpreting such complex, high-dimensional neural signals is well-recognized, often requiring sophisticated models to de-mix the contributions of various underlying sources Brette (2019). Understanding the manifold properties of LFPs is therefore not just a technical exercise, but a necessary step toward a complete picture of the geometric constraints governing neural information processing.

Here, we bridge this gap by developing an analytical formulation for the LFP manifold. Our approach is made possible by leveraging two well-established properties of the LFP

signal: its inherent boundedness and its extensively documented power spectrum, which typically consists of distinct oscillatory peaks superimposed on a 1/f-like aperiodic background Buzsáki and Draguhn (2004). While the power spectral density (PSD) is a valuable tool, it is intrinsically a compressed, lossy representation; it quantifies the power at different frequencies but discards the phase relationships required to reconstruct the signal or understand its underlying geometry. To overcome this limitation, we turn to a powerful modeling tradition rooted in physics and stochastic calculus. The complex, noisy dynamics observed in nature are often modeled with continuous-time stochastic differential equations (SDEs), such as the Ornstein-Uhlenbeck process, which elegantly describe the evolution of a system under the influence of both deterministic drift and random diffusion Gardiner (1985). Our key insight is to recognize that the discrete-time autoregressive (AR) process is the direct counterpart to these SDEs, providing a principled way to model the LFP's evolution from one moment to the next Stoica and Moses (2005); Kaminski and Blinowska (1991). By modeling the LFP as a time-varying AR process, we transform the complex biological signal into a tractable dynamical system, allowing us to derive a closed-form description of its generative dynamics without needing to know all the hidden biophysical parameters.

With this analytical framework, we provide a formal basis for the geometry of the recorded LFP. We explore the geometrical properties of the resulting state space and prove that the generative manifold of an LFP signal dominated by oscillatory components is a k-torus. We show that its intrinsic dimension, k, is determined directly by the number of dominant, pure oscillatory components in the signal, while the aperiodic 1/f component contributes to the thickness of the data around the manifold but not to its essential topology. We validate our analytical results first with a synthetic benchmark with known dynamics, confirming the computational feasibility and theoretical correctness of our pipeline. Critically, we then demonstrate the practical utility of our method by applying it to a real LFP recording, successfully retrieving the underlying manifold structure from a noisy biological signal. We expect this analytical investigation and methodology to provide a useful and principled view for the neuroscience community, enabling a deeper interpretation of future experiments and a geometric re-interpretation of existing datasets.

2. Modeling Framework

LFPs change smoothly, exhibit short-term memory, and often contain rhythms. Assuming *dynamics* formalizes this: the next value depends on the recent past, so trajectories concentrate on a low-dimensional manifold rather than filling $\mathbb{R}^N$. Random inputs and measurement noise add thickness around this set but do not determine its shape. A dynamic model thus separates lawful structure from randomness and makes the geometry easier to interpret. Autoregressive (AR) models are a natural choice: they use few parameters, are straightforward to fit, and capture both smooth $1/f$-like backgrounds and narrowband peaks, yielding an interpretable, low-complexity summary of the signal.

2.1. AR Setup

Consider an N-dimensional, discrete-time process

$$X(t) \;=\; [\,x_1(t),\, x_2(t),\, \ldots,\, x_N(t)\,]^\top \in \mathbb{R}^N, \qquad t \in \mathbb{Z}. \tag{1}$$

Each channel follows a AR model

$$x_i(t) \;=\; \sum_{m=1}^{p_i} \phi_{i,m}\, x_i(t-m) \;+\; \varepsilon_i(t), \qquad i = 1, \ldots, N, \tag{2}$$

with order p_i, coefficients $\{\phi_{i,m}\}$ are given, and noise process $\varepsilon_i(t)$. We collect the samples as $\varepsilon(t) = [\varepsilon_1(t), \ldots, \varepsilon_N(t)]^\top$ and assume they are zero-mean, temporally white, and possibly cross–correlated across channels, i.e., $\varepsilon(t) \sim \mathcal{N}(0, R_\varepsilon(t))$ with a (possibly) time-varying covariance $R_\varepsilon(t) \in \mathbb{R}^{N\times N}$.[1]

For compactness, we will occasionally refer to the block–diagonal, stacked form of (2); the precise statement and proof of the equivalence appear in **Appendix A, Lemma 1**. All additional algebra (including the state–space companion form and stability/cyclostationarity details) is deferred to the appendix, so that the exposition here can remain focused on the modeling intuition and geometric consequences.

2.2. State-Space View

Empirically, each LFP channel can be *predictively summarized* by a linear combination of its previous samples. To make that prediction problem explicit and easy to work with, we gather the current sample and the last $p-1$ samples from all N channels into one memory vector:

$$s(t) \;=\; \big[\, X(t)^\top,\; X(t-1)^\top,\; \ldots,\; X(t-p+1)^\top \,\big]^\top \in \mathbb{R}^{Np}. \tag{3}$$

With this stacked state, the evolution of the data becomes a linear update plus a small, unpredictable term:

$$s(t{+}1) \;=\; A\, s(t) \;+\; B\, \varepsilon(t{+}1), \qquad X(t) \;=\; C\, s(t). \tag{4}$$

Equation (4) turns the multi–channel LFP into a standard, low–parameter state-space model: A describes how yesterday's memory becomes today's; B injects the new, not-yet-explained part of the signal; and C simply selects the current observation from the memory stack. Working in this state-space form brings two immediate benefits. First, it cleanly separates what must be *learned* from data (how the past predicts the next step) from what is merely *bookkeeping* (the lag shift that carries history forward), which in practice leads to simple estimators and transparent gradients. Second, with coefficients the basic questions of stability and variance are routine: stability is governed by the eigenvalues of A (Schur stability), and the covariance follows the standard discrete Lyapunov relation $\Sigma = A\,\Sigma\,A^\top + B\,R_\varepsilon\,B^\top$, making second-moment analysis straightforward.[2]

2.3. Z-transform and pole analysis

Assume each channel $x_i(t)$ follows an AR model as defined in (2) and focus on the noise-free dynamics. Taking the Z-transform of the deterministic part yields

$$\left(1 - \sum_{m=1}^{p_i} \phi_{i,m}\, z^{-m}\right) X_i(z) \;=\; 0, \tag{5}$$

1. If desired, $R_\varepsilon(t)$ may be taken constant.
2. The companion-form matrices (A, B, C) and their algebraic equivalence to a AR(p) model are given in Lemma 2 (Appendix A). We retain only the high-level relations (3) and (4) in the main text to keep the narrative focused.

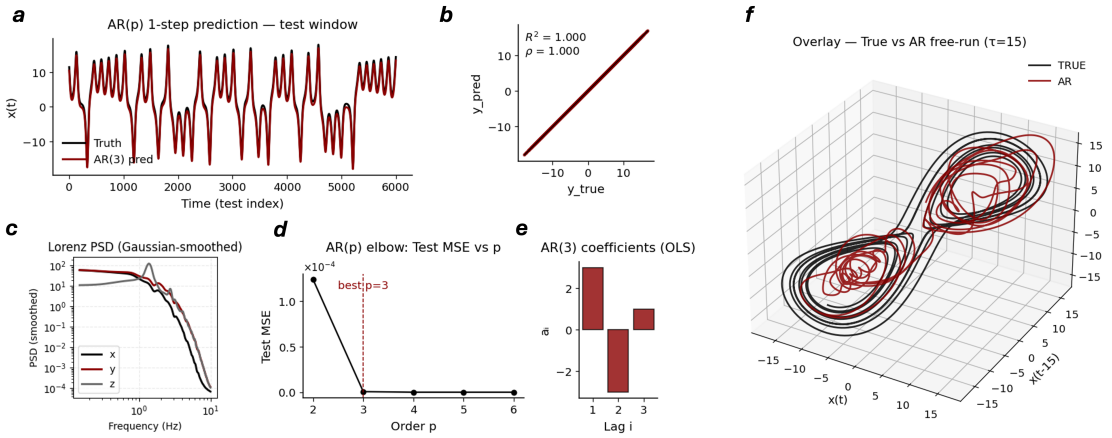

Figure 2: **Manifold reconstruction of synthetic Lorenz signal. (a)** One–step prediction on the test window for an AR(p) model (here p=3 selected by the elbow in **d**). The teacher–forced predictor closely tracks the held–out trajectory $x(t)$. **(b)** Pointwise agreement on the test set: scatter of y_{pred} vs. y_{true} with annotated R^2 and Pearson correlation ρ, confirming the strong short–horizon fit. **(c)** Gaussian–smoothed power spectra of the Lorenz coordinates (x, y, z) highlighting the broadband/chaotic content that an AR(p) must approximate. **(d)** Model–order selection curve ("elbow"): test MSE versus p, with the chosen $p = 3$. **(e)** Estimated AR(3) coefficients (a_1, a_2, a_3) from OLS. **(f)** Reconstructed manifold via delay embedding of a *free–run* simulation from the fitted AR model (no teacher forcing): overlay of the true attractor and the AR-generated trajectory in the 3D state space $[x(t), x(t-\tau), x(t-2\tau)]$ with τ=15.

so the long-term behavior is governed by the roots of the characteristic polynomial $P_i(z^{-1}) = 1 - \sum_{m=1}^{p_i} \phi_{i,m} z^{-m}$. Empirically, LFP-like signals have *bounded variance* (finite PSD), which rules out poles outside the unit circle. Poles strictly inside ($|z| < 1$) generate exponentially decaying components that die out; poles exactly on the unit circle in complex-conjugate pairs ($z = e^{\pm i\omega}$) generate sustained sinusoids $\sin(\omega t)$ and $\cos(\omega t)$. In other words, at long times each channel reduces to a sum of a finite number K_i of narrowband oscillations (plus bounded fluctuations), and the network-wide $K = \sum_i K_i$ *distinct* undamped frequencies are the only ones that survive asymptotically. A precise factorization statement is given in Lemma 3 (Appendix A).

2.4. Manifold view: from oscillations to a torus

The pole picture from the z–transform has a direct geometric interpretation. Each undamped oscillatory pair on the unit circle ($z = e^{\pm i\omega}$ with $0 < \omega < \pi$) contributes one independent *phase* that winds around a circle. If there is a single such mode, the noise–free state (e.g., a short lag stack of the signal) traces out a deformed circle. With two independent modes, the state moves on the Cartesian product of two circles—a 2–torus. In general, if there are K *distinct* undamped oscillatory pairs (no duplicates modulo π), the

latent geometry is the K–fold product of circles, i.e., a K–torus. Poles strictly inside the unit circle ($|z| < 1$) correspond to damped components; they add a thin "thickness" around the torus but do not change its intrinsic topology.

Informal proposition (K–torus geometry). In the AR setting, assume the signal has bounded variance and exactly K distinct undamped oscillatory pairs on the unit circle (no roots at $z = \pm 1$ and no frequency collisions modulo π). Then, ignoring noise, the trajectory of a lag–embedded state obtained by stacking *two* consecutive samples per oscillatory component evolves on a smooth K–dimensional torus. Each sustained oscillation contributes one circular phase; K phases yield the product of K circles. Damped poles ($|z| < 1$) do not create new circle factors; they only thicken the observed point cloud. If two oscillation frequencies coincide (mod π), one circle factor collapses and the intrinsic dimension drops accordingly. The formal version of this statement appears in Proposition 5 (Appendix A), and the "two lags per oscillation" requirement is made precise in Lemma 4 (Appendix A).

Practical reading of the proposition. If the spectrum exhibits, say, alpha and beta peaks (two sustained, distinct oscillations), then after stacking only two lags per peak, the underlying noise–free trajectory lives on a 2–torus; adding a third narrowband peak promotes it to a 3–torus. Background $1/f$–like activity and estimation noise merely blur the torus, without altering its circle count.

What can reduce the torus dimension? Frequency collisions (two peaks at the same frequency modulo π within the same channel) merge circle factors; severe damping moves poles inside the unit disk and removes the corresponding circle in the long run. Amplitude modulation or slow drifts change the local shape and thickness but, as long as the poles remain on the unit circle and distinct, the multi–circular topology persists.

2.5. Spectral density of the AR model

In LFP analysis, the power spectral density (PSD) is the common language for describing signal structure. It disentangles narrowband rhythms (peaks) from broadband background (the $1/f$-like floor). The PSD of the AR(p) signal is given by:

$$S_x(f) \;=\; \frac{\sigma_\varepsilon^2}{\left|\,1 - \sum_{k=1}^{p} a_k\, e^{-i\,2\pi f k/F_s}\,\right|^2}. \tag{6}$$

How to read (6). Write the roots of $P\big(z^{-1}\big)$ as $z = \rho e^{\pm i\omega}$. Bounded variance excludes $|z| > 1$ and real unit roots $z = \pm 1$. Unit–modulus conjugate pairs ($\rho = 1$, $0 < \omega < \pi$) generate persistent narrowband oscillations (sharp spectral peaks); subunit poles ($0 < \rho < 1$) produce damped dynamics that broaden the background without creating sustained rhythms. Thus the PSD compactly encodes the same structure as the AR dynamics: peaks $\leftrightarrow$ near–unit oscillatory modes, background $\leftrightarrow$ strictly stable poles. This link lets us use standard spectral tools while keeping a direct line to the geometric picture that follows in the manuscript.

2.6. Data–driven PSD for LFP

Empirically, LFP PSDs are well captured by a smooth aperiodic background noise plus oscillatory peaks. We adopt

$$S_{\mathrm{LFP}}(f) \;=\; A\big(f^2+f_k^2\big)^{-\kappa/2} \;+\; \sum_{j=1}^{J} B_j\,\frac{\gamma_j^2}{(f-f_j)^2+\gamma_j^2}\,, \tag{7}$$

where the first term is the background noise term and the sum collects J narrowband bumps at centers f_j with half–widths γ_j and heights B_j. This decomposition mirrors common LFP practice and will map cleanly to AR poles.

2.7. From the PSD to AR poles

Equation (7) admits a natural autoregressive realization that makes the underlying dynamics and geometry explicit. In particular, each oscillatory peak centered at (f_j, γ_j) can be represented by a conjugate pole pair

$$z_{j,\pm} = \rho_j\, e^{\pm i\omega_{0,j}}, \qquad \omega_{0,j} = \frac{2\pi f_j}{F_s}, \qquad \rho_j \approx e^{-\,2\pi\gamma_j/F_s} \;\; (\gamma_j \ll F_s),$$

which corresponds to an AR(2) factor

$$P_j(z^{-1}) \;=\; 1 - 2\rho_j \cos\omega_{0,j}\, z^{-1} + \rho_j^2 z^{-2}.$$

When $\rho_j \approx 1$, the mode is nearly undamped and yields a sustained narrowband oscillation (one circular phase in the manifold view).

The broadband *background noise* term $A\big(f^2+f_k^2\big)^{-\kappa/2}$ can be captured by a short product of real, strictly stable poles,

$$P_{\mathrm{bg}}(z^{-1}) \;\approx\; \prod_{r=1}^{m}\big(1-\rho_r z^{-1}\big), \qquad \rho_r \in (0,1), \;\; \rho_r \approx e^{-2\pi\lambda_r/F_s}, \tag{8}$$

with $\{\lambda_r\}$ chosen to match the location and slope of the aperiodic floor over the frequency band of interest. Because $|\rho_r| < 1$, these factors shape the broadband spectrum without introducing persistent oscillations.

One may realize the overall process either as a *parallel* sum,

$$x(t) = x_{\mathrm{bg}}(t) + \sum_{j=1}^{J} x_j(t) + \eta(t),$$

where each x_j and x_{bg} is generated by its own AR block (so the PSDs add exactly), or as a single all–pole model with

$$P_{\mathrm{total}}(z^{-1}) \;=\; P_{\mathrm{bg}}(z^{-1}) \prod_{j=1}^{J} P_j(z^{-1}), \qquad S_x(f) \;=\; \frac{\sigma_\varepsilon^2}{\big|P_{\mathrm{total}}(e^{-i2\pi f/F_s})\big|^2}. \tag{9}$$

In either formulation, the correspondence is direct: spectral peaks map to near–unit conjugate pole pairs (sustained rhythms), while the background noise maps to real subunit poles (aperiodic floor). This correspondence is useful because it preserves the intuitive spectral picture while furnishing a generative, pole–zero description that ties seamlessly into the dynamical and manifold analyses developed in the rest of the paper.

2.8. Manifold implications (data–driven dimension)

In our signal AR model poles with $|z| < 1$ are contracting and add no circle factors; each undamped conjugate pair $z = e^{\pm i\omega}$ $(0 < \omega < \pi)$ contributes one independent phase circle, so with two lags per undamped mode the noiseless lag state lies on a Cartesian product of circles (a K-torus) in $\mathbb{R}^{2K}$. Let K be the number of *distinct* undamped pairs (counted modulo π within channels) and let m be the number of within–channel coincidences; then

$$\dim \mathcal{M} = K_{\text{eff}} = K - m. \tag{10}$$

If $m = 0$, $\mathcal{M}$ is an embedded K-torus; otherwise it is a smooth immersed submanifold of dimension K_{eff}. Aperiodic/background components (subunit real poles) only thicken the cloud around $\mathcal{M}$ and do not alter its topology. This motivates using prediction error as a practical proxy for geometric thickness under colored (OU/Gauss–Markov) noise (Fallah, 2024). **Figure 2.** We use the *Lorenz* system as a controlled testbed to demonstrate our modeling paradigm on data with LFP-like structure. In **a**–**b**, a low-order AR model (AR(3)) achieves essentially perfect one-step prediction on held-out data (time-trace overlay and identity scatter). The power spectra in **c** display a broadband, approximately $1/f^{\kappa}$ decay with a shallow resonance—qualitatively similar to empirical LFP spectra—showing that the synthetic signal matches the spectral regime of interest. The order selection curve in **d** exhibits a clear elbow at $p = 3$, and the corresponding coefficients are shown in **e**. Crucially, when run freely, the fitted AR(3) reproduces the delay-embedded geometry of the underlying dynamics (**f**), indicating that the same parametric model captures both spectral content and global structure. Together, these panels support our claim that AR-based pole and manifold analysis is appropriate for LFP-like signals.

2.9. Betti numbers of K-torus manifold

Betti numbers give a compact, dimension-by-dimension summary of a space's "holes." Here b_0 counts connected components, b_1 counts independent loops, b_2 counts hollow surfaces, and higher b_j continue this pattern. In our AR setting, the noiseless lag–embedded state lies on a K-torus $\mathcal{M} \cong (S^1)^K$, whose Betti numbers follow a simple rule:

$$b_j\big((S^1)^K\big) = \binom{K}{j}, \qquad j = 0, \ldots, K.$$

Thus $b_0 = 1$ (one piece), $b_1 = K$ (one loop per undamped oscillation), $b_2 = \binom{K}{2}$, and so on (e.g., for K=1: $(1, 1)$; for K=2: $(1, 2, 1)$; for K=3: $(1, 3, 3, 1)$). Practically, these counts are intrinsic and robust: strictly stable (background) AR components may thicken the data cloud but do not create or destroy loops, so the Betti numbers—and hence the manifold's intrinsic structure—remain unchanged.

3. Discussion

In this work, we developed an analytical framework to characterize the manifold geometry of local field potentials. By modeling the LFP as an autoregressive process, we showed that its underlying noise-free structure is a K-torus. Crucially, we derived a closed-form expression

for the manifold's intrinsic dimension, demonstrating it is determined by the number of distinct oscillatory peaks in the signal. We established the computational feasibility of this approach on a synthetic benchmark, confirming that an AR model can effectively capture the local dynamics of a complex signal. This framework provides a new, theoretically grounded perspective on the geometric structure of collective neural activity.

A key conceptual implication of our work is the idea of the LFP manifold as a candidate "input manifold." Much of the existing literature has focused on manifolds derived from spiking activity, which are often interpreted as "output manifolds" that represent the computational state or behavioral intent of a neural circuit Vyas et al. (2020). In contrast, the LFP primarily reflects synaptic and dendritic activity, representing the inputs to a local population. The toroidal manifold we describe can therefore be seen as the geometric structure of the possible input states or contexts for a circuit. This provides a crucial missing piece of the puzzle, suggesting that the dynamics on output manifolds are constrained and guided by trajectories on these input manifolds, linking the geometry of inputs to the geometry of computation. (Gedankien et al., 2025)

This framework also offers a geometric re-interpretation of the role of neural oscillations. The rich literature on this topic has assigned functional roles to different frequency bands and their interactions Buzsáki and Draguhn (2004); Siegel et al. (2012). Our work provides a unifying geometric language: distinct, sustained oscillations are the fundamental coordinate axes of the LFP manifold. The presence of alpha, beta, and gamma rhythms, for example, would generate a 3-torus. Phenomena such as cross-frequency coupling, where the phase of a slow oscillation modulates the amplitude of a faster one, can be viewed as specific trajectories on this multi-dimensional torus. This perspective shifts the view of oscillations from being mere rhythms to being the basis vectors that define the dimensions of the neural input space.

Our results rely on a linear, stationary AR model; we allow slowly time-varying coefficients to capture mild nonstationarities, but abrupt regime changes remain challenging. Single-contact LFPs mix multiple sources Einevoll et al. (2013). To date our validation is synthetic; applying the framework to biological recordings is a necessary next step.

These limitations, however, point toward clear and exciting future directions. The immediate next step is to apply this framework to LFP data recorded from animals performing cognitive tasks. This would allow us to test specific, falsifiable predictions. For instance, we hypothesize that the dimensionality of the LFP manifold in the prefrontal cortex will increase during a working memory task that requires maintaining multiple distinct items, reflecting an expansion of the input state space. We further predict that attentional modulation may manifest as a change in the manifold's "thickness," where increased attention tightens trajectories around the core torus, corresponding to a reduction in the variance of the aperiodic signal components. By providing a bridge between the spectral properties of LFPs and the geometry of neural dynamics, this work lays the theoretical foundation for answering such questions.

References

Romain Brette. Is coding a relevant metaphor for the brain? *Behavioral and Brain Sciences*, 42, 2019.

György Buzsáki and Andreas Draguhn. Neuronal oscillations in cortical networks. *Science*, 304(5679):1926–1929, 2004.

György Buzsáki, Costas A Anastassiou, and Christof Koch. The origin of extracellular fields and currents—eeg, ecog, lfp and spikes. *Nature Reviews Neuroscience*, 13(6):407–420, 2012.

H. Choubdar, M. Mahdavi, Z. Rostami, E. Zabeh, M. J. Gillies, A. L. Green, T. Z. Aziz, and R. Lashgari. Neural oscillatory characteristics of feedback-associated activity in globus pallidus interna. *Scientific Reports*, 13(1):4141, March 2023. doi: 10.1038/s41598-023-30832-4.

SueYeon Chung and Larry F Abbott. Neural population geometry: An approach for understanding biological and artificial neural networks. *Current opinion in neurobiology*, 70: 137–144, 2021.

SueYeon Chung, Daniel D Lee, and Haim Sompolinsky. Linear readout of object manifolds. *Physical Review E*, 93(6):060301, 2016.

Uri Cohen, SueYeon Chung, Daniel D Lee, and Haim Sompolinsky. Separability and geometry of object manifolds in deep neural networks. *Nature communications*, 11(1):746, 2020.

John P Cunningham and Byron M Yu. Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*, 17(11):1500–1509, 2014.

Gaute T Einevoll, Alain Destexhe, Markus Diesmann, Sonja Grün, Viktor Jirsa, Marc de Kamps, Michele Migliore, Torbjørn V Ness, Hans E Plesser, and Anno Schüth. The scientific case for brain simulations. *Neuron*, 79(4):595–600, 2013.

Kasra Fallah. On the calculation of mutual information for channels with gauss-markov noise. Master's thesis, University of Waterloo, 2024. URL https://hdl.handle.net/10012/20776.

Zeinab Fazlali, Yadollah Ranjbar-Slamloo, Erfan Zabeh, and Ehsan Arabzadeh. Stimulation of Locus Coeruleus noradrenergic system modulates sensory processing and brain state in two different time scales. *bioRxiv*, page 2020.07.09.188615, 2020. doi: 10.1101/2020.07.09.188615. Preprint.

Peiran Gao and Surya Ganguli. On the theory of neural dimensionality, dynamics, and measurement. *Current Opinion in Neurobiology*, 46:203–213, 2017.

Crispin W Gardiner. *Handbook of Stochastic Methods*, volume 3. Springer, Berlin, 1985.

T. Gedankien, J. Kriegel, E. Zabeh, D. McDonagh, B. Lega, and J. Jacobs. Cholinergic blockade reveals role for human hippocampal theta in encoding but not retrieval. *bioRxiv*, page 2025.05.12.653487, May 2025. doi: 10.1101/2025.05.12.653487. Preprint.

M Kaminski and K J Blinowska. A method for the analysis of the transient brain responses based on the autoregressive moving average model. *Journal of Neuroscience Methods*, 38 (2-3):135–144, 1991.

Eliott Robert Joseph Levy, Simón Carrillo-Segura, Eun Hye Park, William Thomas Redman, José Rafael Hurtado, SueYeon Chung, and André Antonio Fenton. A manifold neural population code for space in hippocampal coactivity dynamics independent of place fields. *Cell reports*, 42(10), 2023.

Earl K Miller, Mikael Lundqvist, and André M Bastos. Working memory revived in working models. *Neuron*, 98(5):1012–1026, 2018.

Markus Siegel, Tobias H Donner, and Andreas K Engel. Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience*, 13(2):121–134, 2012.

Petre Stoica and Randolph L Moses. *Spectral Analysis of Signals*. Pearson Prentice Hall, 2005.

Saurabh Vyas, Matthew D Golub, David Sussillo, and Krishna V Shenoy. Computation through neural population dynamics. *Annual Review of Neuroscience*, 43:249–275, 2020.

Byron M Yu, John P Cunningham, Gopal Santhanam, Stephen I Ryu, Krishna V Shenoy, and Maneesh Sahani. Gaussian process factor analysis for low-dimensional single-trial analysis of neural population activity. In *Advances in Neural Information Processing Systems 22*, pages 1881–1889, 2009.

E. Zabeh, N. C. Foley, J. Jacobs, and J. P. Gottlieb. Beta traveling waves in monkey frontal and parietal areas encode recent reward history. *Nature Communications*, 14(1): 5428, September 2023. doi: 10.1038/s41467-023-41125-9.

Appendix A. Proofs

Lemma 1 (Block–diagonal AR(p) reduction) *Let $X(t) = [x_1(t), \dots, x_N(t)]^\top \in \mathbb{R}^N$ and suppose each component satisfies a autoregression*

$$x_i(t) \;=\; \sum_{m=1}^{p_i} \phi_{i,m}\, x_i(t-m) \;+\; \varepsilon_i(t), \qquad i = 1, \dots, N,$$

with orders $p_i \in \mathbb{N}$, coefficients $\{\phi_{i,m}\}$, and innovations $\varepsilon_i(t)$. Let $p := \max_i p_i$ and set $\phi_{i,m} := 0$ for all $m > p_i$ (zero padding). Define the diagonal lag blocks

$$\Phi_m \;:=\; \operatorname{diag}\big(\phi_{1,m}, \dots, \phi_{N,m}\big) \in \mathbb{R}^{N\times N}, \qquad m = 1, \dots, p. \tag{11}$$

Then the stacked process obeys the block–diagonal VAR(p)

$$X(t) \;=\; \sum_{m=1}^{p} \Phi_m\, X(t-m) \;+\; \varepsilon(t), \qquad \varepsilon(t) := [\varepsilon_1(t), \dots, \varepsilon_N(t)]^\top. \tag{12}$$

Proof The ith entry of $\sum_{m=1}^{p} \Phi_m X(t-m)$ equals $\sum_{m=1}^{p} \phi_{i,m} x_i(t-m)$, which, by zero padding, reduces to $\sum_{m=1}^{p_i} \phi_{i,m} x_i(t-m)$, i.e. the ith scalar AR. Stacking over i yields (12); conversely, (12) read componentwise recovers the scalar equations. ■

Lemma 2 (Companion form and equivalence to VAR(p)) *Let $X(t)$ satisfy (12). Define the stacked lag state (as in the main text, Eq. (3))*

$$s(t) \;=\; \big[\, X(t)^\top,\; X(t-1)^\top,\; \dots,\; X(t-p+1)^\top \,\big]^\top \in \mathbb{R}^{Np},$$

and the time–invariant companion matrices

$$\begin{aligned} A &= \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{p-1} & \Phi_p \\ I_N & 0 & \cdots & 0 & 0 \\ & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & I_N & 0 & 0 \end{bmatrix} \in \mathbb{R}^{Np\times Np}, \\ B &= \begin{bmatrix} I_N \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^{Np\times N}, \qquad C = \begin{bmatrix} I_N & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{N\times Np}. \end{aligned} \tag{13}$$

Then the state–space recursion in the main text (Eq. (4))

$$s(t{+}1) \;=\; A\, s(t) \;+\; B\, \varepsilon(t{+}1), \qquad X(t) \;=\; C\, s(t),$$

is algebraically equivalent to the VAR(p) model (12). Moreover, A admits the selection/shift decomposition

$$A \;=\; \mathcal{E}_1\, \mathcal{F} + \mathcal{J}, \qquad \mathcal{F} := \begin{bmatrix} \Phi_1 & \cdots & \Phi_p \end{bmatrix} \in \mathbb{R}^{N\times Np},\; \mathcal{E}_1 := \begin{bmatrix} I_N \\ 0 \\ \vdots \\ 0 \end{bmatrix},\; \mathcal{J} := \begin{bmatrix} 0 & 0 & \cdots & 0 \\ I_N & 0 & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & \cdots & I_N & 0 \end{bmatrix},$$

so that all learnable dynamics reside in the top block row $\mathcal{F}$, while $\mathcal{J}$ is a fixed lag shift.

Proof The top block of $As(t)$ equals $\sum_{m=1}^{p} \Phi_m X(t+1-m)$ and the top block of $B\,\varepsilon(t{+}1)$ is $\varepsilon(t{+}1)$, giving $X(t{+}1) = \sum_{m=1}^{p} \Phi_m X(t{+}1-m) + \varepsilon(t{+}1)$, i.e. (12). The lower blocks of A implement a down–shift so that the remaining entries of $s(t{+}1)$ are $X(t), X(t{-}1), \ldots, X(t{-}p{+}2)$. Conversely, reading off the first block of $s(t{+}1)$ from the state recursion recovers (12). The decomposition follows by inspection of the first block row and the fixed shift structure. ■

A.1. Spectral factorization and minimal-lag embedding

For channel $i \in \{1, \ldots, N\}$ with order p_i and AR coefficients $\{\phi_{i,m}\}_{m=1}^{p_i}$, define the AR operator z-form:

$$P_i\big(z^{-1}\big) \;=\; 1 - \sum_{m=1}^{p_i} \phi_{i,m}\, z^{-m}.$$

Roots are written as $z = \rho e^{\pm i\omega}$ with $\rho > 0$ and $\omega \in [0, \pi]$.

Lemma 3 (pole factorization) *For each channel i there exist an integer $K_i \geq 0$ and frequencies $\{\omega_{i,r}\}_{r=1}^{K_i} \subset (0, \pi)$ such that*

$$P_i\big(z^{-1}\big) \;=\; \left(\prod_{r=1}^{K_i} \left[\, 1 - 2\cos\omega_{i,r}\, z^{-1} + z^{-2} \,\right] \right) R_i\big(z^{-1}\big)\,,$$

where all roots of the residual factor $R_i\big(z^{-1}\big)$ lie strictly inside the open unit disk. Let $K := \sum_{i=1}^{N} K_i$ be the total number of unit-modulus conjugate pairs across channels.

Proof By the Fundamental Theorem of Algebra, $P_i\big(z^{-1}\big)$ factors over $\mathbb{C}$ as a product of terms $(1 - \zeta z^{-1})$, one for each root $z = \zeta$. Real coefficients imply complex roots occur in conjugate pairs $(\zeta, \bar{\zeta})$. Group every unit-modulus pair $\zeta = e^{i\omega}$, $0 < \omega < \pi$, into

$$(1 - \zeta z^{-1})(1 - \bar{\zeta} z^{-1}) = 1 - (\zeta + \bar{\zeta}) z^{-1} + z^{-2} = 1 - 2\cos\omega\, z^{-1} + z^{-2}.$$

Since the variance of the signal assumed to be bounded $|z| > 1$ and the real unit roots ± 1; all remaining roots obey $|z| < 1$ and contribute the strictly stable residual factor R_i. ■

Lemma 4 (Minimal-lag linearization and invertibility) *Suppose channel i admits K_i undamped factors as in Lemma **??** with distinct frequencies modulo π: $\omega_{i,r} \not\equiv \omega_{i,r'} \pmod{\pi}$ for $r \neq r'$. Consider the noiseless sum of these factors with amplitudes $A_{i,r} > 0$ and phases $\theta_{i,r} \in \mathbb{R}$,*

$$x_i(t) \;=\; \sum_{r=1}^{K_i} A_{i,r}\, \sin\theta_{i,r}(t), \qquad \theta_{i,r}(t+1) = \theta_{i,r}(t) + \omega_{i,r}.$$

Stack $2K_i$ consecutive lags $s^{(i)}(t) = [x_i(t), x_i(t-1), \ldots, x_i(t-2K_i+1)]^\top \in \mathbb{R}^{2K_i}$. Then there exist a real, square matrix $M_i \in \mathbb{R}^{2K_i \times 2K_i}$, a positive diagonal $D_i \in \mathbb{R}^{2K_i \times 2K_i}$, and the phase vector $u_i(\theta) = [\sin\theta_{i,1}, \cos\theta_{i,1}, \ldots, \sin\theta_{i,K_i}, \cos\theta_{i,K_i}]^\top$ such that

$$s^{(i)}(t) \;=\; M_i\, D_i\, u_i\big(\theta_i(t)\big), \qquad \textit{with} \qquad \det M_i \neq 0.$$

Hence two lags per undamped factor are sufficient and necessary in the AR setting.

Proof Using $\sin(\alpha-\beta)=\sin\alpha\cos\beta-\cos\alpha\sin\beta$,

$$x_i^{(r)}(t-k)=A_{i,r}\,\sin\big(\theta_{i,r}(t)-k\,\omega_{i,r}\big)=A_{i,r}\big(\sin\theta_{i,r}(t)\cos k\omega_{i,r}-\cos\theta_{i,r}(t)\sin k\omega_{i,r}\big).$$

Thus each lag is linear in $(\sin\theta_{i,r},\cos\theta_{i,r})$ with coefficients $\cos k\omega_{i,r}$ and $-\sin k\omega_{i,r}$. Stacking $k=0,\dots,2K_i-1$ over r yields $s^{(i)}(t)=M_iD_iu_i(\theta_i(t))$ with the advertised block real Vandermonde M_i. Distinctness modulo π gives $2K_i$ distinct nodes on the unit circle, so M_i is full rank (a realified Vandermonde argument), and D_i is invertible. ■

Global minimal-lag embedding. Stack the per-channel vectors $s^{(i)}(t)$ to form

$$s_{\min}(t)\;=\;\big[s^{(1)}(t)^\top,\dots,s^{(N)}(t)^\top\big]^\top\in\mathbb{R}^{2K},\tag{14}$$

and define

$$M:=\mathrm{blkdiag}\big(M_1,\dots,M_N\big),\qquad D:=\mathrm{blkdiag}\big(D_1,\dots,D_N\big),\qquad u(\theta):=\big[u_1(\theta_1)^\top,\dots,u_N(\theta_N)^\top\big]^\top.$$

Then

$$s_{\min}(t)\;=\;M\,D\,u\big(\theta(t)\big).\tag{15}$$

Proposition 5 (K-torus) *Assume the hypotheses of Lemma 4 hold for all channels, so that M is invertible. Define $F:\mathbb{T}^K\to\mathbb{R}^{2K}$ by*

$$F(\theta)\;:=\;M\,D\,u(\theta),\tag{16}$$

where $u(\theta)=[\sin\theta_1,\cos\theta_1,\dots,\sin\theta_K,\cos\theta_K]^\top$ collects all K phases. Then F is an embedding and $\mathcal{M}:=F(\mathbb{T}^K)$ is a smooth, compact manifold diffeomorphic to the K-torus $(S^1)^K$. Equivalently, in whitened coordinates $y=W\,s_{\min}(t)$ with $W:=(MD)^{-1}$,

$$\mathcal{M}\;=\;\Big\{\,s\in\mathbb{R}^{2K}\;:\;\big(y_{2q-1}\big)^2+\big(y_{2q}\big)^2=1,\;q=1,\dots,K,\quad y=Ws\,\Big\}.\tag{17}$$

Proof $u:\mathbb{T}^K\to\mathbb{R}^{2K}$ is smooth and MD is an invertible linear map, so $F=MD\circ u$ is smooth. The Jacobian satisfies $\mathrm{D}F(\theta)=MD\,\mathrm{D}u(\theta)$ with $\mathrm{D}u(\theta)=\mathrm{blkdiag}\big(J(\theta_1),\dots,J(\theta_K)\big)$ and $J(\vartheta)=\begin{bmatrix}\cos\vartheta\\-\sin\vartheta\end{bmatrix}$, hence $\operatorname{rank}\mathrm{D}F(\theta)=K$ for all θ. If $F(\theta)=F(\theta')$, applying $W=(MD)^{-1}$ gives $u(\theta)=u(\theta')$, so each $(\sin\theta_q,\cos\theta_q)=(\sin\theta'_q,\cos\theta'_q)$ and thus $\theta_q\equiv\theta'_q\pmod{2\pi}$. Therefore F is injective. A continuous injective map from compact $\mathbb{T}^K$ into Hausdorff $\mathbb{R}^{2K}$ is a homeomorphism onto its image; combined with immersion, F is an embedding. In $y=Ws$ coordinates, each pair (y_{2q-1},y_{2q}) lies on the unit circle, yielding (17) and the torus structure. ■